### University of Nebraska - Lincoln

### DigitalCommons@University of Nebraska - Lincoln

Virology Papers

Virology, Nebraska Center for

2-20-2007

## Supplementary Data for "Sequence and annotation of the 314-kb MT325 and the 321-kb FR483 viruses that infect *Chlorella* Pbi": Appendix B: Gene Names M001L through M807R

Lisa A. Fitzgerald University of Nebraska-Lincoln, lisa.fitzgerald@nrl.navy.mil

Michael V. Graves University of Massachusetts-Lowell, Michael\_Graves@uml.edu

Xiao Li University of Massachusetts-Lowell

Tamara Feldblyum The Institute for Genomic Research, Rockville, MD

James Hartigan Agencourt Bioscience Corporation, Beverly, MA

See next page for additional authors

Follow this and additional works at: https://digitalcommons.unl.edu/virologypub



Part of the Virology Commons

Fitzgerald, Lisa A.; Graves, Michael V.; Li, Xiao; Feldblyum, Tamara; Hartigan, James; and Van Etten, James L., "Supplementary Data for "Sequence and annotation of the 314-kb MT325 and the 321-kb FR483 viruses that infect Chlorella Pbi": Appendix B: Gene Names M001L through M807R" (2007). Virology Papers. 3.

https://digitalcommons.unl.edu/virologypub/3

This Article is brought to you for free and open access by the Virology, Nebraska Center for at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Virology Papers by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

| <b>Authors</b><br>Lisa A. Fitzgerald, Michael V. Graves, Xiao Li, Tamara Feldblyum, James Hartigan, and James L. Van Etten |
|----------------------------------------------------------------------------------------------------------------------------|
|                                                                                                                            |
|                                                                                                                            |
|                                                                                                                            |
|                                                                                                                            |
|                                                                                                                            |
|                                                                                                                            |
|                                                                                                                            |
|                                                                                                                            |
|                                                                                                                            |
|                                                                                                                            |
|                                                                                                                            |
|                                                                                                                            |
|                                                                                                                            |
|                                                                                                                            |

#### SUPPLEMENTARY DATA FOR

# Sequence and annotation of the 314-kb MT325 and the 321-kb FR483 viruses that infect *Chlorella* Pbi

Lisa A. Fitzgerald<sup>a</sup>, Michael V. Graves<sup>b</sup>, Xiao Li<sup>b</sup>, Tamara Feldblyum<sup>c</sup>, James Hartigan<sup>d</sup>, and James L. Van Etten<sup>e, f, \*</sup>

**Abstract:** Viruses MT325 and FR483, members of the family Phycodnaviridae, genus *Chlorovirus*, infect the fresh water, unicellular, eukaryotic, chlorella-like green alga, *Chlorella* Pbi. The 314,335-bp genome of MT325 and the 321,240-bp genome of FR483 are the first viruses that infect *Chlorella* Pbi to have their genomes sequenced and annotated. Furthermore, these genomes are the two smallest chlorella virus genomes sequenced to date, MT325 has 331 putative protein-encoding and 10 tRNA-encoding genes and FR483 has 335 putative protein-encoding and 9 tRNA-encoding genes. The protein-encoding genes are almost evenly distributed on both strands, and intergenic space is minimal. Approximately 40% of the viral gene products resemble entries in public databases, including some that are the first of their kind to be detected in a virus. For example, these unique gene products include an aquaglyceroporin in MT325, a potassium ion transporter protein and an alkyl sulfatase in FR483, and a dTDP–glucose pyrophosphorylase in both viruses. Comparison of MT325 and FR483 protein-encoding genes with the prototype chlorella virus PBCV-1 indicates that approximately 82% of the genes are present in all three viruses.

Supplementary data associated with this article is archived in this repository as 4 separate files: Appendices A–D. Each document, in spreadsheet format, shows Gene Name, Genome Position, A.A. length, Peptid e Mw, pI, CDD Hit Number, COGs, COG Definition, Bit Score, E-value, % Identity, % Positive, Query from-to, Hit from-to, BLASTp Hit Number, Hit Accession, BLASTp Definition, Bit Score, E-value, % Identity, % Positive, Query from-to, and Hit from-to.

Appendix A: Gene Names m002R through m843L Appendix B: Gene Names M001L through M807R Appendix C: Gene Names n001L through n849R Appendix D: Gene Names N003L through N847R

<sup>&</sup>lt;sup>a</sup>Department of Chemistry, University of Nebraska–Lincoln, Lincoln, NE 68588-0304

<sup>&</sup>lt;sup>b</sup>Department of Biological Sciences, University of Massachusetts-Lowell, Lowell, MA 01854

<sup>&</sup>lt;sup>c</sup>The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, MD 20850

<sup>&</sup>lt;sup>d</sup>Agencourt Bioscience Corporation, 500 Cummings Center, Suite 2450, Beverly, MA 01915

<sup>&</sup>lt;sup>e</sup>Department of Plant Pathology, University of Nebraska–Lincoln, Lincoln, NE 68583-0722

<sup>&</sup>lt;sup>f</sup>Nebraska Center for Virology, University of Nebraska, Lincoln, NE 68588-0666

<sup>\*</sup>Corresponding author. Email: jvanetten@unlnotes.unl.edu

| Gene<br>Name | Genome<br>Position | A.A.<br>length | Peptide<br>Mw | pl    | CDD Hit<br>Number | COGs         | COG Definition                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | Bit<br>Score | E-value   | %<br>Identity P | %<br>ositive | Query<br>from-to | Hit<br>from-to | BLASTp<br>Hit<br>Number | Hit<br>Accession                                                | BLASTp Definition                                                                   | Bit<br>Score     | E-value Id                          | %<br>lentity Po   | %<br>ositive      | Query<br>from-to         | Hit from-<br>to               |
|--------------|--------------------|----------------|---------------|-------|-------------------|--------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|-----------|-----------------|--------------|------------------|----------------|-------------------------|-----------------------------------------------------------------|-------------------------------------------------------------------------------------|------------------|-------------------------------------|-------------------|-------------------|--------------------------|-------------------------------|
| M001L        | 1052-384           | 223            | 26,253        | 6.70  |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |           |                 |              |                  |                | 1<br>2<br>3             |                                                                 | intigen                                                                             |                  | 1.41E-06<br>2.03E-05<br>7.73E-05    | 25%<br>24%<br>27% | 59%<br>51%<br>54% | 14113<br>1136 1<br>27137 | 61–160<br>1993–2129<br>91–194 |
| M003L        | 15211147           | 125            | 13,988        | 9.87  |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |           |                 |              |                  |                | 3                       | No Hit Found No Hit Found                                       | GF 110                                                                              | 49.00            | 7.732-03                            | 2170              | 5470              | 21-131                   | 51-154                        |
| M005L        | 22941755           | 180            | 20,424        | 10.16 |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |           |                 |              |                  |                | 1                       | NP_048429 A81L                                                  |                                                                                     | 110.92           | 1.71E-23                            | 40%               | 60%               | 33158                    | 21-163                        |
| M007L        | 28392330           | 170            | 19,447        | 6.12  |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |           |                 |              |                  |                | 1                       | NP 048432 A84L                                                  |                                                                                     | 73.94            | 1.97E-12                            | 33%               | 57%               | 11135                    | 10-149                        |
| M009L        | 37362930           | 269            | 31,504        | 6.02  | 1                 | pfam03016    | Exostosin, Exostosin family. The EXT family is a family of tumour suppressor genes. Mutations of EXT1 on 8q24.1, EXT2 on 11p11-13, and EXT3 on 19 have been associated with the autosomal dominant disorder known as hereditary multiple exostoses (HME). This is the most common known skeletial dysplasia. The chromosomal coations of other EXT genes suggest association with other forms of neoplasia. EXT1 and EXT2 have both been shown to nordoe a heparan sulphate polymerase with both D-glucuronyl (GloA) and N-acetyl-D-glucosaminoglycan (GloKAO) transferase activities. The nature of the defect in heparan                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | 45.82        | 4.20E-06  | 25%             | 51%          | 159235           | 5 221–292      | 1                       | NP_048423 A75L                                                  |                                                                                     | 233.42           | 5.34E-60                            | 40%               | 61%               | 8-268                    | 7276                          |
| M010L        | 41463790           | 119            | 13,199        | 9.97  | 1                 | cd01285      | subnate hisconthesis in HME is unclear nucleoside, dearniases Nucleoside dearniases include adenosine, quantine and optication dearniases. These enzymes are Zn deportedrat and catalyze the dearniasions. These enzymes are Zn deportedrat and catalyze the dearniasion of nucleosidists. The zno on in the active site plays a central role in the proposed catalytic mechanism, activating a water molecule to form a hydroxide on that performs a nucleophitic attack on the substrate. The functional enzyme is a homodimer. Cybosine dearniases catalyzes the dearnianism of cybosine to enazio and ammonia and is a member of the pyrimidine salvage pathway. Cybosine dearniases is found in bacteria and fungi but is not present in mammals; for this reason, the enzyme is currently of interest for antimicrobial drug design and gene therapy applications against tumors. Some emembers of this family are IRNA-specific adenosine dearniases that generate inosine at the first position of their anticolon opeosition 30 of specific RNAss, this modification is thought the enlarge the color mecognition capacity during such of Assemblace causines in ventifica as a reaf of the utilization of registrate | 57.97        | 9.98E-10  | 35%             | 51%          | 4106             | 3 1–94         | 1                       | NP_048547 contains cytidine ar<br>signature                     | nd deoxycytidine deaminase Zn-binding region                                        | 161.38           | 7.13E-39                            | 62%               | 83%               | 1–118                    | 1–118                         |
|              |                    |                |               |       | 2                 | COG0590      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 52.65        | 3.33E-08  | 31%             | 47%          | 2106             | 10105          | 2                       | AAR26853 FirrV-1-A29                                            |                                                                                     | 50.83            | 1.36E-05                            | 27%               | 50%               | 22111                    | 24105                         |
|              |                    |                |               |       | 3                 | pfam00383    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 43.43        | 2.04E-05  | 30%             | 47%          | 4106             | 7101           | 3                       | AAX51127 tRNA-specific adenos                                   | ine deaminase                                                                       | 48.52            | 6.74E-05                            | 31%               | 49%               | 4108                     | 10-104                        |
|              |                    |                |               |       | 4                 | cd01284      | Riboflawin, deaminase-veucutase. Riboflawin-specific deaminase.<br>Riboflawin biospirhesis protein Ribi (Diaminohydroxyphosphonbosylaminopyrimidine deaminase) catalyzes<br>the deamination of 2.5-diamino-6-thosylamino-(49/h)-pyrimidinone<br>5-'_phosphate, which is an intermediate step in the biosymhesis of<br>riboflawin. The ribi opene of Bacillaw subtilis and the ribi open of E. coli<br>comain which catalyzes the subsequent reduction of the ribosyl side<br>chain                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 36.76        | 0.002101  | 26%             | 45%          | 4107             | 7 1–95         | 4                       | AAC68441 cytosine deaminase                                     |                                                                                     | 48.14            | 8.80E-05                            | 31%               | 49%               | 4108                     | 10104                         |
| M011L        | 45244237           | 96             | 10,637        | 10.49 |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |           |                 |              |                  |                | 1                       | NP_048546 A199R                                                 |                                                                                     | 73.17            | 2.59E-12                            | 43%               | 68%               | 179                      | 182                           |
| M012R        | 45525019           | 156            | 17,894        | 7.66  |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |           |                 |              |                  |                | 1                       | NP_048543 A196L                                                 |                                                                                     | 185.27           | 4.65E-46                            | 56%               | 72%               | 1151                     | 1151                          |
|              |                    |                |               |       |                   |              | PCNA N, Proliferating cell nuclear antigen, N-terminal domain. N-                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |              |           |                 |              |                  |                |                         | _                                                               |                                                                                     |                  |                                     |                   |                   |                          |                               |
| M014R        | 50765867           | 264            | 29,728        | 5.00  | 1                 |              | terminal and C-terminal domains of PCNA are topologically identical. Three PCNA molecules are tightly associated to form a closed ring encircling duplex DNA. PCNA_C, Proliferating cell nuclear antigen, C-terminal domain.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |              | 4.08E-20  | 28%             | 57%          |                  | 2 1-125        |                         |                                                                 | NA, corresponds to Swiss-Prot Accession Number                                      |                  | 1.34E-100                           | 68%               | 85%               | 9261                     | 7259                          |
|              |                    |                |               |       | 2                 |              | Three PCNA molecules are tightly associated to form a closed ring                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | 79.98        | 2.27E-16  | 34%             | 55%          | 137261           | 2127           | 2                       | XP_534355 PREDICTED: similar                                    | to proliferating cell nuclear antigen                                               | 158.69           | 1.61E-37                            | 32%               | 55%               | 5-264                    | 202461                        |
|              |                    |                |               |       | 3                 | COG0592      | DnaN, DNA polymerase sliding clamp subunit (PCNA homolog) [DNA replication. recombination. and repair].                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | 56.08        | 3.34E-09  | 19%             | 37%          | 2126             | 72321          | 3                       |                                                                 | ar antigen                                                                          | 157.53           | 3.59E-37                            | 34%               | 56%               | 9264                     | 1256                          |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |           |                 |              |                  |                | 4<br>5                  | NR 172217 PCNA1 (PROLIFER                                       | ATING CELLULAR NUCLEAR ANTIGEN); DNA                                                | 156.76<br>156.76 | 6.13E-37<br>6.13E-37                | 32%<br>32%        | 57%<br>57%        | 9264<br>9264             | 1256<br>1256                  |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |           |                 |              |                  |                | 6                       | AAG10077 proliferating cell nucle                               | ar antigen                                                                          | 155.99           | 1.05E-36                            | 32%               | 55%               | 9264                     | 1256                          |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |           |                 |              |                  |                | 7 8                     | CAA37243 unnamed protein prod<br>BAB28355 unnamed protein prod  | luct<br>luct                                                                        | 155.61<br>155.61 | 1.37E-36<br>1.37E-36                | 32%<br>32%        | 55%<br>55%        | 9-264<br>9-264           | 1256<br>1256                  |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |           |                 |              |                  |                | 9                       |                                                                 | ar antigen (DNA polymerase delta auxiliary protein)                                 | 155.61           | 1.37E-36                            | 32%               | 55%               | 9-264                    | 1256                          |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |           |                 |              |                  |                | 10                      | AAD10528 proliferating cell nucle                               | ar antigen                                                                          | 155.22           | 1.78E-36                            | 34%               | 56%               | 9261                     | 1253                          |
| M015L        | 97345877           | 1286           | 141,489       | 11.47 | 1                 |              | DUF874, Helicobacter pylori protein of unknown function (DUF874). This<br>7 family consists of several hypothetical proteins specific to Helicobacter<br>nylori. The function of this family is unknown<br>RPR9, US snRNP splice                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              | 0.000259  | 18%             |              | 10941283         |                | 1                       |                                                                 | aromatin remodeling complex subunit OSA2                                            |                  | 0.00E+00                            | 39%               | 51%               | 21171                    |                               |
|              |                    |                |               |       | 2                 |              | modification1.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |              | 0.000753  | 26%             |              | 12081285         |                | 2                       |                                                                 | 192R - Chlorella virus PBCV-1                                                       |                  | 1.96E-13                            | 30%               |                   | 9811171                  | 2200                          |
|              |                    |                |               |       | 3                 |              | and repairl.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | 37.34        | 0.001476  | 15%             | 38%          | 10371278         |                | 3                       |                                                                 |                                                                                     | 67.78            | 2.93E-09                            | 24%               | 31%               |                          | 6661122                       |
|              |                    |                |               |       | 4                 | COG3206      | biosvnthesis (Cell envelope biogenesis, outer membrane).<br>SCP-1, Synaptonemal complex protein 1 (SCP-1). Synaptonemal<br>complex protein 1 (SCP-1) is the major component of the transverse                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | 35.44        |           | 12%             |              | 10551286         |                | 4                       |                                                                 | gen                                                                                 |                  | 2.93E-09                            | 25%               | 32%               | 106493                   |                               |
|              |                    |                |               |       | 5                 | ptam05483    | 3 filaments of the synaptonemal complex. Synaptonemal complexes are<br>structures that are formed between homologous chromosomes during<br>meiatric prophase                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | 35.17        | 0.006547  | 24%             | 44%          | 10481212         | 2 600-772      | 5                       |                                                                 |                                                                                     |                  | 9.42E-08                            | 25%               | 34%               | 120468                   | 161-492                       |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |           |                 |              |                  |                | 6                       | T17681 hypothetical protein a<br>YP 783728 PREDICTED: similar t | 191R - Chlorella virus PBCV-1<br>to Protein transport protein Sec24C (SEC24-related | 62.00<br>59.31   | 1.61E-07<br>1.04E-06                | 46%<br>23%        | 72%<br>34%        | 918977<br>207494         | 263<br>6311                   |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |           |                 |              |                  |                | 8                       | YP 394285 DDEDICTED similar                                     | to GA11046 PA                                                                       | 58.92            | 1.36E-06                            | 28%               | 36%               | 209477                   | 124394                        |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |           |                 |              |                  |                |                         | XP_697998 PREDICTED: similar t                                  | o AT rich interactive domain 1B (SWI1-like) isoform                                 |                  | 3.96E-06                            | 25%               | 35%               | 144460                   | 506844                        |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |           |                 |              |                  |                |                         | NP_571089 calymmin                                              |                                                                                     |                  | 5.17E-06                            | 25%               | 33%               | 204479                   | 282-573                       |
| M019L        | 124899769          | 907            | 102,974       | 7.91  | 1                 |              | recombination, and repair).  DNA pol B, DNA polymerase family B. This region of DNA polymerase                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |              | 9.23E-100 | 30%             | 48%          |                  | 11772          |                         | NP_048532 PBVC-1 DNA polyme                                     | rase                                                                                |                  | 0.00E+00                            | 72%               | 84%               | 1903                     | 1913                          |
|              |                    |                |               |       | 2                 |              | B appears to consist of more than one structural domain, possibly including elongation, DNA-binding and dNTP binding activities  POLBC, DNA polymerase type-B family: DNA polymerase alpha delta                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 336.58       | 1.19E-93  | 38%             | 55%          | 427850           | 1439           | 2                       | BAA35142 DNA polymerase                                         |                                                                                     | 1348.18          | 0.00E+00                            | 72%               | 83%               | 1-903                    | 1913                          |
|              |                    |                |               |       | 3                 | smart00486   | g epsilon and zeta chain (eukaryota), DNA polymerases in archaea, DNA<br>polymerase II in e. coli, mitochondrial DNA polymerases and and virus<br>DNA polymerases.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 306.37       | 1.55E-84  | 33%             | 51%          | 176638           | 3 1475         | 3                       | P30320 DNA polymerase                                           |                                                                                     | 1342.41          | 0.00E+00                            | 71%               | 83%               | 1-903                    | 1913                          |
|              |                    |                |               |       | 4                 | cd00145      | POLBc, DNA polymerase type-B family; DNA directed DNA polymerase. 5 Posseses DNA binding, polymerase and 3'-5', exonuclease activity.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 283.74       | 9.61E-78  | 33%             | 49%          | 176675           | 5 1511         | 4                       | AAB49748 DNA polymerase                                         |                                                                                     | 443.74           | 1.33E-122                           | 96%               | 97%               | 460689                   | 1230                          |
|              |                    |                |               |       | 5                 | pfam03104    | DNA_pol_B_exo, DNA polymerase family B, exonuclease domain. This<br>4 domain has 3' to 5' exonuclease activity and adopts a<br>ribonuclease H type fold                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | 194.94       | 4.75E-51  | 26%             | 41%          | 27354            | 1334           | 5                       | AAK28935 DNA polymerase                                         |                                                                                     | 431.80           | 5.23E-119                           | 100%              | 100%              | 468682                   | 1215                          |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |           |                 |              |                  |                | 6<br>7                  | AAK28933 DNA polymerase<br>AAX86472 DNA polymerase              |                                                                                     |                  | 3.39E-118<br>3.39E-118              | 99%<br>99%        | 100%<br>100%      | 468682<br>469682         | 1215<br>1214                  |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |           |                 |              |                  |                | 8<br>9<br>10            | AAK28936 DNA polymerase<br>XP_757605 hypothetical protein U     |                                                                                     | 427.56<br>412.15 | 9.86E-118<br>4.29E-113<br>3.07E-111 | 99%<br>32%<br>32% | 99%<br>49%<br>52% | 468682<br>17851<br>17863 | 1215<br>83929                 |
| M023L        | 12664-12446        | 73             | 8,575         | 8.65  |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |           |                 |              |                  |                |                         | No Hit Found No Hit Found                                       |                                                                                     |                  |                                     |                   |                   |                          |                               |

| Gene<br>Name | Genome<br>Position | A.A.<br>length | Peptide<br>Mw | pl    | CDD Hit<br>Number | COGs         | COG Definition                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | Bit<br>Score   | E-value              | %<br>Identity P | %<br>ositive | Query<br>from-to | Hit<br>from-to | BLASTp<br>Hit<br>Number | Hit<br>Accession                           | n BLASTp Definition                                                                                                                                                                                                                 | Bit<br>Score                                             | E-value I                                                            | %<br>dentity P                         | %<br>ositive                           | Query I<br>from-to                                       | Hit from-<br>to                                     |
|--------------|--------------------|----------------|---------------|-------|-------------------|--------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------|----------------------|-----------------|--------------|------------------|----------------|-------------------------|--------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------|----------------------------------------------------------------------|----------------------------------------|----------------------------------------|----------------------------------------------------------|-----------------------------------------------------|
| M024L        | 1288512685         | 67             | 7,956         | 9.03  |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                |                      |                 |              |                  |                |                         | No Hit Foun                                | d No Hit Found                                                                                                                                                                                                                      |                                                          |                                                                      |                                        |                                        |                                                          |                                                     |
| M025R        | 1280813545         | 246            | 27,971        | 7.96  |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                |                      |                 |              |                  |                | 1<br>2<br>3             | NP_04900                                   |                                                                                                                                                                                                                                     | 137.12<br>135.58                                         | 2.33E-56<br>4.42E-31<br>1.29E-30                                     | 57%<br>34%<br>32%                      | 79%<br>56%<br>51%                      | 73243<br>1227<br>2242                                    | 3173<br>4227<br>8253                                |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                |                      |                 |              |                  |                | 4                       |                                            | similar to PBCV-1 ORF A79R, corresponds to GenBank Accession                                                                                                                                                                        | 133.27                                                   | 6.39E-30                                                             | 31%                                    | 52%                                    | 1-244                                                    | 1248                                                |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                |                      |                 |              |                  |                | 6                       | NP_04842                                   | Number 017055<br>7 A79R                                                                                                                                                                                                             | 131.72<br>120.94                                         | 1.86E-29<br>3.28E-26                                                 | 31%<br>31%                             | 52%<br>52%                             | 1-241<br>1-221                                           | 4248<br>1218                                        |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                |                      |                 |              |                  |                | 7                       |                                            | 4 hypothetical protein A275R<br>1 hypothetical protein A275R                                                                                                                                                                        | 81.65<br>80.49                                           | 2.21E-14<br>4.91E-14                                                 | 31%<br>30%                             | 52%<br>52%                             | 78241<br>78241                                           | 4167<br>4167                                        |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                |                      |                 |              |                  |                | 10                      | AAU0630<br>NP_04900                        | hypothetical protein A275R     similar to Chlorella virus PBCV-1 ORF A450R, corresponds to GenBank Accession Number U42580                                                                                                          | 67.01<br>57.00                                           | 5.62E-10<br>5.82E-07                                                 | 32%<br>46%                             | 54%<br>68%                             | 110242<br>147                                            | 1132<br>147                                         |
| M026L        | 1462813591         | 346            | 38,210        | 7.11  | 1                 |              | LdhA, Lactate dehydrogenase and related dehydrogenases [Energy production and conversion / Coenzyme metabolism / General function prediction onlvl.  SerA, Phosphoglycerate dehydrogenase and related dehydrogenases                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |                | 7.39E-68             |                 | 57%          |                  | 3321           | 1                       |                                            | contains D-isomer specific 2-hydroxyacid dehydrogenase signature;<br>1 similar to E. coli D-lactate dehydrogenase, corresponds to Swiss-Prot<br>Accession Number P52643                                                             |                                                          | 3.41E-127                                                            | 67%                                    | 80%                                    | 9345                                                     | 25-361                                              |
|              |                    |                |               |       | 2                 | COG0111      | [Amino acid transport and metaholism]  2-Hacid_dh_C, D-isomer specific 2-hydroxyacid dehydrogenase, NAD                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 189.38         | 2.72E-49             | 35%             | 56%          | 87333            | 74303          | 2                       | XP_36314                                   | 1 hypothetical protein MG08725.4                                                                                                                                                                                                    | 240.35                                                   | 6.56E-62                                                             | 38%                                    | 57%                                    | 16345                                                    | 3332                                                |
|              |                    |                |               |       | 3                 | pfam02826    | large denydrogenase and D-lactate denydrogenase families in SCOP. Neterminal portion of which is represented by family ofam00389  2-Hacid_dh, D-isomer specific 2-hydroxyacid dehydrogenase, catalytic                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | 168.10         | 7.14E-43             | 39%             | 64%          | 116312           | 1184           | 3                       | CAE8193                                    | 7 related to D-lactate dehydrogenase                                                                                                                                                                                                | 228.79                                                   | 1.97E-58                                                             | 38%                                    | 56%                                    | 16332                                                    | 3-333                                               |
|              |                    |                |               |       | 4                 | pfam00389    | domain. This family represents the largest portion of the catalytic domain of 2-hydroxyacid dehydrogenases as the NAD binding domain is inserted                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | 55.65          | 4.81E-09             | 26%             | 45%          | 16110            | 195            | 4                       | BAE5910                                    | 9 unnamed protein product                                                                                                                                                                                                           | 228.02                                                   | 3.37E-58                                                             | 40%                                    | 56%                                    | 16332                                                    | 3317                                                |
|              |                    |                |               |       |                   |              | within the structural domain.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |                |                      |                 |              |                  |                | 5<br>6<br>7<br>8<br>9   | AAL2056<br>YP 21662<br>YP_40811<br>CAG1962 | D-lactate dehydrogenase     NAD-dependent fermentative D-lactate dehydrogenase     D-lactate dehydrogenase     Sementative D-lactate dehydrogenase, NAD-dependent     P-turtive D-Lactate dehydrogenase     D-lactate dehydrogenase | 226.87<br>226.48<br>224.94<br>223.02<br>222.63<br>222.25 | 7.50E-58<br>9.80E-58<br>2.85E-57<br>1.08E-56<br>1.41E-56<br>1.85E-56 | 40%<br>39%<br>39%<br>39%<br>39%<br>40% | 57%<br>57%<br>57%<br>57%<br>55%<br>55% | 16-332<br>16-332<br>16-332<br>16-332<br>16-345<br>16-332 | 3-315<br>3-315<br>3-315<br>3-315<br>13-338<br>3-317 |
|              |                    |                |               |       |                   |              | MIP, Major intrinsic protein (MIP) superfamily. Members of the MIP superfamily function as membrane channels that selectively transport water, small neutral molecules, and ions out of and between cells. The                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |                |                      |                 |              |                  |                |                         |                                            |                                                                                                                                                                                                                                     |                                                          |                                                                      |                                        |                                        |                                                          |                                                     |
| M030R        | 1479815607         | 270            | 29,491        | 7.20  | 1                 | cd00333      | channel proteins share a common fold: the N-terminal cytosolic portion followed by six transmembrane helices, which might have arisen through gene duplication. On the basis of sequence similarity and functional characteristics, the superfamily can be subdivided into two major groups: water-selective channels called aquaporins (AQPs) and glycerol uptake facilitators (GIPs). AQPs are found in all three kingdoms of life, while                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | 103.35         | 2.08E-23             | 26%             | 45%          | 9262             | 3228           | 1                       | ABA4076                                    | 3 aquaglyceroporin                                                                                                                                                                                                                  | 463.38                                                   | 3.20E-129                                                            | 88%                                    | 88%                                    | 1270                                                     | 1270                                                |
|              |                    |                |               |       |                   |              | GlpFs have been characterized only within microorganisms                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                |                      |                 |              |                  |                | 2                       |                                            | Slycerol uptake facilitator protein, GLPF     glycerol uptake facilitator protein                                                                                                                                                   |                                                          | 3.88E-34<br>9.90E-30                                                 | 34%<br>33%                             | 51%<br>48%                             | 6261<br>12261                                            | 1231<br>12237                                       |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                |                      |                 |              |                  |                |                         | ZP 0090848                                 |                                                                                                                                                                                                                                     | 132.49                                                   | 1.29E-29<br>1.09E-28                                                 | 32%<br>34%                             | 49%                                    | 6-256                                                    | 1227                                                |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                |                      |                 |              |                  |                | 6                       | BAD6591                                    | 5 glycerol uptake facilitator<br>2 Major intrinsic protein                                                                                                                                                                          | 128.26                                                   | 2.44E-28<br>4.16E-28                                                 | 31%<br>34%                             | 47%<br>47%                             | 11261<br>9261                                            | 6232<br>4231                                        |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                |                      |                 |              |                  |                | 8<br>9<br>10            | BAB8226<br>NP 78256                        | probable qiyoerol uptake facilitator protein     qiyoerol uptake facilitator protein     qiyoerol uptake facilitator protein                                                                                                        | 127.10<br>125.56<br>125.56                               | 5.43E-28<br>1.58E-27<br>1.58E-27                                     | 34%<br>30%<br>32%                      | 47%<br>46%<br>48%                      | 9259<br>6261<br>9261                                     | 5231<br>1232<br>4234                                |
| M032L        | 16521–15835        | 229            | 26,057        | 10.08 | 1                 | cd00283      | GIY-YIG-Clerm, GIYX(10-11)YIG family of class I homing endonucleases promote the mobility of intron or inteln by recognizing and cleaving a homologous allele that lacks the sequence. They catalyze a double-strand break in the DNA near the insertion site of that element to facilitate homing at that site. Class I homing endonucleases are sorted into four families based on the presence of these motifs in their respective. Nor insertion site of the clement to activate the minimization of the control of the CIV-YIG family. The Chemminus of the control of the CIV-YIG family. The Chemminus of a terminal by a long, flexible linker. The DNA-briding domain consists of a minor-groove binding alpha-helix, and a helix-furni-helix. Some also contain a zize finger (e. I. Felly which is not required for DNA briding or calabysis, but is a component of the linker and directs the catalytic domain to cleave the homings let a la fixed distance from the introl ordamin to cleave the homings let a la fixed distance from the introl ordamin to cleave the homings let a la fixed distance from the introl ordamin to cleave the homings let a la fixed distance from the introl ordamin to cleave the homings let a la fixed distance from the introl | 70.03          | 1.93E-13             | 42%             | 60%          | 118225           | 10113          | 1                       | NP_04867                                   | 1 A315L                                                                                                                                                                                                                             | 205.30                                                   | 1.15E-51                                                             | 45%                                    | 61%                                    | 1–225                                                    | 1–240                                               |
|              |                    |                |               |       |                   |              | IFNR1 Intron encoded nuclease reneat motif: Reneat of unknown                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |                |                      |                 |              |                  |                |                         |                                            | 7 similar to Chlorella virus PBCV-1 ORF A315L, corresponds to GenBank                                                                                                                                                               |                                                          |                                                                      |                                        |                                        |                                                          |                                                     |
|              |                    |                |               |       | 2                 |              | function, but possibly DNA-binding via helix-turn-helix motif (Ponting, unpublished)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 42.81<br>39.29 | 3.01E-05<br>0.000363 | 40%<br>34%      | 58%<br>55%   | 174227           | 1-53           |                         |                                            | Accession Number 042300                                                                                                                                                                                                             | 193.74                                                   | 3.47E-48<br>2.04E-40                                                 | 45%<br>41%                             | 62%<br>58%                             | 1-228<br>5-225                                           | 1226                                                |
|              |                    |                |               |       | 4                 |              | GIYc, GIY-YIG type nucleases (URI domain); . NUMOD1, NUMOD1 domain                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |                | 0.000363             |                 | 55%<br>69%   | 191<br>174206    | 1-83           | 3                       | NP_04864<br>NP_04885                       | 1 PBCV-1 33kd peptide<br>1 similar to PBCV-1 ORF A315L, corresponds to GenBank Accession<br>Number M74440                                                                                                                           |                                                          | 2.04E-40<br>2.11E-29                                                 | 41%<br>38%                             | 58%<br>51%                             | 5225<br>1200                                             | 7247<br>1196                                        |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                |                      |                 |              |                  |                | 5                       | AAC4924                                    | 4 ORF301<br>5 intron encoded Bmol                                                                                                                                                                                                   | 78.18<br>72.79                                           | 2.12E-13<br>8.92E-12                                                 | 31%<br>31%                             | 46%<br>41%                             | 9162<br>4200                                             | 80-235<br>5-242                                     |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                |                      |                 |              |                  |                | 7                       | CAA3880                                    | 4 GIY COII i1 grp IB protein<br>8 ORF211                                                                                                                                                                                            |                                                          | 2.20E-10<br>1.42E-09                                                 | 30%<br>35%                             | 46%<br>47%                             | 4193<br>55169                                            | 74-259<br>50-140                                    |
|              |                    |                |               |       |                   |              | Thy1, Thymidylate synthase complementing protein. Thymidylate                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |                |                      |                 |              |                  |                | 10                      | YP_29379<br>NP_89939                       | 5 putative endonuclease<br>3 SeqD                                                                                                                                                                                                   | 62.00<br>60.85                                           | 1.57E-08<br>3.51E-08                                                 | 40%<br>29%                             | 58%<br>47%                             | 281<br>1166                                              | 381<br>1176                                         |
| M034L        | 17228-16587        | 214            | 24,469        | 8.64  | 1                 |              | synthase complementing protein (Thy1) complements the thymidine growth requirement of the organisms in which it is found, but shows no homoloay to thymidivate synthase                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |                | 7.83E-54             |                 | 55%          |                  | 1216           |                         |                                            | 0 Synechocystis ORF s111635, corresponds to GenBank Accession Number D90903                                                                                                                                                         |                                                          | 4.18E-74                                                             | 60%                                    | 80%                                    | 1-213                                                    | 1215                                                |
|              |                    |                |               |       | 2                 | COG1351      | THY1, Predicted alternative thymidylate synthase [Nucleotide transport and metabolism].                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 134.76         | 7.22E-33             | 30%             | 44%          | 16213            | 29237          |                         |                                            | 8 thymidilate synthase<br>5 Thymidylate synthase complementing protein                                                                                                                                                              |                                                          | 3.01E-56<br>5.13E-56                                                 | 53%<br>51%                             | 68%<br>69%                             | 1214<br>1214                                             | 3213<br>21236                                       |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                |                      |                 |              |                  |                | 4                       | ZP 0100603                                 | 3 thymidylate synthase 7 Thymidylate synthase complementing protein ThyX                                                                                                                                                            | 219.55<br>219.16                                         | 5.13E-56<br>6.70E-56                                                 | 54%<br>54%                             | 71%<br>71%                             | 12213<br>12213                                           | 6210<br>6210                                        |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                |                      |                 |              |                  |                | ē                       | ZP_0053130                                 | Thymidylate synthase (FAD)     Predicted alternative thymidylate synthase                                                                                                                                                           | 216.47<br>214.54                                         | 4.34E-55<br>1.65E-54                                                 | 53%<br>52%                             | 68%<br>71%                             | 1-213<br>12-213                                          | 1215<br>6210                                        |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                |                      |                 |              |                  |                | 8                       |                                            | 9 thymidylate synthase, flavin-dependent                                                                                                                                                                                            | 213.77                                                   | 2.81E-54<br>2.81E-54                                                 | 51%<br>58%                             | 67%<br>70%                             | 1214<br>19213                                            | 1216<br>12210                                       |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                |                      |                 |              |                  |                | 10                      | ABB2766                                    | 3 Thymidylate synthase complementing protein ThyX                                                                                                                                                                                   | 213.39                                                   | 3.68E-54                                                             | 50%                                    | 69%                                    | 1-214                                                    | 1216                                                |
| M037R        | 1736319141         | 593            | 64,754        | 6.02  | 1                 | COG0449      | GlmS, Glucosamine 6-phosphate synthetase, contains amidotransferase and phosphosugar isomerase domains [Cell envelope biogenesis, outer membrane].  GFAT, Glutamine amidotransferases class-II (Gn-AT)_GFAT-type. This                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | 616.81         | 5.16E-178            | 44%             | 62%          | 1593             | 1597           | 1                       | BAD1529                                    | 9 glutamine:fructose-6-phosphate amidotransferase GFAT                                                                                                                                                                              | 729.55                                                   | 0.00E+00                                                             | 61%                                    | 78%                                    | 1591                                                     | 1594                                                |
|              |                    |                |               |       | 2                 | cd00714      | domain is found at the N-terminus of glucosamine 6-phosphate (GlcN-6-<br>P) synthase (GlN-6-grAF). The glutaminase domain catalyzes amide<br>nitrogen transfer from glutamine to the appropriate substrate. In this<br>process, glutamine is hydrolyzed to glutamic acid and ammonia. GFAT<br>catalyzes the formation of glucosamine 6-phosphate from fructose 6-<br>phosphate and glutamine, the nitriating step in the biosynthesis of UDP-                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 253.13         | 1.55E-68             | 47%             | 66%          | 2208             | 1210           | 2                       | NP_04844                                   | 8 PBCV-1 glucosamine synthetase                                                                                                                                                                                                     | 728.01                                                   | 0.00E+00                                                             | 60%                                    | 77%                                    | 1591                                                     | 1593                                                |
|              |                    |                |               |       | 3                 | COG2222      | GlcN-6-P<br>AgaS, Predicted phosphosugar isomerases [Cell envelope biogenesis,                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | 179 32         | 2.71E-46             | 29%             | 48%          | 248583           | 2329           | 1                       | CAE3949                                    | 3 glucosaminefructose-6-phosphate aminotransferase                                                                                                                                                                                  | 511 15                                                   | 4.12E-143                                                            | 46%                                    | 63%                                    | 1593                                                     | 1610                                                |
|              |                    |                |               |       | 4                 | nfam00310    | Outer membranel.  GATase 2 Glutamine amidotransferases class-II                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | 128.88         | 3.74E-31             | 36%             | 59%          | 2133             | 1136           | 4                       | CAE4499                                    | 2 glucosamine-fructose-6-phosphate aminotransferase                                                                                                                                                                                 | 508.06                                                   | 3.49E-142                                                            | 45%                                    | 63%                                    | 1593                                                     | 1610                                                |
|              |                    |                |               |       | 5                 | COG0034      | PurF, Glutamine phosphoribosylpyrophosphate amidotransferase [Nucleotide transport and metabolism].                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | 128.02         | 7.24E-31             | 31%             | 54%          | 1228             | 4232           | 5                       | ZP_0059423                                 | 2 Glucosamine-fructose-6-phosphate aminotransferase, isomerising                                                                                                                                                                    | 506.52                                                   | 1.02E-141                                                            | 45%                                    | 64%                                    | 1-593                                                    | 1612                                                |

| Gene<br>Name | Genome<br>Position | A.A.<br>length | Peptide<br>Mw | pl  | CDD Hit<br>Number | COGs                          | COG Definition                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | Bit<br>Score            | E-value                         | %<br>Identity F   | %<br>ositive      | Query<br>from-to           | Hit<br>from-to | BLASTP Hit<br>Hit Accession                                                                            | BLASTp Definition                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | Bit E-value lo                                                                                                                                                                                                    | %<br>dentity Po                                             | %<br>ositive f                                                     | Query H<br>from-to                                                                               | lit from-<br>to                                                                                  |
|--------------|--------------------|----------------|---------------|-----|-------------------|-------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------|---------------------------------|-------------------|-------------------|----------------------------|----------------|--------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------|--------------------------------------------------------------------|--------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------|
|              |                    |                |               |     | 6                 |                               | GPATase. N. Gultamine amidotanaferaese class-II (GN-AT), GPAT-<br>type. This domain is found at the N-terminus of gultamine<br>phosphorbosylyryophosphate (Prpp) amidotransferase (GPATase). The<br>gultaminase domain catalyzes amide mitogen transfer from gultamine to<br>the appropriate substrate. In this process, gultamine is hydrolyzed to<br>gultamic acid and ammonia. GPATase catalyzes the first step in purine<br>phosphorbosylamine, pryophosphate and gultamate. GPATase<br>crystalizes as a homoteramer, but can also exist as a homdimer.<br>Gn_AT.II, Glutamine amidotransferases class-II (GATase). The<br>gultaminase domain catalyzes an amide nitrogen transfer from gultamine<br>to the appropriate substrate. In this process, gultamine is hydrolyzed to<br>gultamic acid and ammonia. This column belongs to the Nith hydrolase<br>gultamic acid and ammonia. This domain belongs to the Nith hydrolase<br>gulcosamine-fructose 6-phosphate synthase (GLMS or GPAT), gultamine<br>plosporbroshygyrophosphate (Pyp) middoransferase (GPATase).                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |                         | 5.16E-30                        | 32%               | 52%               |                            | 1224           | 6 ZP_0094299                                                                                           | Clucosamine-fructose-6-phosphate aminotransferase (isomerizing)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 504.21 5.04E-141                                                                                                                                                                                                  | 45%                                                         | 64%                                                                | 1–593                                                                                            | 1-612                                                                                            |
|              |                    |                |               |     | 7                 |                               | a sparagine synthetase B (AnrB), beta lactam synthetase (beta-LS) and<br>glutamate synthesa (GIRS) CLMS catalyzes the formation of glucosamine<br>6-phosphate from fructose 6-phosphate and glutamine in amino sugar<br>synthesis. GPA fase catalyzes the first step in purine biosynthesis, an<br>amide transfer from glutamine to PRPP, resulting in<br>phosphorbobyatime, pyrophosphate and glutamine. Beta<br>LS catalyzes the formation of the beta-lactamin right in the beta-lactamine<br>synthesizes and synthesizes asparagine from sparate and glutamine. Beta<br>LS catalyzes the formation of the beta-lactamin right in beta-lactamine<br>lack the synthesizes and the synthesizes of the synthesizes<br>lack the synthesizes and synthesizes<br>lack the synthesizes and synthesizes<br>lack the synthesizes and synthesizes<br>lack the synthesizes and synthesizes<br>lack the synthesizes<br>lac |                         | 3.32E-25                        | 33%               | 47%               |                            | 1203           | 7 CAD1370                                                                                              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                                                                                                                                                                                                   | 44%                                                         | 63%                                                                |                                                                                                  | 1-612                                                                                            |
|              |                    |                |               |     | 8                 |                               | domains are also found in proteins that regulate the expression of genes involved in synthesis of phosphosugars. Presumably the SIS domains bind to the end-croduct of the nathway.  AsnB, Asparagine synthase (glutamine-hydrolyzing) [Amino acid                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 91.50                   | 7.42E-20<br>4.88E-16            | 37%<br>25%        | 55%               | 283410                     | 3131           |                                                                                                        | 2 Glucosamine-fructose-6-phosphate aminotransferase, isomerising COG0449: Glucosamine 6-phosphate synthetase, contains                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | 497.28 6.16E-139<br>496.89 8.05E-139                                                                                                                                                                              | 45%                                                         | 63%                                                                | 1593                                                                                             | 1616                                                                                             |
|              |                    |                |               |     | 10                |                               | AsnB, Glutamine amidotransferases class-II (GATase) asparagine synthase_B type. Asparagine synthetase B catalyses the ATP-dependent conversion of asparate to asparagine. This enzyme is a homodimer with                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |                         | 4.00E-10                        | 29%               | 49%               |                            | 1163           |                                                                                                        | amidotransferase and phosphosugar isomerase domains glucosaminefructose-6-phosphate aminotransferase, isomerizing                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | 496.09 0.05E-139                                                                                                                                                                                                  | 44%                                                         | 64%                                                                | 1-593                                                                                            | 1619                                                                                             |
|              |                    |                |               |     |                   |                               | each monomer composed of a glutaminase domain and a synthetase<br>domain. The N-terminal glutaminase domain hydrolyzes glutamine to<br>alutamic acid and ammonia                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |                         | 1.112.11                        | 2070              | 4070              | 2 100                      | . 100          | 10 11 _10004                                                                                           | y glocosamine—nuclose-o-priosphate animou ansierase, isomerizing                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 400.12 1.072 100                                                                                                                                                                                                  | 4470                                                        | 0470                                                               | 1 000                                                                                            | . 000                                                                                            |
| M042R        | 1918822055         | 956            | 3 96,592      | 5.7 | 8 1               | COG2911                       | COG2911, Uncharacterized protein conserved in bacteria (Function unknown).                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | 35.67                   | 0.005043                        | 16%               | 30%               | 21-749                     | 304-991        | 2 BAB8346<br>3 BAB8346<br>4 BAB8346<br>5 BAB8347<br>6 NP_04847<br>7 AAA630<br>8 NP_04836<br>9 NP_04836 | Vp260 like protein  PBCV-1 Vp260 protein  Glovoprotein Vp260 like protein  Glovoprotein Vp260  Aar/Thr/Sar/Val rich protein  Aar/Thr/Sar/Val rich protein  Aar/Thr/Sar/Val rich protein  Commenconist in Kuises-Prot Accreation Number P15021  cell aufrica antigen ChingA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | 1353.96 0.00E+00<br>694.50 0.00E+00<br>632.48 2.15E-179<br>626.32 1.54E-177<br>620.54 8.44E-177<br>621.53 8.16E-62<br>241.59 8.16E-62<br>215.31 8.16E-62<br>168.32 1.15E-39<br>166.01 5.69E-39<br>158.30 1.19E-36 | 71% 42% 41% 41% 40% 30% 30% 25% 25%                         | 81%<br>60%<br>57%<br>57%<br>42%<br>41%<br>41%<br>40%<br>37%        | 1-956<br>2-956<br>10-950<br>10-950<br>10-950<br>27-831<br>51-831<br>20-850<br>22-823<br>28-828 2 | 1-955<br>3-954<br>1-930<br>1-930<br>1-931<br>166-1021<br>15-835<br>19-961<br>18-810<br>2551026   |
| M047R        | 2209626496         | 1467           | 7 151,716     | 6.6 | 1 1               | COG2911                       | . COG2911, Uncharacterized protein conserved in bacteria [Function unknown].                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | 40.68                   | 0.000137                        | 20%               | 38%               | 636897                     | 7661006        | 2 BAB8346<br>3 BAB8346<br>4 BAB8346<br>5 BAB8347<br>6 NP_04847<br>7 AAA8630<br>8 NP_04837              | Vy250 like protein PSCV-1 Vy250 protein S                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | 1853.95 0.00E+00<br>1542.32 0.00E+00<br>1515.75 0.00E+00<br>583.56 1.83E-164<br>527.71 1.19E-147<br>190.66 3.45E-46<br>179.87 6.09E-43<br>120.17 5.72E-25<br>115.16 1.84E-23<br>114.39 3.14E-23                   | 68%<br>55%<br>53%<br>35%<br>36%<br>28%<br>30%<br>24%<br>21% | 79%<br>68%<br>67%<br>52%<br>53%<br>40%<br>42%<br>40%<br>36%<br>36% | 101346<br>101467<br>101467<br>41114<br>1-914<br>35857<br>35691<br>19687<br>631308<br>65904       | 1-1336<br>1-1464<br>1-1462<br>5-1121<br>1-923<br>59-973<br>59-748<br>15-755<br>16-1183<br>19-966 |
| M055R        | 26539-30921        | 1461           | 1 149,623     | 5.5 | 5 1               | COG2911                       | COG2911, Uncharacterized protein conserved in bacteria [Function unknown].                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | 41.45                   | 8.60E-05                        | 19%               | 34%               | 374893                     | ¥791000        | 2 BAB8347<br>3 BAB8346<br>4 BAB8346<br>5 BAB8347<br>6 NP 04847<br>7 AAA8630<br>8 NP_04837              | Vy250 like protein PBCV-1 Vy250 pice protein PBCV-1 Vy250 pice protein AsorthySerVal rich protein: AsorthySerVal rich protein: similar to Rickettsia cell surface antigen, corresponds to Swiss-Prot Accession Number P19521 AsorthySerVal rich protein; millar to Rickettsia cell surface antigen, corresponds to Swiss-Prot Accession Number P19521 AsorthySerVal rich protein; millar to Rickettsia cell surface antigen, corresponds to Swiss-Prot Accession Number P15021 AsorthySerVal rich protein                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 1578.53 0.00E+00<br>1574.30 0.00E+00<br>1570.44 0.00E+00<br>628.63 4.93E-178<br>570.85 1.22E-160<br>197.59 2.81E-48<br>174.10 3.33E-41<br>135.96 1.00E-29<br>130.57 4.22E-28<br>121.32 2.56E-25                   | 56%<br>61%<br>55%<br>37%<br>38%<br>28%<br>27%<br>25%<br>23% | 69%<br>73%<br>69%<br>54%<br>53%<br>40%<br>40%<br>40%<br>38%<br>37% | 101461<br>101335<br>101461<br>41110<br>1940<br>54957<br>54806<br>63815<br>17854<br>34855 2       | 1-1464<br>1-1331<br>1-1462<br>5-1118<br>1-949<br>21-922<br>21-793<br>19-805<br>10-951<br>2921156 |
| M061R        | 31058-35458        | 1467           | 7 150,571     | 4.9 |                   |                               | COG2911, Uncharacterized protein conserved in bacteria [Function unknown]. COG1315. Predicted polymerase, most proteins contain PALM domain.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | 41.06                   | 0.0001                          | 21%               | 40%               | 636891                     |                |                                                                                                        | 9 Vp260 like protein                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | 1596.25 0.00E+00                                                                                                                                                                                                  | 57%                                                         | 69%                                                                |                                                                                                  | 1-1464                                                                                           |
|              |                    |                |               |     | 2                 | COG1315                       | recombination, and repair).  DUE342 Protein of unknown function (DUE342). This family of bacterial                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |                         | 0.001453                        | 26%               | 38%               | 636775                     |                |                                                                                                        | Vp260 like protein                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 1587.39 0.00E+00                                                                                                                                                                                                  | 62%                                                         | 73%                                                                | 101322                                                                                           | 1-1311                                                                                           |
| M070L        | 36685–36332        | 118            | 3 13,453      | 9.5 | 3                 | ptam03961                     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 34.54                   | 0.009318                        | 30%               | 45%               | 636784                     | 191–313        | 4 BAB8346<br>5 BAB8347<br>6 NP 04847<br>7 AAA8630<br>8 NP_04837<br>9 NP_04836<br>10 NP_04836           | 3 Vp280 like protein  Vp280 like protein  Vp280 like protein  PBCV-1 Vp280 protein  AsnTIN/Ser/Val rich protein; similar to Rickettsia cell surface antigen, corresconds to Swiss-Prot Accession Number P15921  AsnTIN/Ser/Val rich protein  No Hit Found                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | 1582.39 0.00E+00 620.16 1.76E-175 570.08 2.09E-160 213.77 3.80E-63 194.13 3.12E-47 134.04 3.83E-29 125.95 1.04E-26 123.25 6.76E-26                                                                                | 56%<br>36%<br>38%<br>30%<br>31%<br>23%<br>21%<br>21%        | 70%<br>53%<br>53%<br>42%<br>43%<br>38%<br>38%<br>36%               |                                                                                                  | 1-1462<br>5-1147<br>1-949<br>21-844<br>48-749<br>19-760<br>19-1089<br>1661332                    |
| M071R        | 36740-41230        | 1497           | 7 156,173     | 5.2 | 7 1 2 3           | COG1664<br>COG4801<br>COG2911 | ComA, Inlegral membrane protein ComA involved in cell shape retermination (Call envolven binomenies in dare membrane) code401, Predicted acyltransferase (General function prediction only). COG2911, Uncharacterized protein conserved in bacteria (Function unknown).                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | 37.21<br>36.45<br>35.67 | 0.001544<br>0.00259<br>0.004858 | 19%<br>23%<br>17% | 36%<br>44%<br>34% | 714811<br>514751<br>198815 | 40232          | 2 BAB8346<br>3 BAB8347<br>4 BAB8347<br>5 BAB8347<br>6 NP 04847<br>7 AAA8630<br>8 NP_04336              | Vp260 like protein Vp260 like pr | 1362.05 0.00E+00<br>676.01 0.00E+00<br>653.67 0.00E+00<br>626.32 2.51E-177<br>609.37 3.18E-172<br>609.37 3.39E-65<br>261.685 4.60E-54<br>148.29 2.01E-33<br>144.82 2.22E-32<br>131.72 1.94E-28                    | 62%<br>33%<br>32%<br>33%<br>39%<br>27%<br>27%<br>22%<br>22% | 72%<br>49%<br>49%<br>49%<br>55%<br>41%<br>41%<br>36%<br>36%<br>37% |                                                                                                  | 1-1142<br>3-1455<br>3-1448<br>3-1306<br>2-954<br>24-977<br>24-828<br>6-1328<br>92-1232<br>26-796 |

| Gene<br>Name | Genome<br>Position | A.A.<br>length | Peptide<br>Mw | pl    | CDD Hit<br>Number | COGs                 | COG Definition                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | Bit<br>Score    | E-value              | %<br>dentity F    | %<br>Positive     | Query<br>from-to | Hit<br>from-to    | BLASTp<br>Hit<br>Number                    | BLASTp Definition                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | Bit<br>Score                                                                                     | E-value lo                                                                                               | %<br>lentity Po                                                    |                                                      | Query<br>from-to                                                                                         | Hit from-<br>to                                                                                                    |
|--------------|--------------------|----------------|---------------|-------|-------------------|----------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------|----------------------|-------------------|-------------------|------------------|-------------------|--------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------|------------------------------------------------------|----------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------|
| M078R        | 41286-42515        | 410            | 46,456        | 6.32  | 1                 | pfam04451            | Capsid Iridovir, Iridovirus major capsid protein. This family includes the<br>major capsid protein of iridoviruses, chlorella virus and Spodoptera<br>ascovirus, which are all dsDNA viruses with no RNA stage. This is the<br>most abundant structural protein and can account for up to 45% of virion<br>protein. In Chlorella virus MT325 the major capsid protein is                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | 286.05          | 2.10E-78             | 37%               | 54%               | 5406             | i 3443            | 1                                          | 1 AAC27494 putative capsid protein                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 483.80                                                                                           | 4.34E-135                                                                                                | 99%                                                                | 100%                                                 | 21250                                                                                                    | 1230                                                                                                               |
|              |                    |                |               |       | 2                 | COG0770              | alvonratein                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | 37.21           | 0.001475             | 32%               | 47%               | 10-68            | 382441            | 3<br>4<br>5<br>6<br>7<br>8                 | 7 1M4X C Chain C, Pbcv-1 Virus Capsid, Quasi-Atomic Model                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | 298.13<br>293.51<br>292.35<br>289.27<br>271.55<br>271.55<br>233.03                               | 5.21E-80<br>3.38E-79<br>8.32E-78<br>1.85E-77<br>1.57E-76<br>3.39E-71<br>3.39E-71<br>1.33E-59<br>8.65E-59 | 41%<br>40%<br>40%<br>40%<br>41%<br>39%<br>39%<br>34%<br>32%        | 55%<br>55%<br>54%<br>54%<br>54%<br>54%<br>54%<br>53% | 5-410<br>5-410<br>5-410<br>5-410<br>5-410<br>27-410<br>27-410<br>5-410<br>5-410                          | 3-437<br>3-437<br>3-437<br>3-436<br>3-432<br>1-413<br>1-413<br>3-403<br>2-400                                      |
| M081R        | 4254943169         | 207            | 23,405        | 11.45 |                   | No Hit Found         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                 |                      |                   |                   |                  |                   |                                            | No Hit Found No Hit Found                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |                                                                                                  |                                                                                                          |                                                                    |                                                      |                                                                                                          |                                                                                                                    |
| M083R        | 4334443598         | 85             | 9,762         | 7.49  |                   | No Hit Found         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                 |                      |                   |                   |                  |                   |                                            | No Hit Found No Hit Found                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |                                                                                                  |                                                                                                          |                                                                    |                                                      |                                                                                                          |                                                                                                                    |
| M085R        | 4371645284         | 523            | 57,632        | 9.71  | 1 2 3             | pfam00704<br>COG0810 | Givon 18, Givon 19 domain.  Glivon hufon 18, Glivonari hydrodiases family 18.  TonB, Periplasmic protein 1018, links inner and outer membranes [Cell everlices biosensess, outer membranes, outer membranes, outer membranes, outer membranes, outer membranes, outer membranes, outer membrane l'Trypan, PARP. Procyclic acidici repetitive protein (PARP), This family consists of several Trypanosome bruce procyclic acidic repetitive protein (PARP) like sequences. The procyclic acidic repetitive protein (parp) genes of Trypanosoma brucei encode a small family of abundant surface                                                                                                                                                                                                                                                                                                                                                                                                                                   | 209.60<br>68.25 |                      | 38%<br>35%<br>45% | 57%<br>50%<br>51% |                  |                   | 3                                          | NP_048613 PBCV-1 chitinase     CAB4709 chitinase     AA232788 chitinase  4 EAA73155 hypothetical protein FG03591.1                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 225.33<br>224.94                                                                                 | 3.37E-178<br>3.86E-57<br>5.04E-57<br>2.50E-56                                                            | 74%<br>34%<br>34%<br>35%                                           | 82%<br>52%<br>52%                                    | 1-390<br>1-372<br>1-372                                                                                  | 1-392<br>19-384<br>19-384                                                                                          |
|              |                    |                |               |       | •                 |                      | proteins whose expression is restricted to the procyclic form of the<br>parasite. They are found at two unlinked loci, parpA and parpB;<br>transcription of both loci is developmentally regulated                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 00.10           | 2.041-12             | 37.10             | 30 /6             | 384-408          | 47-122            | •                                          | 4 ENTOTO hypothetical protein PG03991.1                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | 222.03                                                                                           | 2.502-50                                                                                                 | 33 /6                                                              | 3270                                                 | 14-572                                                                                                   | 30-377                                                                                                             |
|              |                    |                |               |       | 5                 |                      | Neisseria_TspB, Neisseria meningitidis TspB protein. This family consists of several Neisseria meningitidis TspB virulence factor proteins                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 58.60           | 6.68E-10             | 29%               | 34%               | 405492           | 323-410           | 5                                          | 5 AAP04616 chitinase                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | 221.86                                                                                           | 4.26E-56                                                                                                 | 35%                                                                | 51%                                                  | 1-372                                                                                                    | 19-384                                                                                                             |
|              |                    |                |               |       | 6                 |                      | DUF1210, Protein of unknown function (DUF1210). This family represents a conserved region within plant proline-rich proteins.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 50.88           |                      | 41%               | 46%               |                  | 124202            |                                            | 6 AAY32603 chitinase                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                                                                                                  | 8.04E-55                                                                                                 | 33%                                                                | 51%                                                  | 1372                                                                                                     | 19-384                                                                                                             |
|              |                    |                |               |       | 7                 |                      | DeedD, Uncharacterized protein conserved in bacteria [Function unknown].  RPH3A, effector, Rabphilin-3A effector domain. This is a family of proteins involved in potein transport in synaptic vesicles. Rabphilin-3A has been shown to contact Rab3A, a small G protein important in                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 46.94<br>46.99  | 1.95E-06<br>2.05E-06 | 31%<br>27%        | 35%<br>36%        |                  | 65149<br>i 165246 |                                            | 7 AAT37496 putative chitinase  8 CAA45468 chitinase                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                                                                                                  |                                                                                                          | 33%                                                                | 50%<br>49%                                           | 12372<br>3389                                                                                            | 2352<br>25398                                                                                                      |
|              |                    |                |               |       | 9                 | pfam05539            | Pneumo_att_G, Pneumovirinae attachment membrane glycoprotein G                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | 45.54           | 5.82E-06             | 20%               | 31%               | 395485           | 228-323           | 9                                          | 9 AAF19616 42 kDa endochitinase                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 213.00                                                                                           | 1.98E-53                                                                                                 | 34%                                                                | 48%                                                  | 11389                                                                                                    | 34399                                                                                                              |
|              |                    |                |               |       | 10                |                      | MCPVI, Minor capsid protein VI. This minor capsid protein may act as a link between the external capsid and the internal DNA-protein core. The C-terminal 11 residues may function as a protease cofactor leading to enzyme activation.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 45.40           | 6.34E-06             | 29%               | 37%               | 396484           | 110205            | 10                                         | 10 AAF19613 42 kDa endochitinase                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 213.00                                                                                           | 1.98E-53                                                                                                 | 34%                                                                | 48%                                                  | 11389                                                                                                    | 34–399                                                                                                             |
| M086R        | 4534445931         | 196            | 22,100        | 8.10  |                   | No Hit Found         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                 |                      |                   |                   |                  |                   | 1                                          | 1 NP_048655 A301L                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | 59.69                                                                                            | 5.64E-08                                                                                                 | 30%                                                                | 52%                                                  | 3-132                                                                                                    | 2135                                                                                                               |
| M088R        | 4596746605         | 213            | 23,543        | 4.68  |                   | No Hit Found         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                 |                      |                   |                   |                  |                   | 1                                          | 1 NP_048652 A298L                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | 158.30                                                                                           | 1.39E-37                                                                                                 | 42%                                                                | 59%                                                  | 1208                                                                                                     | 1220                                                                                                               |
| M090L        | 4717946685         | 165            | 18,423        | 12.17 |                   | No Hit Found         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                 |                      |                   |                   |                  |                   | 1                                          | 1 NP_048650 A296R                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | 80.49                                                                                            | 1.94E-14                                                                                                 | 64%                                                                | 76%                                                  | 80144                                                                                                    | 72-135                                                                                                             |
| M091R        | 47173-48201        | 343            | 39,379        | 9.13  | 2                 |                      | Glyco, pytro, 48, Glycosyl hydrolase family 46. This family are chicosanase enzymes.  chitosanase glyco, lydro, 48, Glycosyl hydrolase family 46 chitosanase droman. This family are composed of the chitosanase enzymes which hydrolyzes chitosan, a biopolymer of beta (14,4) linked -D-glucosamine (Gloth) residues produced by partial or full decebylation of chitin. Chitosanases play a role in defense against pathogens such as fungi and are found in microorganisms, fungi, vurses, and plants. Microbial chitosanases who members are the most prevalent can be divided into 3 subclasses based on the specificity of the cleavage positions for partial subclasses haded on the specificity of the cleavage positions for partial Gloth-Gloth and Gloth-Gloth linkages, whereas subclass II chitosanases such as Bacillus sp. no. 7-14 can cleave only Gloth-Gloth linkages. Subclass III chitosanases such as MH-H1 chitosanase are the most versatile and can applie bot file-NG-Gloth and Gloth-Gloth, linkages. |                 | 1.06E-63<br>2.90E-53 | 39%               | 50%<br>44%        |                  | 1215              |                                            | 1 BAA20342 vChta-1 2 NP_048646 PBCV-1 chitosanase                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |                                                                                                  | 9.50E-114<br>1.79E-112                                                                                   | 64%                                                                | 73%                                                  | 18-335<br>18-335                                                                                         | 5-321<br>5-321                                                                                                     |
| M093L        | 4956548435         | 377            | 43,656        | 9.08  |                   | No Hit Found         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                 |                      |                   |                   |                  |                   | 4<br>5<br>6<br>7<br>8<br>9<br>10           | 3         BAA94840 chilosanase           4         CAB14630 chilosanase           5         AAA19895 chilosanase precursor           7         ZP_00900309 chilosanase           8         1CHK, B. Chain B. Shreptomyces N174 Chilosanase Ph5.5 298k           9         ABC17783 secreted chilosanase precursor           1         NP 048711 A354R                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 83.96<br>81.65<br>81.26<br>79.72<br>79.34<br>78.95<br>76.26                                      | 5.93E-15<br>7.74E-15<br>3.84E-14<br>5.02E-14<br>1.46E-13<br>1.91E-13<br>2.49E-13<br>1.61E-12<br>3.03E-47 | 29%<br>29%<br>28%<br>28%<br>31%<br>29%<br>27%<br>28%               | 47%<br>42%<br>44%<br>46%<br>48%<br>45%<br>44%<br>60% | 96-304<br>107-328<br>93-329<br>101-329<br>117-324<br>117-329<br>107-324<br>107-328                       | 41-252<br>37-264<br>17-266<br>39-266<br>38-246<br>15-226<br>51-274<br>38-265<br>1-235                              |
|              |                    |                |               |       |                   |                      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                 |                      |                   |                   |                  |                   |                                            | 2 NP_048779 similar to Bacteriophage SP01 gene 31 intron, corresponds to Swiss-Prot<br>Accession Number P34081<br>3 NP_048435 A87R                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |                                                                                                  | 2.05E-11<br>2.96E-10                                                                                     | 28%<br>31%                                                         | 47%<br>46%                                           | 85345<br>194373                                                                                          | 58-319<br>270-447                                                                                                  |
| M097L        | 5062549549         | 359            | 39,960        | 8.73  |                   | No Hit Found         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                 |                      |                   |                   |                  |                   | 1                                          | 1 NP 048640 A286R                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | 363.23                                                                                           | 7.11E-99                                                                                                 | 51%                                                                | 69%                                                  | 17344                                                                                                    | 22-362                                                                                                             |
| M099R        | 5066451170         | 169            | 18,220        | 8.23  | 1                 | pfam00080            | Sod_Cu, Copperizinc superoxide dismutase (SODC), superoxide dismutases (SODS) cataliyes the conversion of superoxide radicals to hydrogen peroxide and molecular oxygen. Three evolutionarily distinct families of SODs are known, of which the copperizinc-binding family is one. Defects in the human SODI gene cause familial amyrotrophic lateral sclerosis (Lou Gehrigäapors, disease). Structure is an eight-stranded beta sandvich. Smill or the immunolobulin fold.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | 179.67          | 2.21E-46             | 58%               | 74%               | 16165            | i 1152            |                                            | contains Cu/Zn superoxide dismutase signatures 1 and 2; similar to 1 NP_048593 Neurospora Cu/Zn-superoxide dismutase, corresponds to Swiss-Prot 2 XP_503850 humothetical protein                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                                                                                                  | 5.00E-69                                                                                                 | 76%                                                                | 85%                                                  | 10169                                                                                                    | 28-187                                                                                                             |
| M100L        | 5175451179         | 192            | 22,991        | 7.15  |                   | No Hit Found         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                 |                      |                   |                   |                  |                   | 2<br>3<br>4<br>5<br>6<br>7<br>8<br>9<br>10 | 0751L8 Superviole dismutase   1791L8 Superviole dismutase   1790870 Superviole dismutase   179088116 Superviole dismutase   1790870 Superviole dismutase | 179.87<br>179.49<br>179.10<br>177.95<br>177.18<br>177.18<br>177.18<br>177.18<br>177.18<br>177.18 | 2.50E-44<br>3.26E-44<br>4.26E-44                                                                         | 60%<br>57%<br>59%<br>58%<br>58%<br>57%<br>58%<br>58%<br>58%<br>58% | 71% 72% 71% 72% 70% 71% 70% 71% 70% 71% 72% 72% 50%  | 16-162<br>16-164<br>18-164<br>16-162<br>18-165<br>16-162<br>18-165<br>18-165<br>17-162<br>4-172<br>4-174 | 1-151<br>1-153<br>3-150<br>1-151<br>4-152<br>1-151<br>4-152<br>4-152<br>4-152<br>1-150<br>3-171<br>6-119<br>26-158 |
| M101R        | 51793-52629        | 279            | 31,108        | 6.29  | 1                 | cd00542              | PVA, Penicillin V acytase (PVA), also known as conjugated bile salt acid hydrolase (CBAH), catalyzes the trytrolysis of penicillin V to yeld 6-marino penicillina caid (6-APA), an important key intermediate of semisynthetic penicillins. PVA has an N-terminal nucleophilic cysteine as do other NH hydrolases which is exposed by prosessing of the PVA research.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 191.58          | 5.57E-50             | 33%               | 48%               | 2275             | i 1297            | 5<br>4<br>5                                | 4 EAL86798 phosphoglycerate mutase family protein, putative                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | 49.29<br>49.29                                                                                   | 7.21E-05<br>7.21E-05<br>2.38E-90                                                                         | 24%<br>27%<br>57%                                                  | 43%<br>47%<br>74%                                    | 4164<br>4117                                                                                             | 22-174<br>9-123<br>1-279                                                                                           |

| Ger<br>Nan |           |         |        | Peptide<br>Mw | pl    | CDD Hit     | COGs         | COG Definition                                                                                                                                                                                                                                                          | Bit    | E-value              | %          | %          | Query         | f 4-             | BLASTp<br>Hit | Hit            |                                                                                                                                       | Bit                     | E-value                          | %<br>Identity F   | %                 |                         | Hit from-                  |
|------------|-----------|---------|--------|---------------|-------|-------------|--------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------|----------------------|------------|------------|---------------|------------------|---------------|----------------|---------------------------------------------------------------------------------------------------------------------------------------|-------------------------|----------------------------------|-------------------|-------------------|-------------------------|----------------------------|
| Ivaii      | ne Posi   | ition   | length | IVIV          |       | Number<br>2 | COG3049      | COG3049, Penicillin V acylase and related amidases [Cell envelope                                                                                                                                                                                                       | 172 05 | 3.72E-44             | dentity F  | 46%        |               | from-to<br>22323 | Number        | Access         | 8651 COG3049: Penicillin V acylase and related amidases                                                                               | Score<br>145.98         | 1.20E-33                         | Identity F        | 48%               | from-to<br>2276         | to<br>4304                 |
|            |           |         |        |               |       | -           | 000004       | biodenesis. outer membranel.  CBAH, Linear amide C-N hydrolases, choloylglycine hydrolase family.                                                                                                                                                                       | 172.00 | 0.722 44             | 2070       | 40,0       | . 270         | 22 020           | -             |                | 5500 COGSO46. Perilciliii V acylase and related amidases                                                                              | 140.00                  | 1.202.00                         | 0170              | 4070              | 2 2/0                   | 4 004                      |
|            |           |         |        |               |       | 3           | pfam02275    | This family includes several hydrolases which cleave carbon-niftrogen bonds, other than peptide bonds, in linear amides. These include hotoloylglycine hydrolase (conjugated bile acid hydrolase, CBAH) EC:3.5.1.24, penicillin acylase EC:3.5.1.11 and acid ceramidase | 147.84 | 7.49E-37             | 32%        | 50%        | 2275          | 1298             | 3             | AAP08          | 8002 Choloylglycine hydrolase                                                                                                         | 143.67                  | 5.96E-33                         | 31%               | 48%               | 1-275                   | 1-298                      |
|            |           |         |        |               |       | 4           | cd01935      | FC:3.5.1.23 CGH_like, Chololylglycine hydrolase (CGH)_like. This family of chololylglycine hydrolases-like proteins includes conjugated bile acid hydrolase (CBAH), penicillin acylase and acid ceramidase which cleave                                                 | 116.83 | 1.54E-27             | 29%        | 42%        | 2268          | 1270             | 4             | ABB1           | 1030 Penicillin amidase                                                                                                               | 141.35                  | 2.96E-32                         | 29%               | 46%               | 1-278                   | 1-315                      |
|            |           |         |        |               |       |             |              | carbon-nitrogen bonds, other than peptide bonds, in linear amides.  CGH, CGH Choloylglycine hydrolase (also known as bile salt hydrolase) is an intestinal bacterial enzyme responsible for the deconjugation and                                                       |        |                      |            |            |               |                  |               |                |                                                                                                                                       |                         |                                  |                   |                   |                         |                            |
|            |           |         |        |               |       | 5           | cd01902      | 2 subsequent dehydroxylation of conjugated cholic acid (CA) to form deoxycholic acid (DCA). CGH has a conserved Ntn hydrolase fold similar to those of penicillin v acylase (PVA) and acid ceramidase (AC)                                                              | 80.82  | 1.27E-16             | 27%        | 46%        | 2261          | 1270             | 5             |                | 0567 choloylglycine hydrolase                                                                                                         |                         | 5.04E-32                         | 30%               | 47%               | 1275                    |                            |
|            |           |         |        |               |       |             |              |                                                                                                                                                                                                                                                                         |        |                      |            |            |               |                  | 6<br>7        | AAL51          | 8968 choloylglycine hydrolase family protein<br>1724 CHOLOYLGLYCINE HYDROLASE                                                         | 138.27<br>133.65        | 2.50E-31<br>6.17E-30             | 29%<br>29%        | 47%<br>45%        | 1-275<br>2-276          | 1298<br>37337              |
|            |           |         |        |               |       |             |              |                                                                                                                                                                                                                                                                         |        |                      |            |            |               |                  | 9             | YP_414         | 10379 choloylqlycine hydrolase family protein<br>4856 Choloylqlycine hydrolase                                                        | 133.65                  | 6.17E-30<br>6.17E-30             | 29%<br>29%        | 45%<br>45%        | 2276<br>2276            | 3303<br>3303               |
|            |           |         |        |               |       |             |              |                                                                                                                                                                                                                                                                         |        |                      |            |            |               |                  | 10            |                | 3011 choloylglycine hydrolase family protein                                                                                          | 132.88                  | 1.05E-29                         | 28%               | 47%               | 1275                    | 1298                       |
| M10        | 03L 53704 | 4-52817 | 296    | 32,914        | 7.01  | 1           | COG0388      | 8 COG0388, Predicted amidohydrolase [General function prediction only]. CN hydrolase, Carbon-nitrogen hydrolase. This family contains                                                                                                                                   | 181.89 | 4.09E-47             | 34%        | 52%        | 1296          | 1272             | 1             | NP_048         | 8426 contains ATP/GTP-binding site motif A; similar to rat beta-alanine synthetase. corresponds to Swiss-Prot Accession Number Q03248 | 420.24                  | 3.63E-116                        | 68%               | 79%               | 2-295                   | 4297                       |
|            |           |         |        |               |       | 2           | pfam00795    | hydrolases that break carbon-nitrogen bonds. The family includes:<br>Nitrilase EC:3.5.5.1, Aliphatic amidase EC:3.5.1.4, Biotidinase                                                                                                                                    | 143.56 | 1.70E-35             | 34%        | 53%        | 4177          | 1174             | 2             | ZP_00403       | 3767 COG0388: Predicted amidohydrolase                                                                                                | 278.87                  | 1.31E-73                         | 46%               | 65%               | 5294                    | 6289                       |
|            |           |         |        |               |       | 3           | COG0815      | EC:3.5.1.12. Beta-ureidopropionase EC:3.5.1.6  Lnt, Apolipoprotein N-acyltransferase [Cell envelope biogenesis, outer membranel.                                                                                                                                        | 51.17  | 1.10E-07             | 23%        | 36%        | 35215         | 269451           | 3             | BAC60          | 0037 putative carbon-nitrogen hydrolase                                                                                               | 278.10                  | 2.23E-73                         | 49%               | 65%               | 6291                    | 8280                       |
|            |           |         |        |               |       |             |              | membraner.                                                                                                                                                                                                                                                              |        |                      |            |            |               |                  | 4 5           | AAK99          | 9627 Beta-alanine synthase or beta-ureidopropionase<br>5873 beta-alanine synthase                                                     | 276.94<br>276.17        | 4.96E-73<br>8.46E-73             | 46%<br>47%        | 65%<br>63%        | 5294<br>4295            | 6289<br>9291               |
|            |           |         |        |               |       |             |              |                                                                                                                                                                                                                                                                         |        |                      |            |            |               |                  | 6 7           | ABB08          | 18352 Nitrilase/cyanide hydratase and apolipoprotein N-acyltransferase 1984 COG0388: Predicted amidohydrolase                         | 273.86<br>273.86        | 4.20E-72<br>4.20E-72             | 47%<br>46%        | 62%<br>63%        | 4295<br>2294            | 7294<br>3285               |
|            |           |         |        |               |       |             |              |                                                                                                                                                                                                                                                                         |        |                      |            |            |               |                  | 8<br>9        | ABA40<br>CAG77 | 0443 beta-alanine synthase-like protein<br>7171 putative carbon-nitrogen hydrolase                                                    | 273.86<br>271.94        | 4.20E-72<br>1.60E-71             | 46%<br>46%        | 63%<br>62%        | 4-295<br>2-294          | 9291<br>3285               |
|            |           |         |        |               |       |             |              |                                                                                                                                                                                                                                                                         |        |                      |            |            |               |                  | 10            |                | 3658 putative carbon-nitrogen hydrolase                                                                                               | 271.94                  | 1.60E-71                         | 46%               | 63%               | 2294                    | 3285                       |
| M10        |           | 454709  | 302    |               | 10.66 |             | No Hit Found |                                                                                                                                                                                                                                                                         |        |                      |            |            |               |                  |               |                | 8591 A243R                                                                                                                            |                         | 4.72E-135                        | 76%               | 87%               | 1302                    | 1302                       |
| M10        | J9R 55071 | 1-55517 | 149    | 16,949        | 8.37  |             | No Hit Found |                                                                                                                                                                                                                                                                         |        |                      |            |            |               |                  | 1 2           | AAU06          | 8604 A253R<br>6294 hypothetical protein A253R                                                                                         | 72.02                   | 5.73E-12<br>5.73E-12             | 34%<br>35%<br>33% | 53%<br>53%<br>53% | 25147<br>25147<br>25149 | 20-146<br>20-146<br>20-148 |
|            |           |         |        |               |       |             |              |                                                                                                                                                                                                                                                                         |        |                      |            |            |               |                  | 4             | AAU06          | 6289 hypothetical protein A253R<br>16293 hypothetical protein A253R<br>16292 hypothetical protein A253R                               | 70.86<br>51.22<br>50.83 | 1.28E-11<br>1.05E-05<br>1.37E-05 | 38%<br>38%        | 57%<br>57%        | 25149<br>2597<br>2593   | 20148<br>2091<br>2087      |
|            |           |         |        |               |       |             |              |                                                                                                                                                                                                                                                                         |        |                      |            |            |               |                  | 6             |                | 16291 hypothetical protein A253R<br>16291 hypothetical protein A253R                                                                  | 50.83                   | 2.33E-05                         | 38%               | 56%               | 25-93<br>25-97          | 2091                       |
|            |           |         |        |               |       |             |              | GIY-YIG_Cterm, GIYX(10-11)YIG family of class I homing<br>endonucleases C-terminus (GIY-YIG_Cterm). Homing endonucleases                                                                                                                                                |        |                      |            |            |               |                  |               |                |                                                                                                                                       |                         |                                  |                   |                   |                         |                            |
|            |           |         |        |               |       |             |              | promote the mobility of intron or intein by recognizing and cleaving a<br>homologous allele that lacks the sequence. They catalyze a double-                                                                                                                            |        |                      |            |            |               |                  |               |                |                                                                                                                                       |                         |                                  |                   |                   |                         |                            |
|            |           |         |        |               |       |             |              | strand break in the DNA near the insertion site of that element to facilitate<br>homing at that site. Class I homing endonucleases are sorted into four<br>families based on the presence of these motifs in their respective N-                                        |        |                      |            |            |               |                  |               |                |                                                                                                                                       |                         |                                  |                   |                   |                         |                            |
| M11        | 11R 55571 | 1-56287 | 239    | 27,298        | 9.91  | 1           | cd00283      | termini: LAGLIDADG, His-Cys box, HNH, and GIY-YIG. This CD contains several but not all members of the GIY-YIG family. The C-terminus of                                                                                                                                | 61.17  | 9.29E-11             | 38%        | 59%        | 110235        | 3113             | 1             | NP_048         | 8671 A315L                                                                                                                            | 200.29                  | 4.06E-50                         | 43%               | 61%               | 1-237                   | 1242                       |
|            |           |         |        |               |       |             |              | GIY-YIG is a DNA-binding domain which is separated from the N-<br>terminus by a long, flexible linker. The DNA-binding domain consists of a                                                                                                                             |        |                      |            |            |               |                  |               |                |                                                                                                                                       |                         |                                  |                   |                   |                         |                            |
|            |           |         |        |               |       |             |              | minor-groove binding alpha-helix, and a helix-turn-helix. Some also contain a zinc finger (i.e. I-TevI) which is not required for DNA binding or catalysis, but is a component of the linker and directs the catalytic                                                  |        |                      |            |            |               |                  |               |                |                                                                                                                                       |                         |                                  |                   |                   |                         |                            |
|            |           |         |        |               |       |             |              | domain to cleave the homing site at a fixed distance from the intron<br>insertion site<br>IENR1, Intron encoded nuclease repeat motif; Repeat of unknown                                                                                                                |        |                      |            |            |               |                  |               |                | 9007 Similar to Chlorella virus PBCV-1 ORF A315L, corresponds to GenBank                                                              |                         |                                  |                   |                   |                         |                            |
|            |           |         |        |               |       | 2           |              | 7 function, but possibly DNA-binding via helix-turn-helix motif (Ponting, unpublished).                                                                                                                                                                                 | 39.35  | 0.000336             | 40%        | 63%<br>49% | 182237        | 1-53             |               |                | Accession Number 042560                                                                                                               | 192.59                  | 8.47E-48                         | 44%               | 59%<br>53%        | 1-236                   | 1224                       |
|            |           |         |        |               |       | 3           |              | 5 GIYc, GIY-YIG type nucleases (URI domain); . 3 NUMOD1, NUMOD1 domain                                                                                                                                                                                                  |        | 0.001287<br>0.001811 | 33%<br>50% | 68%        | 189<br>182210 | 1-83<br>1-29     | 3             | NP_048         | 8641 PBCV-1 33kd peptide<br>8851 similar to PBCV-1 ORF A315L, corresponds to GenBank Accession<br>Number M74440                       | 132.88<br>128.64        | 7.96E-30<br>1.50E-28             | 36%<br>37%        | 53%               | 11236<br>1208           | 15-248<br>1196             |
|            |           |         |        |               |       |             |              |                                                                                                                                                                                                                                                                         |        |                      |            |            |               |                  | 5<br>6        |                | 19393 SegD<br>19244 ORF301                                                                                                            | 66.63<br>65.08          | 7.01E-10<br>2.04E-09             | 28%<br>28%        | 45%<br>43%        | 1228<br>11171           | 1231<br>84236              |
|            |           |         |        |               |       |             |              |                                                                                                                                                                                                                                                                         |        |                      |            |            |               |                  | 7 8           | YP_293         | 3795 putative endonuclease<br>I8813 GIY ND1 i4 grp IB protein b                                                                       | 65.08<br>63.54          | 2.04E-09<br>5.93E-09             | 41%<br>29%        | 58%<br>41%        | 279<br>11198            | 381<br>77247               |
|            |           |         |        |               |       |             |              |                                                                                                                                                                                                                                                                         |        |                      |            |            |               |                  | 9<br>10       | NP_074         | 4961 orf261<br>19248 ORF211                                                                                                           | 63.54<br>60.08          | 5.93E-09<br>6.56E-08             | 29%<br>33%        | 41%<br>46%        | 11198<br>53177          | 76-246<br>50-140           |
| M1         | 14L 56971 | 1-56294 | 226    | 24,793        | 8.60  |             | No Hit Found |                                                                                                                                                                                                                                                                         |        |                      |            |            |               |                  | 1             |                | 8616 A262L                                                                                                                            | 122.48                  | 9.53E-27                         | 47%               | 66%               | 111215                  | 2103                       |
| M1         | 16L 57762 | 2-56998 | 255    | 29,094        | 9.41  |             | No Hit Found |                                                                                                                                                                                                                                                                         |        |                      |            |            |               |                  | 1             | NP_048         | 8619 A265L                                                                                                                            | 198.36                  | 1.74E-49                         | 56%               | 68%               | 85254                   | 1170                       |
| M1:        | 20L 58145 | 557789  | 119    | 13,637        | 4.55  |             | No Hit Found |                                                                                                                                                                                                                                                                         |        |                      |            |            |               |                  |               | No Hit Fo      | ound No Hit Found                                                                                                                     |                         |                                  |                   |                   |                         |                            |
| M12        | 21R 58225 | 5-58503 | 93     | 10,134        | 9.41  |             | No Hit Found |                                                                                                                                                                                                                                                                         |        |                      |            |            |               |                  |               | No Hit Fo      | ound No Hit Found                                                                                                                     |                         |                                  |                   |                   |                         |                            |
| M12        | 22R 58523 | 3-58945 | 141    | 16,343        | 11.34 |             | No Hit Found |                                                                                                                                                                                                                                                                         |        |                      |            |            |               |                  | 1             | NP_048         | 8484 a136R                                                                                                                            | 58.92                   | 5.04E-08                         | 31%               | 48%               | 1141                    | 1146                       |
|            |           |         | 348    | 00.750        | 4.70  |             |              | GH16_laminarinase, Laminarinase, also known as glucan endo-1,3-beta-<br>D-glucosidase, is a glycosyl hydrolase family 16 member that hydrolyzes                                                                                                                         | 440.74 | 0.045.05             | 0001       | 540/       | 00.040        | 4 000            |               | ND 04          | 0410                                                                                                                                  | 440.00                  |                                  | 540/              | 700/              | 0.045                   | 40.004                     |
| MI         | 24L 60186 | 6-59143 | 348    | 38,758        | 4.79  | 1           | C0U218U      | 1,3-beta-D-glucosidic linkages in 1,3-beta-D-glucans such as laminarins, curdians, paramylons, and pachymans, with very limited action on mixed-link (1,3-1,4-)-heta-D-plusans                                                                                          | 142.71 | 2.61E-35             | 36%        | 51%        | 88-343        | 1236             | 1             | NP_048         | 8442 PBCV-1 beta-1,3 glucanase                                                                                                        | 410.22                  | 4.83E-113                        | 54%               | 70%               | 9345                    | 12-361                     |
|            |           |         |        |               |       |             |              | Glyco_hydrolase_16, The O-Glycosyl hydrolases are a widespread group of enzymes that hydrolyse the glycosidic bond between two or                                                                                                                                       |        |                      |            |            |               |                  |               |                |                                                                                                                                       |                         |                                  |                   |                   |                         |                            |
|            |           |         |        |               |       |             |              | more carbohydrates, or between a carbohydrate and a non-carbohydrate                                                                                                                                                                                                    |        |                      |            |            |               |                  |               |                |                                                                                                                                       |                         |                                  |                   |                   |                         |                            |
|            |           |         |        |               |       | 2           | cd00413      | 3 similarity has led to the definition of more than 95 different families<br>inlouding glycoside hydrolase family 16. Family 16 includes lichenase,                                                                                                                     | 94.29  | 1.14E-20             | 34%        | 48%        | 92-343        | 1216             | 2             | P23            | 3903 Glucan endo-1,3-beta-glucosidase A1 precursor ((1->3)-beta-glucan endohydrolase) ((1->3)-beta-glucanase A1)                      | 117.86                  | 4.95E-25                         | 31%               | 46%               | 74343                   | 410-678                    |
|            |           |         |        |               |       |             |              | xyloglucan endotransglycosylase (XET), beta-agarase, kappa-<br>carrageenase, endo-beta-1,3-glucanase, endo-beta-1,3-1,4-glucanase,                                                                                                                                      |        |                      |            |            |               |                  |               |                |                                                                                                                                       |                         |                                  |                   |                   |                         |                            |
|            |           |         |        |               |       |             |              | and endo-beta-galactosidase, all of which have a conserved jelly roll fold with a deep active site channel harboring the catalytic residues                                                                                                                             |        |                      |            |            |               |                  |               |                |                                                                                                                                       |                         |                                  |                   |                   |                         |                            |
|            |           |         |        |               |       |             |              | GH16_laminarinase_like, A beta-1,3-glucanase (laminarinase)-like<br>protein exists in the bacterial genus Streptomyces as well as the fungal<br>class Sordariomycetes. The laminarinases belong to glycosyl hydrolase                                                   | _      |                      | _          |            |               |                  |               |                |                                                                                                                                       |                         |                                  |                   |                   |                         |                            |
|            |           |         |        |               |       | 3           | cd02182      | channel. The bacterial members contain an additional C-terminal                                                                                                                                                                                                         | 79.95  | 2.18E-16             | 29%        | 43%        | 85343         | 1256             | 3             | YP_435         | 5911 Beta-glucanase/Beta-glucan synthetase                                                                                            | 117.09                  | 8.45E-25                         | 32%               | 46%               | 79343                   | 322-571                    |
|            |           |         |        |               |       | 4           | COG2273      | carbohydrate-hinding module (CRM)<br>3 SKN1, Beta-glucanase/Beta-glucan synthetase [Carbohydrate transport                                                                                                                                                              | 64.69  | 8.99E-12             | 26%        | 39%        | 74-348        | 29267            | 4             | AAC60          | i0453 beta-1,3-glucanase                                                                                                              | 115.16                  | 3.21E-24                         | 31%               | 44%               | 84343                   | 424-681                    |
|            |           |         |        |               |       | 5           |              | and metabolisml.  2 Glyco hydro 16, Glycosyl hydrolases family 16  GH16_beta_GRP, Beta-GRP (beta-1,3-glucan recognition protein) is                                                                                                                                     |        | 3.46E-10             | 28%        | 45%        | 190342        |                  |               |                | 1367 Glycoside hydrolase, family 16                                                                                                   |                         | 1.22E-23                         | 32%               | 45%               |                         | 48-325                     |
|            |           |         |        |               |       |             |              | one of several pattern recognition receptors (PRRs), also referred to as biosensor proteins, that complexes with pathogen-associated beta-1.3-                                                                                                                          |        |                      |            |            |               |                  |               |                |                                                                                                                                       |                         |                                  |                   |                   |                         |                            |
|            |           |         |        |               |       | 6           | cd02179      | glucans and then transduces signals necessary for activation of an<br>appropriate immune response. Their structures adopt a jelly roll fold with                                                                                                                        | 58.78  | 4.92E-10             | 24%        | 37%        | 89329         | 2283             | 6             | ZP_00637       | 7494 Glucan endo-1,3-beta-D-glucosidase                                                                                               | 111.31                  | 4.64E-23                         | 31%               | 43%               | 85343                   | 48-325                     |
|            |           |         |        |               |       |             |              | a deep active site channel harboring the catalytic residues, like those of<br>other alvoosyl hydrolase family 16 members                                                                                                                                                |        |                      |            |            |               |                  |               |                |                                                                                                                                       |                         |                                  |                   |                   |                         |                            |
|            |           |         |        |               |       |             |              |                                                                                                                                                                                                                                                                         |        |                      |            |            |               |                  |               |                |                                                                                                                                       |                         |                                  |                   |                   |                         |                            |

| Gene<br>Name | Genome<br>Position | A.A.<br>length | Peptide<br>Mw | pl   | CDD Hit<br>Number | COGs         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | Bit<br>Score   | E-value              | %<br>Identity | %<br>Positive | Query<br>from-to | Hit<br>from-to | BLASTp<br>Hit<br>Number | Hit<br>Accession                                             | BLASTp Definition                                                                                                    | Bit<br>Score               | E-value                          | %<br>Identity P   | %<br>Positive     | Query<br>from-to        | Hit from-<br>to    |
|--------------|--------------------|----------------|---------------|------|-------------------|--------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------|----------------------|---------------|---------------|------------------|----------------|-------------------------|--------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------|----------------------------|----------------------------------|-------------------|-------------------|-------------------------|--------------------|
|              |                    |                |               |      | 7                 | cd02183      | GH16, GPI glucanosyltransferase, GPI (glycosylphosphatidylinosital) -<br>glucanosyltransferase is a GPI-anchroed membrane problen present in<br>the fungal cell wall that is thought to play an important role in cell wall<br>blosynthesis. GPI-pulcanosyltransferase belongs to a family of glycosyl<br>hydrolases that includes lichenase, vyloglucan endotransglycosylase<br>(KET), beta-agrese, kappa-carragenase, endo-beta-jaluchosase, and rother<br>kende beta-1,3-1,4-glucanase, and endo-beta-glasticosidase, all of which<br>have a conserved jelly rol fold with a deep active site channel harboring                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 38.68          | 0.000569             | 28%           | 49%           | 190283           | 58144          | 7                       | BAD63242 endo-beta-1,3-gluc                                  | anase                                                                                                                | 109.00                     | 2.30E-22                         | 31%               | 45%               | 83343                   | 31–277             |
|              |                    |                |               |      | 8                 | cd02175      | the cataloter residues.  GH16 [Licheaus, Licheause, allos known as 1,3-1,4-beta-glucianase, is a member of glycosyl hydroiase family 16, that specifically cleaves 1,4-beta plucosatic on similar specifically cleaves 1,4-beta plucosatic on mixed-linked beta glucans that also contain 1,3-beta-plucosatic initiages. Natural substrates of beta-glucansase are beta more contained to the plucosatic initiages. Natural substrates of beta-glucansase are beta mores, Certaria instancia. This protein is found not only in bacteria but also in anaerobic fungi. This domain includes two seven-stranded antiparalle blact-ahestis that are adjacent to one another forming a rowward sollowed by the contained of the con | 38.74          | 0.000637             | 29%           | 46%           | 126281           | 30157          | 8                       | BAC06195 1,3-(1,3;1,4)-beta-(                                | D-glucan 3(4)-glucanohydrolase                                                                                       | 107.84                     | 5.12E-22                         | 32%               | 44%               | 89347                   | 34–263             |
|              |                    |                |               |      | 9                 | cd02177      | calisations that make up the call valid of marine and algae such as<br>Rhodophycaean Kappa-carageanese exist in beteria belonging to at<br>least three phylogenetically distant branches, including<br>pseudoalteromans, plantchonycetes, and baceroldetes. This domain<br>adopts a curved beta-sandwich conformation, with a tunnel-shaped<br>active site active referred to as a latitural field.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 35.47          | 0.005539             | 30%           | 47%           | 81212            | 3117           | 9                       | AAO08191 Beta-glucanase/Be                                   |                                                                                                                      |                            | 2.54E-21                         | 28%               | 42%               | 49343                   |                    |
|              |                    |                |               |      |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                |                      |               |               |                  |                | 10                      | BAC96166 conserved hypothe                                   | etical protein                                                                                                       | 104.38                     | 5.67E-21                         | 28%               | 42%               | 49343                   | 23-344             |
| M128R        | 60357-62039        | 561            | 64,127        | 7.38 | 8 1               | COG1215      | COG1215, Glycosyltransferases, probably involved in cell wall binnenesis (Cell envelone binnenesis outer membrane) Chitin_synth_2, Chitin synthase. Members of this family are fungal chitin                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | 91.53          | 8.00E-20             | 22%           | 42%           | 33-483           | 10420          | 1                       | AAD26641 hyaluronan synthat                                  | se                                                                                                                   | 858.98                     | 0.00E+00                         | 70%               | 83%               | 1561                    | 8568               |
|              |                    |                |               |      | 2                 | pfam03142    | synthase EC:2.4.1.16 enzymes. They catalyse chitin synthesis as<br>? follows: UDP-N-acetyl-D-glucosamine + {(1,4)-(N-acetyl-beta-D-glucosaminyl)}(N) <=> UDP + {(1,4)-(N-acetyl-beta-D-glucosaminyl)}(N)+1)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 58.00          | 8.14E-10             | 23%           | 44%           | 186345           | 193-373        | 2                       | NP_048446 PBCV-1 hyaluronic                                  | acid synthetase                                                                                                      | 857.06                     | 0.00E+00                         | 70%               | 82%               | 1-561                   | 8568               |
|              |                    |                |               |      | 3                 | pfam00535    | Glycos_transf_2, Glycosyl transferase. Diverse family, transferring sugar from UDP-glucose, UDP-N-acetyl- galactosamine, GDP-mannose or CDP-abequose, to a range of substrates including cellulose, dolichol phosphate and teichoic acids.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | 44.83          | 9.01E-06             | 24%           | 44%           | 86283            | 2167           | 3                       | AAD26643 hyaluronan synthat                                  | se                                                                                                                   | 855.13                     | 0.00E+00                         | 70%               | 82%               | 1561                    | 8568               |
|              |                    |                |               |      |                   |              | bridsbridge and telephole acids                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                |                      |               |               |                  |                | 4 5                     | NP_037285 hyaluronan synthas<br>NP_005319 hyaluronan synthas | se 2<br>se 2                                                                                                         | 224.17<br>223.79           | 9.43E-57<br>1.23E-56             | 30%<br>30%        | 49%<br>49%        | 7530<br>7530            | 12-532<br>12-532   |
|              |                    |                |               |      |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                |                      |               |               |                  |                | 6                       | BAC37733 unnamed protein p<br>NP_032242 hyaluronan synthat   | roduct                                                                                                               | 223.79<br>223.79           | 1.23E-56<br>1.23E-56             | 30%<br>30%        | 49%<br>49%        | 7530<br>7530            | 12-532<br>12-532   |
|              |                    |                |               |      |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                |                      |               |               |                  |                | 8<br>9<br>10            | XP_528222 PREDICTED: simil<br>NP_776504 hyaluronan synthas   | lar to hyaluronan synthase 2<br>se 2                                                                                 | 223.79<br>222.63<br>222.25 | 1.23E-56<br>2.74E-56<br>3.58E-56 | 30%<br>30%<br>29% | 49%<br>49%<br>49% | 7-530<br>7-530<br>7-530 | 24-544<br>12-532   |
| M133R        | 62232-63188        | 319            | 35,900        | 8.34 | 4 1               | pfam01331    | mRNA_cap_enzyme, mRNA capping enzyme, catalytic domain. This family represents the ATP binding catalytic domain of the mRNA capping enzyme.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 113.84         | 1.26E-26             | 27%           | 44%           | 50-223           | 1192           | 1                       |                                                              |                                                                                                                      | 331.26                     | 2.50E-89                         | 53%               | 69%               | 2-314                   | 12-324             |
|              |                    |                |               |      | 2                 | COG5226      | [RNA processing and modification]                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | 90.48          | 1.33E-19             | 22%           | 37%           |                  | 39362          | 2                       |                                                              | Of Guanylylated Mrna Capping Enzyme Complexed                                                                        |                            | 1.24E-88                         | 52%               | 69%               | 2-314                   |                    |
|              |                    |                |               |      | 3                 | pfam03919    | mRNA cap C, mRNA capping enzyme, C-terminal domain<br>CDC9, ATP-dependent DNA ligase [DNA replication, recombination, and                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | 47.66<br>35.75 | 1.08E-06<br>0.004237 | 26%<br>24%    | 43%<br>41%    |                  | 3109<br>215355 | 3                       |                                                              | pylytransferase                                                                                                      | 74.71<br>72.79             | 4.21E-12<br>1.60E-11             | 27%<br>26%        | 43%<br>42%        | 44311<br>44311          | 221-479<br>220-478 |
|              |                    |                |               |      | -                 | 0001100      | repairl.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 00.70          | 0.004207             | 2470          | 41.2          | 121 201          | 210 000        | 5                       | AAT68133 mRNA capping ena                                    | zyme                                                                                                                 | 72.40                      | 2.09E-11                         | 23%               | 45%               | 8317                    | 234566             |
|              |                    |                |               |      |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                |                      |               |               |                  |                | 6                       | NP 974263 mRNA guanylyltr                                    | ransferase/ phosphoprotein phosphatase/ protein                                                                      |                            | 2.09E-11<br>2.31E-10             | 24%<br>22%        | 43%<br>42%        | 9317<br>33318           |                    |
|              |                    |                |               |      |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                |                      |               |               |                  |                | . 8                     | tvrosine/serine/thre                                         | onine phosphatase<br>zyme alpha subunit (mRNA guanylyltransferase) (GTP-                                             |                            | 5.15E-10                         | 25%               | 39%               | 47316                   |                    |
|              |                    |                |               |      |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                |                      |               |               |                  |                | 9                       | RNA quanvivitrans<br>Chain B. Structur                       | terase)(Glase) re Of An Mrna Capping Enzyme Bound To The                                                             |                            | 1.15E-09                         | 25%               | 38%               | 47316                   |                    |
|              |                    |                |               |      |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                |                      |               |               |                  |                | 10                      |                                                              |                                                                                                                      |                            | 1.50E-09                         | 24%               | 42%               | 47317                   | 276-571            |
| M136L        | 63572-63216        | 119            | 13,803        | 8.00 | 8                 | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                |                      |               |               |                  |                | 1                       | NP_049001 A645R                                              |                                                                                                                      | 80.88                      | 1.22E-14                         | 33%               | 57%               | 1–117                   | 1121               |
| M137L        | 64469-63633        | 279            | 31,701        | 5.23 | 3 1               |              | UCH, Ubiquitin carboxyl-terminal hydrolase Peptidase_C19_Peptidase C19_Contains ubiquitinyl hydrolases. They are intracellular peptidases that remove ubiquitiny molecules from polyubiquinated peptidase by cleavage of isopeptide bonds. They hydrolyse bonds involving the carboxyl group of the C-terminal Gly                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 107.79         | 9.38E-25             | 20%           | 40%           | 4275             | 6312           | 1                       |                                                              | carboxy-terminal hydrolase active sites; similar to<br>arboxy-terminal hydrolase, coresponds to Swiss-Prot<br>Q09879 | 261.54                     | 1.96E-68                         | 44%               | 67%               | 1–277                   | 1-280              |
|              |                    |                |               |      | 2                 |              | residue of ubiquitin The purpose of the de-ubiquitination is thought to be<br>editing of the ubiquitin conjugates, which could rescue them from<br>degradation, as well as recycling of the Ubiquitin. The<br>ubiquitin/proteasing conference is provided to the control to the<br>ubiquitin/proteasing collegates of the<br>largest families of the<br>largest families of an anticases in the human reasone.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | 78.81          | 5.43E-16             | 17%           | 33%           | 4276             | i 2320         | 2                       | ,                                                            |                                                                                                                      | 60.08                      | 8.66E-08                         | 23%               | 41%               | 4-278                   | 68-347             |
| M139L        | 6548364548         | 312            | 34,925        | 7.50 | 0 1               | COG1405      | SUA7, Transcription initiation factor TFIIIB, Brf1 subunit/Transcription initiation factor TFIIB [Transcription].                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | 79.58          | 3.15E-16             | 22%           | 42%           | 4431             | 7282           | 1                       |                                                              | ccus woesei factor TFIIB homolog, corresponds to<br>n Number X70668                                                  |                            | 3.39E-51                         | 35%               | 57%               | 30312                   | 1290               |
|              |                    |                |               |      |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                |                      |               |               |                  |                | 2                       | . AAT93251 YPR086W                                           | on initiation factor IIB (Swiss Prot. accession number                                                               | 51.22<br>51.22             | 4.80E-05<br>4.80E-05             | 21%<br>21%        | 35%<br>35%        | 44304<br>44304          | 27-316<br>27-316   |
|              |                    |                |               |      |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                |                      |               |               |                  |                | 4                       | P29055)<br>CAG61530 unnamed protein p                        | roduct                                                                                                               |                            | 8.19E-05                         | 20%               | 36%               | 44304                   | 28-317             |
| M140L        | 6599065487         | 168            | 19,353        | 9.14 | 4                 | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                |                      |               |               |                  |                | 1                       | NP_048457 A109L                                              |                                                                                                                      | 128.64                     | 6.63E-29                         | 58%               | 73%               | 66168                   | 1103               |
| M143L        | 6785566041         | 605            | 67,784        | 10.8 | 8 1               | smart00220   | S_TKc, Serine/Threonine protein kinases, catalytic domain;<br>Phosphotransferases. Serine or threonine-specific kinase subfamily.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | 47.90          | 1.00E-06             | 33%           | 48%           | 63257            | 4151           | 1                       | NP_048636 similar to PBCV-1<br>Number U17055                 | 1 ORF A34R, corresponds to GenBank Accession                                                                         | 483.80                     | 7.23E-135                        | 44%               | 59%               | 1605                    | 15-568             |
|              |                    |                |               |      | 2                 |              | Pkinase, Protein kinase domain  S_TKc, Serine/Threonine protein kinases, catalytic domain.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | 46.43          | 2.48E-06             | 29%           | 50%           | 63-257           | 4151           | 2                       | NP_048632 similar to bovine ci                               | ylicin I, corresponds to Swiss-Prot Accession Number                                                                 | 459.91                     | 1.12E-127                        | 42%               | 57%               | 1-539                   | 1608               |
|              |                    |                |               |      | 3                 | cd00180      | Phosphotransferases of the serine or threonine-specific kinase subfamily.<br>The enzymstic society of these protein kinases is controlled by<br>phosphorylation of specific residues in the activation segment of the<br>catalytic domain, sometimes combined with reversible conformational<br>catalytic domain, sometimes combined with reversible conformational<br>APH, Phosphotransferase enzyme family. This family consists of<br>bacterial antibiotic resistance proteins, which confer resistance to various                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 44.43          | 1.02E-05             | 32%           | 48%           | 63-25            | 5152           | 3                       | NP_048970 RPQT-like (9x)                                     |                                                                                                                      | 311.23                     | 6.43E-83                         | 46%               | 64%               | 11344                   | 7336               |
|              |                    |                |               |      | 4                 | pfam01636    | aminoglycosides they include:—minoglycoside 3':—phosphotransferase or kanamycin kinsae / neomycin-haramycin phosphotransferase and streptomycin 3&apos.&aposkinsae or testpomycin 3&apos.&aposkinsae or steptomycin 3&apos.&aposkinsae or testpomycin 3&apos.&aposkinsae or testpomycin 3&apos.&aposkinsae or the aminoglycoside steptomycin 3&apos.&apos.apos.apos.apos.apos.apos.apos.apos.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | 40.52          | 0.000146             | 44%           | 67%           | 225251           | 171198         |                         | Accession Number                                             |                                                                                                                      | 83.96                      | 1.67E-14                         | 52%               | 69%               | 453539                  | 185                |
| M148L        | 00474 07001        |                | 40.50*        |      | 5                 | COG2334      | IGeneral function prediction only!                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 39.56          | 0.000339             | 40%           | 57%           | 225255           | 200-230        | 5                       | ALL ACCUMULATION                                             | virus PBCV-1 ORF A282L, corresponds to GenBank 1142580                                                               | 66.63                      | 2.76E-09                         | 79%               | 89%               | 566604                  | 139                |
|              | 68171-67884        |                | 10,594        | 8.4  |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                |                      |               |               |                  |                |                         | No Hit Found No Hit Found                                    |                                                                                                                      |                            |                                  |                   |                   |                         |                    |
| M149R        | 6826068730         | 157            | 17,621        | 5.00 | 3                 | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                |                      |               |               |                  |                |                         |                                                              | virus CVK2 unknown ORF, corresponds to GenBank                                                                       |                            | 3.21E-34<br>2.72E-25             | 51%<br>39%        | 72%<br>67%        | 16153<br>3152           |                    |
| M150R        | 68760-69071        | 104            | 12,366        | 10.9 | 5                 | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                |                      |               |               |                  |                | 1<br>2<br>3             | NP_048469 A121R<br>! AAL73473 TIr 6Fp protein                |                                                                                                                      | 60.85                      | 1.24E-27<br>1.29E-08<br>2.21E-08 | 73%<br>44%<br>42% | 86%<br>64%<br>64% | 33104<br>3290<br>3290   |                    |
|              |                    |                |               |      |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                |                      |               |               |                  |                | ,                       |                                                              |                                                                                                                      | 00.00                      | 00                               |                   | 0-70              | OL 30                   | 0 <u>2</u> 00      |

| Gene<br>Name | Genome<br>Position | A.A.<br>length | Peptide<br>Mw | pl   | CDD Hit<br>Number | COGs             | COG Definition                                                                                                                                                                                                                                                                                                                                                                                                         | Bit<br>Score | E-value              | %<br>Identity | %<br>Positive | Query<br>from-to | fram to | BLASTp<br>Hit<br>Number | Hit Accession BL AAL73477 Tir 6Fp protein                                             | ASTp Definition                                                                            | Score            | E-value  <br>2.07E-06            | %<br>dentity P | %<br>ositive | Query<br>from-to<br>32–90 | Hit from-<br>to<br>3288    |
|--------------|--------------------|----------------|---------------|------|-------------------|------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|----------------------|---------------|---------------|------------------|---------|-------------------------|---------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------|------------------|----------------------------------|----------------|--------------|---------------------------|----------------------------|
| M152R        | 6910874000         | 1631           | 167.109       | 5.1  | 0                 | No Hit Found     |                                                                                                                                                                                                                                                                                                                                                                                                                        |              |                      |               |               |                  |         | 1                       | NP 048470 PBCV-1 Vp260 protein                                                        |                                                                                            |                  | 8.96E-128                        | 36%            | 47%          | 2-970                     | 3-962                      |
|              |                    |                |               |      |                   |                  |                                                                                                                                                                                                                                                                                                                                                                                                                        |              |                      |               |               |                  |         | 2                       | contains a leucine zinner n                                                           | notif; similar to Bacillus subtilis phage P2A                                              |                  | 7.40E-98                         | 35%            | 46%          | 2812                      | 3809                       |
|              |                    |                |               |      |                   |                  |                                                                                                                                                                                                                                                                                                                                                                                                                        |              |                      |               |               |                  |         | 3                       | NP_048471 preneck appendage prote<br>Number P07537                                    | in, corresponds to Swiss-Prot Accession                                                    |                  | 2.76E-44                         | 37%            |              | 14081629                  |                            |
|              |                    |                |               |      |                   |                  |                                                                                                                                                                                                                                                                                                                                                                                                                        |              |                      |               |               |                  |         | 5                       | BAB83471 Vp260 like protein<br>BAB83467 Vp260 like protein                            |                                                                                            | 176.41<br>169.09 | 7.52E-42<br>1.20E-39             | 22%<br>22%     | 37%<br>37%   | 28787<br>11825<br>8637    | 32-861<br>9-838<br>152-798 |
|              |                    |                |               |      |                   |                  |                                                                                                                                                                                                                                                                                                                                                                                                                        |              |                      |               |               |                  |         | 7                       | BAB83469 Vp260 like protein<br>BAB83470 Vp260 like protein                            |                                                                                            | 162.16           | 2.26E-38<br>1.47E-37             | 24%<br>24%     | 38%<br>39%   | 8643                      | 152-805                    |
|              |                    |                |               |      |                   |                  |                                                                                                                                                                                                                                                                                                                                                                                                                        |              |                      |               |               |                  |         | 9<br>10                 | BAB83468 Vp260 like protein<br>XP 637068 hypothetical protein DDB02                   | 15928                                                                                      | 158.30<br>115.16 | 2.12E-36<br>2.06E-23             | 23%<br>19%     | 38%<br>34%   | 8637<br>351459            | 152-798<br>4411786         |
|              |                    |                |               |      |                   |                  |                                                                                                                                                                                                                                                                                                                                                                                                                        |              |                      |               |               |                  |         |                         | <del>-</del>                                                                          | 0014 ODE 40001                                                                             | 109.77           | 8.64E-22                         | 39%            |              | 14481590                  |                            |
| M160R        | 74021-75307        | 429            | 48,922        | 10.9 | 0                 | No Hit Found     |                                                                                                                                                                                                                                                                                                                                                                                                                        |              |                      |               |               |                  |         | 1                       | NP_049032 similar to Chlorella virus PE<br>Accession Number U42580                    | 3CV-1 ORF A282L, corresponds to GenBank<br>A34R, corresponds to GenBank Accession          |                  | 7.32E-56                         | 36%            | 53%          | 52423                     |                            |
|              |                    |                |               |      |                   |                  |                                                                                                                                                                                                                                                                                                                                                                                                                        |              |                      |               |               |                  |         |                         |                                                                                       | presponds to Swiss-Prot Accession Number                                                   |                  | 1.81E-38<br>1.26E-31             | 57%<br>51%     | 72%<br>70%   | 10141<br>6141             |                            |
|              |                    |                |               |      |                   |                  |                                                                                                                                                                                                                                                                                                                                                                                                                        |              |                      |               |               |                  |         | 4                       | NP 048441 similar to Chlamydia histo                                                  | one-like protein, corresponds to GenBank                                                   |                  | 1.26E-31<br>2.55E-16             | 41%            | 54%          | 10143                     |                            |
|              |                    |                |               |      |                   |                  |                                                                                                                                                                                                                                                                                                                                                                                                                        |              |                      |               |               |                  |         |                         | Accession Number D71563                                                               |                                                                                            |                  |                                  |                |              |                           |                            |
| M163L        | 75853-75314        | 180            | 20,783        | 9.3  | 4 :               | 1 pfam01096      | TFIIS, Transcription factor S-II (TFIIS)  ZnF_C2C2, C2C2 Zinc finger, Nucleic-acid-binding motif in                                                                                                                                                                                                                                                                                                                    |              | 2.80E-12<br>2.17E-11 | 68%<br>56%    | 74%           | 141179           | 1-39    |                         | BAA04187 transcription elongation fact                                                |                                                                                            |                  | 4.91E-63<br>1.87E-62             | 62%            | 70%<br>71%   | 1180                      | 1180                       |
|              |                    |                |               |      |                   |                  | transcriptional elongation factor TFIIS and RNA polymerases  RPB9, DNA-directed RNA polymerase, subunit M/Transcription                                                                                                                                                                                                                                                                                                | 63.38        |                      |               | 77%           |                  |         |                         |                                                                                       | sin; similar to Chlorella virus CVU1 TFIIS-like<br>onds to GenBank Accession Number D29631 |                  |                                  |                |              |                           |                            |
|              |                    |                |               |      | \$                | 3 COG1594        | elongation factor TFIIS (Transcription).                                                                                                                                                                                                                                                                                                                                                                               | 59.66        | 3.12E-10             | 56%           | 75%           | 143179           | 75111   | 3                       | S47662 transcription elongation fact<br>BAA04186 transcription elongation fact        | or TFIIS homolog - Chlorella virus CV-U1                                                   | 238.81<br>204.53 | 5.43E-62<br>1.13E-51             | 61%<br>57%     | 70%<br>68%   | 1180<br>1164              | 1180<br>1164               |
|              |                    |                |               |      |                   |                  |                                                                                                                                                                                                                                                                                                                                                                                                                        |              |                      |               |               |                  |         | 5                       | CAG98214 unnamed protein product<br>AAS54500 AGR011Wp                                 |                                                                                            | 74.33<br>74.33   | 1.77E-12<br>1.77E-12             | 33%<br>35%     | 45%<br>44%   | 59179<br>66179            |                            |
|              |                    |                |               |      |                   |                  |                                                                                                                                                                                                                                                                                                                                                                                                                        |              |                      |               |               |                  |         | 7                       | XP_762035 hypothetical protein UM058<br>CAG58585 unnamed protein product              | 88.1                                                                                       | 72.40<br>71.63   | 6.73E-12<br>1.15E-11             | 30%<br>36%     | 42%<br>45%   | 14179<br>82179            | 145313                     |
|              |                    |                |               |      |                   |                  |                                                                                                                                                                                                                                                                                                                                                                                                                        |              |                      |               |               |                  |         | 9<br>10                 | CAG86394 unnamed protein product<br>CAC19733 tfs1                                     |                                                                                            | 71.63<br>71.25   | 1.15E-11<br>1.50E-11             | 31%<br>70%     | 43%<br>82%   | 52179<br>140179           | 161288                     |
| M166R        | 7595876617         | 220            | 24,359        | 10.2 | 8                 | No Hit Found     |                                                                                                                                                                                                                                                                                                                                                                                                                        |              |                      |               |               |                  |         | 1                       | NP 048475 A127R                                                                       |                                                                                            | 291.58           | 1.13E-77                         | 63%            | 76%          | 8-217                     | 30-245                     |
| M167L        | 7737976624         | 252            | 28,463        | 8.4  | 5                 | No Hit Found     |                                                                                                                                                                                                                                                                                                                                                                                                                        |              |                      |               |               |                  |         | 1                       | NP_048357 A9R                                                                         |                                                                                            | 203.37           | 5.28E-51                         | 53%            | 77%          | 87252                     | 8173                       |
|              |                    |                |               |      |                   |                  |                                                                                                                                                                                                                                                                                                                                                                                                                        |              |                      |               |               |                  |         |                         |                                                                                       | CV-1 ORF A450R, corresponds to GenBank                                                     | 101.68           | 2.16E-20                         | 29%            | 45%          | 8-251                     | 13-253                     |
|              |                    |                |               |      |                   |                  |                                                                                                                                                                                                                                                                                                                                                                                                                        |              |                      |               |               |                  |         |                         |                                                                                       | 5R, encoded by GenBank Accession Number                                                    | 91.66            | 2.24E-17                         | 25%            | 46%          | 8251                      | 7246                       |
|              |                    |                |               |      |                   |                  |                                                                                                                                                                                                                                                                                                                                                                                                                        |              |                      |               |               |                  |         | 4                       | NP_048629 similar to PBCV-1 ORF A<br>Number U17055<br>NP_048525 A177R                 | A79R, corresponds to GenBank Accession                                                     | 88.97<br>80.11   | 1.45E-16<br>6.74E-14             | 24%<br>26%     | 44%<br>45%   | 9-250<br>2-235            | 11-248<br>4-226            |
|              |                    |                |               |      |                   |                  |                                                                                                                                                                                                                                                                                                                                                                                                                        |              |                      |               |               |                  |         | 6                       | NP_048427 A79R<br>AAU06304 hypothetical protein A275R                                 |                                                                                            | 77.03            | 5.70E-13<br>9.73E-13             | 23%<br>25%     | 45%<br>51%   | 9-230<br>87-250           | 8218                       |
|              |                    |                |               |      |                   |                  |                                                                                                                                                                                                                                                                                                                                                                                                                        |              |                      |               |               |                  |         | 8                       | AAU06301 hypothetical protein A275R<br>AAU06302 hypothetical protein A275R            |                                                                                            | 75.10            | 2.17E-12<br>2.93E-09             | 25%<br>26%     | 51%<br>53%   | 87250<br>119251           | 4167<br>1132               |
| M170R        | 77524-77889        | 122            | 13,843        | 8.1  | 4                 | No Hit Found     |                                                                                                                                                                                                                                                                                                                                                                                                                        |              |                      |               |               |                  |         | 1                       | NP_048396 A48R                                                                        |                                                                                            |                  | 1.80E-18                         | 38%            | 61%          | 8-121                     | 5122                       |
| M171L        | 7814777896         | 84             |               | 3.5  |                   | No Hit Found     |                                                                                                                                                                                                                                                                                                                                                                                                                        |              |                      |               |               |                  |         |                         | No Hit Found No Hit Found                                                             |                                                                                            |                  |                                  |                |              |                           |                            |
| M173R        | 78230-78679        | 150            |               | 4.7  |                   | No Hit Found     |                                                                                                                                                                                                                                                                                                                                                                                                                        |              |                      |               |               |                  |         | 1                       | AAZ21386 hypothetical protein SAR11                                                   | 0565                                                                                       | 94.36            | 1.07E-18                         | 39%            | 61%          | 1117                      | 1115                       |
| M174L        | 8003478661         | 458            | 51,959        | 5.8  |                   | 1 COG0637        | COG0637, Predicted phosphatase/phosphohexomutase [General                                                                                                                                                                                                                                                                                                                                                              | 121 27       | 7.78E-29             | 31%           | 47%           | 0 100            | 2190    |                         | CAH09375 putative nucleotidyltransfera                                                |                                                                                            | 206.46           | 1.56F-51                         | 48%            | 64%          | 220450                    | 2-236                      |
| WITTE        | 00034-70001        | 450            | 31,838        | 3.0  |                   |                  | function prediction only).  Gph, Predicted phosphatases [General function prediction only].                                                                                                                                                                                                                                                                                                                            | 101.37       | 7.60E-23             | 29%           | 46%           | 7196             |         |                         | ZP 00368786 lipopolysaccharide biosynth                                               |                                                                                            |                  | 1.12E-49                         | 47%            | 63%          | 220450                    |                            |
|              |                    |                |               |      |                   |                  | Hydrolase, haloacid dehalogenase-like hydrolase. This family are structurally different from the alpha/ beta hydrolase family (pfam00561).                                                                                                                                                                                                                                                                             |              |                      |               |               |                  |         |                         |                                                                                       |                                                                                            |                  |                                  |                |              |                           |                            |
|              |                    |                |               |      | 3                 |                  | This family includes L-2-haloacid dehalogenase, epoxide hydrolases and phosphalases. The structure of the family consists of two domains. One is an inserted four helix bundle, which is the least well conserved region of the alignment. The rest of the fold is composed of the core alpha/beta                                                                                                                     | 90.43        | 1.42E-19             | 25%           | 39%           | 10190            | 1191    | 3                       | ZP_00371454 lipopolysaccharide biosynth                                               | esis protein, putative                                                                     | 196.82           | 1.23E-48                         | 49%            | 63%          | 220450                    | 1236                       |
|              |                    |                |               |      |                   |                  | domain<br>COG1011, Predicted hydrolase (HAD superfamily) [General function                                                                                                                                                                                                                                                                                                                                             | 77.64        | 1.07E-15             | 25%           | 37%           | 7203             | 2208    | 4                       | AAT91796 putative nucleotidyl transfer                                                | ase family protein                                                                         | 196.05           | 2.10E-48                         | 44%            | 63%          | 220450                    | 2240                       |
|              |                    |                |               |      |                   |                  | RfbA, dTDP-glucose pyrophosphorylase [Cell envelope biogenesis,                                                                                                                                                                                                                                                                                                                                                        | 62.88        |                      | 25%           | 42%           | 223446           |         | 5                       |                                                                                       |                                                                                            |                  | 3.14E-44                         | 44%            | 59%          | 217450                    | 2240                       |
|              |                    |                |               |      | 6                 | 6 COG1208        | outer membranel.  GCD1, Nucleoside-diphosphate-sugar pyrophosphorylase involved in lipopolysaccharide biosynthesis/translation initiation factor 2B, gamma/epsilon subunits (elF-2Bgamma/elF-2Bepsilon) [Cell envelope                                                                                                                                                                                                 | 58.05        | 8.08E-10             | 25%           | 42%           | 222447           | 4228    | 6                       | NP_790553 lipopolysaccharide biosynth                                                 | esis protein, putative                                                                     | 173.33           | 1.46E-41                         | 40%            | 57%          | 220450                    | 2239                       |
|              |                    |                |               |      |                   |                  | biogenesis, outer membrane / Translation, ribosomal structure and<br>hionenesis1<br>NTP_transferase, Nucleotidyl transferase. This family includes a wide                                                                                                                                                                                                                                                              |              |                      |               |               |                  |         |                         |                                                                                       |                                                                                            |                  |                                  |                |              |                           |                            |
|              |                    |                |               |      | 1                 |                  | range of enzymes which transfer nucleotides onto phosphosucars NagD, Predicted sugar phosphatases of the HAD superfamily                                                                                                                                                                                                                                                                                               | 56.37        | 3.20E-09             | 20%           | 38%           | 223392           |         |                         | ZP_00207801 COG1209: dTDP-glucose p                                                   | yrophosphorylase                                                                           |                  | 3.25E-41                         | 42%            | 59%          | 226450                    | 1229                       |
|              |                    |                |               |      | 8                 | 5 COG0647        | [Carbohydrate transport and metabolism]. LicC. CTP:phosphocholine cytidylyltransferase involved in choline                                                                                                                                                                                                                                                                                                             | 43.29        |                      | 23%           | 36%           | 10187            | 9228    |                         | YP_233720 Nucleotidyl transferase                                                     |                                                                                            |                  | 4.70E-40                         | 38%            | 58%          | 220450                    | 2239                       |
|              |                    |                |               |      | 9                 |                  | phosphorylation for cell surface LPS epitopes [Cell envelope biogenesis, outer membrane].<br>CTP_transf_3, Cytidylyltransferase. This family consists of two main                                                                                                                                                                                                                                                      | 41.05        | 0.00012              | 45%           | 68%           | 220258           | 1-39    | 9                       | AAZ34091 lipopolysaccharide biosynth                                                  | esis protein, putative                                                                     | 164.47           | 6.78E-39                         | 38%            | 55%          | 220450                    | 2239                       |
|              |                    |                |               |      |                   |                  | Cytidylyltransferase activities: 1) 3-deoxy-manno-octulosonate<br>cytidylyltransferase, EC:2.7.7.38 catalysing the reaction:- CTP + 3-deoxy-                                                                                                                                                                                                                                                                           |              |                      |               |               |                  |         |                         |                                                                                       |                                                                                            |                  |                                  |                |              |                           |                            |
|              |                    |                |               |      | 10                | 0 ptam02348      | D-manno-octulosonate (===== diphosphate + CMP-3-deoxy-D-manno-octulosonate, 2) acyineuraminate voctulosonate, 2) acyineuraminate see Ec.2.7.7.43, catalysing the reaction: CTP + N-acyineuraminate === diphosphate + CMP-N-acyineuraminate. Neu&c-yiqyidilytransferse of Mannheimia haemolytica has been characterised describing kinetics and regulation by substrate charme energetic charme and amino-supra dremand | 40.31        | 0.000193             | 19%           | 36%           | 222392           | 2175    | 10                      | BAE48961 Nucleoside-diphosphate-su                                                    | gar pyrophosphorylase                                                                      | 148.29           | 5.03E-34                         | 35%            | 54%          | 220456                    | 2-243                      |
| M177L        | 8148380074         | 470            | 53.951        | 4.9  | 6                 | No Hit Found     | sunstrate charme, enemetic charde and amino-sudar demand                                                                                                                                                                                                                                                                                                                                                               |              |                      |               |               |                  |         | 1                       | ZP_00520036 Methyltransferase FkbM                                                    |                                                                                            | 67.40            | 1.17E-09                         | 29%            | 45%          | 2200                      | 402-595                    |
| M178L        | 82188-81529        | 220            | 25,947        | 7.3  | 0                 | No Hit Found     |                                                                                                                                                                                                                                                                                                                                                                                                                        |              |                      |               |               |                  |         |                         | YP 163203 hypothetical protein ZMO14                                                  | 68                                                                                         | 51.60            | 1.97E-05                         | 33%            | 54%          | 69168                     | 428-533                    |
|              | 82706-82224        |                | 18,288        |      |                   | No Hit Found     |                                                                                                                                                                                                                                                                                                                                                                                                                        |              |                      |               |               |                  |         | 1                       | NP 049018 A662I                                                                       |                                                                                            |                  | 1.17E-53                         | 64%            | 79%          |                           | 16171                      |
|              |                    |                |               |      |                   |                  |                                                                                                                                                                                                                                                                                                                                                                                                                        |              |                      |               |               |                  |         | 2                       | NP_565983 unknown protein<br>BAD87006 unknown protein                                 |                                                                                            | 63.93<br>62.77   | 1.77E-09<br>3.94E-09             | 24%<br>25%     | 49%<br>50%   | 27160<br>34160            | 98-232<br>112-239          |
|              |                    |                |               |      |                   |                  |                                                                                                                                                                                                                                                                                                                                                                                                                        |              |                      |               |               |                  |         | 4<br>5                  | XP_328621 hypothetical protein<br>AAM62733 contains similarity to 22 kDa              | a peroxisomal membrane protein                                                             | 49.29            | 2.64E-05<br>4.51E-05             | 23%<br>25%     | 44%<br>51%   | 34156<br>34135            | 122225                     |
|              |                    |                |               |      |                   |                  |                                                                                                                                                                                                                                                                                                                                                                                                                        |              |                      |               |               |                  |         | 6                       | BAB08278 unnamed protein product<br>NP_568621 unknown protein                         |                                                                                            | 49.29            | 4.51E-05<br>4.51E-05             | 25%<br>25%     | 51%<br>51%   | 34135<br>34135            | 121224                     |
| M180I        | 83379-82906        | 158            | 18.513        | 10.3 | 4                 | No Hit Found     |                                                                                                                                                                                                                                                                                                                                                                                                                        |              |                      |               |               |                  |         | 1                       |                                                                                       | intron-associated endonuclease, corresponds mber P13299                                    |                  | 1.63E-33                         | 52%            | 64%          | 1148                      |                            |
|              | 02000              | .50            | .0,010        | 10.0 |                   |                  |                                                                                                                                                                                                                                                                                                                                                                                                                        |              |                      |               |               |                  |         | 2                       | YP 293795 putative endonuclease                                                       |                                                                                            | 51.99            | 6.47E-06                         | 34%            | 52%          | 892                       | 288                        |
|              |                    |                |               |      |                   |                  |                                                                                                                                                                                                                                                                                                                                                                                                                        |              |                      |               |               |                  |         | 3                       | NP_049007 similar to Chlorella virus PE<br>Accession Number U42580<br>NP_048671 A315L | 3CV-1 ORF A315L, corresponds to GenBank                                                    |                  | 4.19E-05<br>9.34E-05             | 34%<br>26%     | 53%<br>50%   | 993<br>9134               | 285<br>2128                |
| M403D        | 83430-83714        | 0.0            | 10,293        | 7.7  | R                 | No Hit Found     |                                                                                                                                                                                                                                                                                                                                                                                                                        |              |                      |               |               |                  |         |                         | ABA40764 potassium ion channel                                                        |                                                                                            |                  | 9.34E-05                         | 100%           | 100%         | 195                       | 195                        |
| MCOI W       | JUNJU-03/ 14       | 95             | 10,293        | 1.1  | -                 | INO I IIL FOUIIG |                                                                                                                                                                                                                                                                                                                                                                                                                        |              |                      |               |               |                  |         | 2                       |                                                                                       |                                                                                            | 83.19            | 1.91E-47<br>2.52E-15<br>3.29E-15 | 47%<br>46%     | 64%<br>64%   | 683<br>683                | 1692<br>1692               |
|              |                    |                |               |      |                   |                  |                                                                                                                                                                                                                                                                                                                                                                                                                        |              |                      |               |               |                  |         | 4 5                     | AAQ16137 potassium channel protein                                                    | tein                                                                                       | 82.80            | 3.29E-15<br>3.29E-15             | 46%<br>48%     | 64%<br>64%   | 683<br>683                | 1692<br>1692               |
|              |                    |                |               |      |                   |                  |                                                                                                                                                                                                                                                                                                                                                                                                                        |              |                      |               |               |                  |         | -                       |                                                                                       |                                                                                            |                  |                                  |                | -            |                           | -                          |

| Gene<br>Name | Genome<br>Position | A.A.<br>length | Peptide<br>Mw |       | CDD Hit<br>Number | COGs         | COG Definition                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | Bit<br>Score    | E-value              | %<br>Identity P | %<br>ositive | Query<br>from-to | Hit<br>from-to | BLASTp<br>Hit<br>Number         | Hit<br>Accession                                                           | BLASTp Definition                                                                                | Bit<br>Score                                                          |                                                                                                          | %<br>Identity P                                             |                                                      | from-to                                                                               | Hit from-<br>to      |
|--------------|--------------------|----------------|---------------|-------|-------------------|--------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------|----------------------|-----------------|--------------|------------------|----------------|---------------------------------|----------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------|-------------------------------------------------------------|------------------------------------------------------|---------------------------------------------------------------------------------------|----------------------|
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                 |                      |                 |              |                  |                | 6<br>7<br>8                     | AAQ16141 potassium chi<br>AAQ16142 potassium chi<br>AAQ16138 potassium chi | annel protein                                                                                    | 82.03                                                                 | 4.29E-15<br>5.61E-15<br>7.32E-15                                                                         | 44%<br>46%<br>44%                                           | 65%<br>64%<br>64%                                    | 683<br>683                                                                            | 1692<br>1692<br>1692 |
| M184L        | 84274-83936        | 113            | 12,957        | 10.01 |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                 |                      |                 |              |                  |                | 1                               | NP_048487 A139L                                                            |                                                                                                  | 113.24                                                                | 2.20E-24                                                                                                 | 55%                                                         | 79%                                                  | 27107                                                                                 | 27-107               |
| M186R        | 84369-85928        | 520            | 58,807        | 8.65  | 1                 | COG1215      | COG1215, Glycosyltransferases, probably involved in cell wall                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 133.91          | 1.37E-32             | 20%             | 35%          | 58460            | 1412           | 1                               | NP 048829 similar to Ac                                                    | etobacter cellulose synthase, corresponds to Swiss-Prot                                          | 802.74                                                                | 0.00E+00                                                                                                 | 72%                                                         | 84%                                                  | 1520                                                                                  | 1517                 |
|              |                    |                |               |       |                   |              | biodenesis [Cell envelode biodenesis, outer membrane].  Glycos_transf_2, Glycosyl transferase. Diverse family, transferring sugar from UDP-glucose, UDP-N-acetyl-galactosamine, GDP-mannose or CDP and the control of th                                                                               |                 |                      |                 |              |                  |                | _                               |                                                                            | mber P21877                                                                                      |                                                                       |                                                                                                          |                                                             |                                                      |                                                                                       |                      |
|              |                    |                |               |       | 2                 | pfam00535    | Corracequose, or a range or souscaries including ceritose, concluding<br>bhosohate and technolic acids.<br>Cellulose synt, Cellulose synthase. Cellulose, an aggregate of<br>unbranched polymers of beta-1,4-linked glucose residues, is the major<br>component of wood and thus paper, and is synthesised by plants, most<br>algae, some bacteria and fungi, and even some animals. The genes that                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 64.09           | 1.19E-11             | 18%             | 37%          | 118280           | 1164           | 2                               | XP_328932 hypothetical p                                                   | rotein                                                                                           | 436.80                                                                | 8.37E-121                                                                                                | 44%                                                         | 60%                                                  | 21516                                                                                 | 168670               |
|              |                    |                |               |       | 3                 |              | synthesise cellulose in higher plants differ greatly from the well-<br>characterized genes found in Acobacter and Aprobacterium; sp. More<br>correctly designated as Aspos.cellulose synthase catalytic<br>subunitsAspos.plant cellulose synthase (CeAl) potents are integral<br>membrane potentia, approximately 1,000 amino acids in length. In<br>the processor for processor for corrections described to the con-<br>stroyed to the processor for corrections deviced the con-<br>traction of the processor for corrections developed the con-<br>traction of the processor for corrections of the con-<br>traction of the contraction of the con-<br>traction of the contraction of the con-<br>traction of the contraction of the contraction of the con-<br>traction of the contraction of the contraction of the con-<br>traction of the contraction of the contraction of the con-<br>traction of the contraction of the contraction of the con-<br>traction of the contraction of the contraction of the con-<br>traction of the contraction of the contraction of the contraction of the con-<br>traction of the contraction of the cont | 43.36           | 2.30E-05             | 30%             | 46%          | 292387           | 517618         | 3                               | ZP_00545426 Cellulose synt                                                 | thase (UDP-forming)                                                                              | 276.56                                                                | 1.45E-72                                                                                                 | 33%                                                         | 52%                                                  | 23497                                                                                 | 60-550               |
|              |                    |                |               |       | 4                 | COG1216      | COG1216, Predicted glycosyltransferases [General function prediction only].                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | 42.81           | 3.38E-05             | 21%             | 34%          | 190312           | 82213          |                                 | NP_790863 cellulose synt                                                   |                                                                                                  |                                                                       | 5.74E-37                                                                                                 | 34%                                                         | 51%                                                  | 111367                                                                                |                      |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                 |                      |                 |              |                  |                | 5<br>6                          | AAQ60348 cellulose synt<br>AAL71842 WssB                                   |                                                                                                  | 157.53<br>154.84                                                      | 9.80E-37<br>6.35E-36                                                                                     | 34%<br>34%                                                  | 51%<br>50%                                           | 115385<br>113367                                                                      | 264540<br>157417     |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                 |                      |                 |              |                  |                | 8                               | AAM38361 celullose synt<br>NP_643825 celullose synt                        | hase                                                                                             | 154.45<br>154.45                                                      | 8.30E-36<br>8.30E-36                                                                                     | 32%<br>32%                                                  | 49%<br>49%<br>49%                                    | 115405<br>115405                                                                      | 153-451<br>155-453   |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                 |                      |                 |              |                  |                | 9<br>10                         | CAJ25375 cellulose synt<br>CAI87218 Cellulose synt                         |                                                                                                  | 154.07<br>153.30                                                      | 1.08E-35<br>1.85E-35                                                                                     | 32%<br>30%                                                  | 47%                                                  | 115405<br>115449                                                                      |                      |
| M190R        | 85998-89204        | 1069           | 113,278       | 10.81 | 1                 | nfam05887    | Trypan_PARP, Procyclic acidic repetitive protein (PARP). This family consists of several Trypanosoma brucei procyclic acidic repetitive protein (PARP) like sequences. The procyclic acidic repetitive protein (parp) genes of Trypanosoma brucei encode a small family of abundant surface                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | 55 37           | 6.06E-09             | 32%             | 56%          | 9511010          | 64123          | 1                               | BAE02830 surface protei                                                    | _                                                                                                | 1631 31                                                               | 0.00E+00                                                                                                 | 89%                                                         | 89%                                                  | 1949                                                                                  | 1953                 |
|              |                    |                | ,             |       |                   |              | proteins whose expression is restricted to the procyclic form of the<br>parasite. They are found at two unlinked loci, parpA and parpB;<br>transcription of both loci is developmentally regulated.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |                 |                      |                 |              |                  |                |                                 | an acceptotes                                                              | "                                                                                                |                                                                       |                                                                                                          |                                                             |                                                      |                                                                                       |                      |
|              |                    |                |               |       | 2                 | pfam05616    | Neisseria_TspB, Neisseria meningitidis TspB protein. This family consists of several Neisseria meningitidis TspB virulence factor proteins                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | 50.13           | 2.28E-07             | 38%             | 39%          | 9531009          | 329-385        | 2                               | NP_048488 PBCV-1 surfa                                                     | ice protein                                                                                      | 737.26                                                                | 0.00E+00                                                                                                 | 46%                                                         | 59%                                                  | 1-865                                                                                 | 1888                 |
|              |                    |                |               |       | 3                 | pfam04625    | DEC-1_N, DEC-1 protein, N terminal region. The defective chorion-1 gene (dec-1) in Drosophila encodes follicle cell proteins necessary for proper eggshell assembly. Multiple products of the dec-1 gene are formed by alternative RNA splicing and proteolytic processing. Cleavage                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | 44.49           | 9.77E-06             | 41%             | 41%          | 9511002          | 94145          | 3                               | BAD22850 surface protei                                                    | in                                                                                               | 734.18                                                                | 0.00E+00                                                                                                 | 45%                                                         | 59%                                                  | 1865                                                                                  | 1888                 |
|              |                    |                |               |       |                   |              | formed by alternative KNA spicing and proteolytic processing. Cleavage products include S80 (80 kDa) which is incorporated into the eggshell, and further proteolysis of S80 aives S80 (60 kDa).  TonB, Periplasmic protein TonB, links inner and outer membranes [Cell                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |                 |                      |                 |              |                  |                |                                 |                                                                            |                                                                                                  |                                                                       |                                                                                                          |                                                             |                                                      |                                                                                       |                      |
|              |                    |                |               |       | 4                 |              | envelope biogenesis, outer membranel.  DedD, Uncharacterized protein conserved in bacteria [Function                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                 | 3.27E-05             | 36%             | 43%          | 9311007          |                | 4                               |                                                                            |                                                                                                  |                                                                       | 0.00E+00                                                                                                 | 45%                                                         | 58%                                                  | 1865                                                                                  | 1888                 |
|              |                    |                |               |       | 5                 |              | unknown1<br>CAP, CAP protein                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | 41.55<br>39.99  | 8.14E-05<br>0.00021  | 33%             | 36%<br>37%   | 9461024          | 219-326        | 5                               |                                                                            | rotein A145R - Chlorella virus PBCV-1<br>n, PAPK (10x); similar to wheat Pro-, Lys-rich protein, |                                                                       | 1.35E-12<br>2.73E-05                                                                                     | 28%<br>43%                                                  | 43%<br>58%                                           | 469698<br>890947                                                                      | 2236<br>206263       |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | 30.00           | 0.00021              | 30 /6           | 37 /0        | 007007           | 210-020        |                                 | corresponds t                                                              | to GenBank Accession Number X52472                                                               |                                                                       |                                                                                                          |                                                             |                                                      |                                                                                       |                      |
| M200L        | 8958689278         | 103            | 11,555        | 5.19  |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                 |                      |                 |              |                  |                | 1                               | NP_048498 A150L                                                            |                                                                                                  | 75.49                                                                 | 5.10E-13                                                                                                 | 38%                                                         | 60%                                                  | 1100                                                                                  | 5105                 |
| M201R        | 89691-91052        | 454            | 50,601        | 8.41  | 1 2               |              | SSL2, DNA or RNA helicases of superfamily II [Transcription / DNA reolication. recombination. and reoair].  DEXDc, DEAD-like helicases superfamily;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 140.23<br>63.71 | 1.66E-34<br>1.91E-11 | 26%<br>23%      | 40%<br>37%   | 54397<br>94233   |                | 1                               | NP_048501 Similar to pn<br>Number P111<br>NP 077551 EsV-1-66               | age T5 helicase, corresponds to Swiss-Prot Accession 107                                         |                                                                       | 9.30E-174<br>3.10E-44                                                                                    | 65%<br>32%                                                  | 81%<br>50%                                           | 3453<br>21426                                                                         | 6459<br>17443        |
|              |                    |                |               |       |                   |              | DEXHo, DEXH-box helicases. A diverse family of proteins involved in<br>ATP-dependent DNA or RNA unwinding, needed in a variety of cellular                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |                 |                      |                 | 34%          |                  |                |                                 |                                                                            |                                                                                                  |                                                                       |                                                                                                          |                                                             |                                                      |                                                                                       |                      |
|              |                    |                |               |       | 3                 |              | processes. The name derives from the sequence of the Walker B motif (motif II). This domain contains the ATP-binding region. DEXDc, DEAD-like helicases superfamily. A diverse family of proteins involved in ATP-dependent RNA or DNA unwinding. This domain                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 55.47           | 5.62E-09<br>1.96E-08 | 23%             | 37%          | 110230           | 2143           |                                 | YP_142750 VV A18 helica                                                    | ase                                                                                              |                                                                       | 1.22E-40<br>7.17E-33                                                                                     | 29%                                                         | 48%                                                  | 4430<br>85430                                                                         | 9453<br>419776       |
|              |                    |                |               |       | 5                 |              | contains the ATP-binding region  HsdR. Type I site-specific restriction-modification system. R (restriction)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | 44.19           | 1.23E-05             | 26%             | 40%          |                  | 190-321        |                                 | YP_437404 DNA or RNA                                                       |                                                                                                  | 141.35                                                                | 6.07E-32                                                                                                 | 29%                                                         | 46%                                                  | 87-430                                                                                |                      |
|              |                    |                |               |       | 6                 |              | subunit and related nelicases liberense mechanismsi.  SNF2_N, SNF2 family N-terminal domain. This domain is found in proteins involved in a variety of processes including transcription regulation (e.g., SNF2, STH1, brahma, MOT1). DNA repair (e.g.,                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | 42.24           | 2.39E-05             | 22%             | 39%          | 110 220          | 18156          |                                 | BAB34728 hypothetical p                                                    |                                                                                                  | 125.06                                                                | 2.55E-30                                                                                                 | 30%                                                         | 49%                                                  | 91 420                                                                                | 417-777              |
|              |                    |                |               |       |                   |              | ERCOS, RADIO, RAZUS, DIVAR ECONIDINATION (E.G., RADISH), after<br>chromatin unwinding (e.g., ISWI) as well as a variety of other proteins<br>with little functional information (e.g. Indextar ETI 1)<br>DEAD, DEAD/DEAH box helicases. Members of this family include the<br>DEAD and DEAH box helicases. Helicases are involved in unwinding                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |                 |                      |                 |              |                  |                |                                 |                                                                            |                                                                                                  |                                                                       |                                                                                                          |                                                             |                                                      |                                                                                       |                      |
|              |                    |                |               |       | 7                 |              | nucleic acids. The DEAD box helicases are involved in various aspects of RNA metabolism, including nuclear transcription, pre mRNA splicing, ribosome biogenesis, nucleocytoplasmic transport, translation, RNA decay and organellar nene expression                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | 43.10           | 2.71E-05             | 18%             | 32%          | 102233           | 19175          |                                 | NP_287072 putative helica                                                  |                                                                                                  |                                                                       | 2.55E-30                                                                                                 | 30%                                                         | 49%                                                  | 81430                                                                                 | 425-785              |
|              |                    |                |               |       | 8                 |              | . Mid. Transcription-repair coupling factor (superfamily il helicase) [DNA reclication, recombination, and moail / Transcription! IELICG. Helicase superfamily c-terminal domain; associated with DEXPO., DEAP, and DEAH-took proteins, yeast initiation factor 4A, Ski2p, and Hepatitis C virus NS3 helicases; this domain is found in a wide variety of helicases and helicase related proteins; may not be an                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 42.17           | 5.73E-05             | 26%             | 43%          | 89-230           | 607-752        | 8                               | terriirai                                                                  | iction enzyme, res subunit:DEAD/DEAH box helicase, N-                                            | 131.34                                                                | 6.28E-29                                                                                                 | 29%                                                         | 45%                                                  | 81430                                                                                 | 481-849              |
|              |                    |                |               |       | 9                 |              | autonomously folding unit, but an integral part of the helicase; 4 helicase superfamilies at present according to the organization of their signature motifs; all helicases share the ability to unwind nucleic acid duplexes with a distinct directional polarity, they utilize the free energy from nucleoside triphosphate hydrolysis to fuel their translocation along DNA, unwinding                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 40.68           | 0.000164             | 23%             | 40%          | 302395           | 17123          | 9                               | EAM63429 Type III restri<br>terminal                                       | iction enzyme, res subunit:DEAD/DEAH box helicase, N-                                            | 129.80                                                                | 1.83E-28                                                                                                 | 30%                                                         | 47%                                                  | 51430                                                                                 | 394-790              |
|              |                    |                |               |       | 10                | COG0610      | COG0610, Type I site-specific restriction-modification system, R (restriction) subunit and related helicases [Defense mechanisms].                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | 39.66           | 0.000293             | 23%             | 36%          | 95-338           | 258-517        | 10                              | ZP_00679441 DEAD/DEAH<br>enzyme. res s                                     | box helicase:Helicase, C-terminal:Type III restriction subunit                                   | 117.86                                                                | 7.19E-25                                                                                                 | 29%                                                         | 46%                                                  | 82430                                                                                 | 427786               |
| M204L        | 91463-91113        | 117            | 12,749        | 3.86  |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                 |                      |                 |              |                  |                | 1                               | NP_048505 A157L                                                            |                                                                                                  | 95.13                                                                 | 6.32E-19                                                                                                 | 44%                                                         | 56%                                                  | 1117                                                                                  | 1109                 |
| M206L        | 91741-91523        | 73             | 8,893         | 3.28  |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                 |                      |                 |              |                  |                |                                 | No Hit Found No Hit Found                                                  |                                                                                                  |                                                                       |                                                                                                          |                                                             |                                                      |                                                                                       |                      |
| M208L        | 92392-91994        | 133            | 14,349        | 10.12 |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                 |                      |                 |              |                  |                |                                 | No Hit Found No Hit Found                                                  |                                                                                                  |                                                                       |                                                                                                          |                                                             |                                                      |                                                                                       |                      |
| M210L        | 92872-92588        | 95             | 10,895        | 10.03 |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                 |                      |                 |              |                  |                |                                 | NP 048513 A165L                                                            |                                                                                                  |                                                                       | 2.60E-12                                                                                                 | 39%                                                         | 54%                                                  | 595                                                                                   | 171-279              |
| M213L        | 93392-92937        |                | 17,762        | 10.22 |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                 |                      |                 |              |                  |                | 1                               | NP_048513 A165L                                                            |                                                                                                  |                                                                       | 7.63E-17                                                                                                 | 37%                                                         | 60%                                                  | 5118                                                                                  |                      |
| M215R        | 93462-94298        | 279            | 32,168        | 5.13  |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                 |                      |                 |              |                  |                | 1<br>2<br>3<br>4<br>5<br>6<br>7 | 1 DOV 1 CAO!!                                                              | ein<br>recursor<br>ein                                                                           | 298.90<br>93.97<br>78.95<br>78.95<br>77.41<br>77.41<br>77.03<br>71.63 | 1.11E-79<br>5.41E-18<br>1.80E-13<br>1.80E-13<br>5.24E-13<br>5.24E-13<br>6.85E-13<br>2.88E-11<br>2.88E-11 | 59%<br>29%<br>30%<br>30%<br>30%<br>30%<br>30%<br>28%<br>29% | 73%<br>50%<br>46%<br>46%<br>45%<br>45%<br>46%<br>46% | 1-245<br>13-213<br>28-229<br>28-229<br>28-216<br>28-216<br>11-211<br>27-208<br>28-213 |                      |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                 |                      |                 |              |                  |                | 10                              | XP_472944 OSJNBa0081                                                       | L15.20                                                                                           | 70.86                                                                 | 4.91E-11                                                                                                 | 28%                                                         | 46%                                                  | 27208                                                                                 | 141-324              |
| M218R        | 94325-94807        | 161            | 17,492        | 4.72  |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                 |                      |                 |              |                  |                | 1                               | NP_048516 A168R                                                            |                                                                                                  | 127.49                                                                | 1.30E-28                                                                                                 | 51%                                                         | 64%                                                  | 26157                                                                                 | 24161                |

| Gene<br>Name | Genome<br>Position | A.A.<br>length | Peptide<br>Mw | pl    | CDD Hit<br>Number | COGs         | COG Definition                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | Bit<br>Score     | E-value I            | %<br>dentity F | %<br>Positive | Query<br>from-to | Hit<br>from-to  | BLASTp<br>Hit<br>Number | Hit<br>Accession     | BLASTp Definition                                                                                     | Bit<br>Score     | E-value                                      | %<br>Identity P          |                          | Query<br>from-to                 | Hit from-<br>to   |
|--------------|--------------------|----------------|---------------|-------|-------------------|--------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------|----------------------|----------------|---------------|------------------|-----------------|-------------------------|----------------------|-------------------------------------------------------------------------------------------------------|------------------|----------------------------------------------|--------------------------|--------------------------|----------------------------------|-------------------|
| M219L        | 95631-94810        | 274            | 30,278        | 9.89  | 1                 | pfam01734    | Patatin, Patatin-like phospholipase. This family consists of various<br>patatin glycoproteins from plants. The patatin protein accounts for up to<br>40% of the total soluble protein in potato tubers. Patatin is a storage<br>protein but it also has the enzymatic activity of lipid acyl hydrolase,                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | 116.20           | 2.61E-27             | 30%            | 50%           | 16189            | 1179            | 1                       | NP_048521            | similar to E. coli hypothetical protein, corresponds to Swiss-Prot Accession Number P39407            | 323.55           | 4.06E-87                                     | 57%                      | 77%                      | 5274                             | 14-288            |
|              |                    |                |               |       |                   |              | catalysing the cleavage of fatty acids from membrane lipids. Members of this family have been found also in vertebrates RssA, Predicted esterase of the alpha-beta hydrolase superfamily                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |                  |                      |                |               |                  |                 |                         |                      |                                                                                                       |                  |                                              |                          |                          |                                  |                   |
|              |                    |                |               |       | 2                 |              | COG4667, Predicted esterase of the alpha-beta hydrolase superfamily                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 101.31<br>51.03  | 7.43E-23<br>1.04E-07 | 31%<br>26%     | 50%<br>42%    |                  | 14193           |                         |                      | Patatin-like phospholipase family     PREDICTED: hypothetical protein XP_795682, partial              | 107.84<br>98.98  | 3.50E-22<br>1.63E-19                         | 34%<br>35%               | 53%<br>50%               | 15185<br>14186                   |                   |
|              |                    |                |               |       | 3                 | 0004007      | [General function prediction only].                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 31.03            | 1.042-07             | 2070           | 42.70         | 10-200           | 14202           | 4                       |                      | hypothetical protein                                                                                  | 89.74            | 9.88E-17<br>2.87E-16                         | 28%<br>30%               | 48%<br>56%               | 14256<br>14186                   | 6261              |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                  |                      |                |               |                  |                 | 5<br>6<br>7             | AAQ66865             | 463L     conserved hypothetical protein     phospholipase, patatin family                             | 88.20            | 2.87E-16<br>2.87E-16<br>3.75E-16             | 32%<br>31%               | 52%<br>49%               | 14185<br>14186                   | 6194              |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                  |                      |                |               |                  |                 | 8                       | XP_789091            | PREDICTED: hypothetical protein XP 783998  esterase of the alpha-beta hydrolase superfamily-like      |                  | 3.75E-16<br>6.40E-16                         | 32%<br>31%               | 51%<br>50%               | 14186<br>15185                   | 86275             |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                  |                      |                |               |                  |                 | 10                      | YP_142800            | patatin-like phospholipase (463L)                                                                     | 82.80            | 1.21E-14                                     | 27%                      | 49%                      | 15238                            | 57-285            |
|              |                    |                |               |       |                   |              | S_TKc, Serine/Threonine protein kinases, catalytic domain.<br>Phosphotransferases of the serine or threonine-specific kinase subfamily.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                  |                      |                |               |                  |                 |                         |                      |                                                                                                       |                  |                                              |                          |                          |                                  |                   |
| M221L        | 96553-95654        | 300            | 34,092        | 4.51  | 1                 |              | The enzymatic activity of these protein kinases is controlled by phosphorylation of specific residues in the activation segment of the catalytic domain, sometimes combined with reversible conformational changes in the C-terminal autoregulatory tail                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 176.55           | 1.66E-45             | 32%            | 54%           | 36295            | 1256            | 1                       | AAU06282             | Protein kinase A248R                                                                                  | 193.74           | 5.66E-48                                     | 38%                      | 59%                      | 21295                            | 32–305            |
|              |                    |                |               |       | 2                 |              | S_TKc, Serine/Threonine protein kinases, catalytic domain;<br>Phosphotransferases, Serine or threonine-specific kinase subfamily                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 176.56           | 1.90E-45             | 31%            | 52%           | 37-295           | 1256            | 2                       |                      | protein kinase A248R                                                                                  |                  | 9.65E-48                                     | 38%                      | 59%                      | 21295                            |                   |
|              |                    |                |               |       | 3                 |              | Pkinase, Protein kinase domain<br>SPS1, Serine/threonine protein kinase [General function prediction only /<br>Signal transduction mechanisms / Transcription / DNA replication,                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 165.46<br>117.18 | 3.71E-42<br>1.39E-27 | 28%            | 51%<br>47%    | 37-295<br>36-296 | 1258            | 3                       |                      | protein kinase A248R  PBCV-1 protein kinase                                                           | 192.97           | 9.65E-48<br>8.17E-47                         | 38%                      | 59%<br>58%               | 21295                            |                   |
|              |                    |                |               |       | 5                 |              | Segina transduction mechanisms / Transcription / DNA replication, recombination, and repair).  TyrKc, Tyrosine kinase, catalytic domain; Phosphotransferases.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                  | 1.32E-20             | 28%            | 48%           |                  | 3-192           | 5                       |                      | serine/threonine protein kinase                                                                       |                  | 4.05E-46                                     | 37%                      | 58%                      | 24295                            |                   |
|              |                    |                |               |       | 3                 | amartouz ra  | Tvrosine-specific kinase subfamilv.  TyrKc, Tyrosine kinase, catalytic domain. Phosphotransferases; tyrosine-specific kinase subfamily. Enzymes with TyrKc domains belong to an extensive family of proteins which share a conserved catalytic core                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 54.05            | 1.321-20             | 20%            | 4070          | 38-221           | 3182            | 3                       | AAA07000             | у зеннелитеонне рготент клазе                                                                         | 107.30           | 4.002-40                                     | 37 /6                    | 30 /6                    | 24-285                           | 20-300            |
|              |                    |                |               |       | 6                 | cd00192      | common to both serine/threonine and tyrosine protein kinases. Enzymatic<br>activity of tyrosine protein kinases is controlled by phosphorylation of<br>specific tyrosine residues in the activation segment of the catalytic                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | 91.40            | 7.55E-20             | 27%            | 46%           | 36227            | 8203            | 6                       | AAU06274             | protein kinase A248R                                                                                  | 162.54           | 1.40E-38                                     | 38%                      | 60%                      | 79295                            | 22-238            |
|              |                    |                |               |       | _                 |              | domain or a C-terminal tyrosine (tail) residue with reversible conformational changes COG3642, Mn2+-dependent serine/threonine protein kinase [Signal                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |                  |                      |                |               |                  |                 | _                       |                      |                                                                                                       |                  |                                              |                          |                          |                                  |                   |
|              |                    |                |               |       | 7 8               |              | transduction mechanismsi. COG4248, Uncharacterized protein with protein kinase and helix-hairpin-                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                  | 2.26E-08<br>0.000635 | 28%            | 41%<br>37%    | 112217           |                 |                         |                      | protein kinase A248R<br>similar to PBCV-1 serine/threonine protein kinase, corresponds to             |                  | 3.44E-37<br>3.81E-28                         | 37%<br>31%               | 59%<br>53%               | 81295<br>41295                   |                   |
|              |                    |                |               |       | ۰                 |              | helix DNA-binding domains l'General function prediction onlyl.  Kdo, Lipopolysaccharide kinase (Kdo/WaaP) family. These lipopolysaccharide kinases are related to protein kinases pfam00069.  This family includes waaP (rfaP) neap product is required for the addition.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | 30.54            | 0.000635             | 20%            | 3776          | 127214           | 114-200         | ٥                       | NF_048643            | GenBank Accession Number U14660                                                                       | 127.07           | 3.01E-20                                     | 3176                     | 5376                     | 41-295                           | 22-219            |
|              |                    |                |               |       | 9                 | pfam06293    | of phosphate to O-4 of the first heptose residue of the lipopolysaccharide (LPS) inner core region. It has previously been shown that WaaP is necessary for resistance to hydrophobic and polycationic antimicrobials in E. coli and that it is required for virulence in invasive strains of S.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 35.63            | 0.004574             | 21%            | 36%           | 13–167           | 1155            | 9                       | AAU06286             | protein kinase A248R                                                                                  | 127.10           | 6.50E-28                                     | 47%                      | 63%                      | 146295                           | 2152              |
|              |                    |                |               |       | 10                | smart00750   | KIND, kinase non-catalytic C-lobe domain; It is an interaction domain identified as being similar to the C-terminal protein kinase catalytic fold (C lobe). Its presence at the N terminus of signalling proteins and the absence of the active-site residues in the catalytic and activation loops suggest that it folds independently and is likely to be non-catalytic. The occurrence of KIND point in metazona implies that it has evolved from the                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 35.35            | 0.006041             | 22%            | 41%           | 135297           | 22175           | 10                      | NP_048631            | similar to PBCV-1 serine/threonine protein kinase, corresponds to GenBank Accession Number U14660     | 123.25           | 9.39E-27                                     | 31%                      | 51%                      | 31292                            | 13–273            |
|              |                    |                |               |       |                   |              | occulrence of initial only in Irrelazation injuries start, it are environed in in ordinary catalytic profesion kinase domain into an interaction domain possibly by keening the substrata-hinding features.  COG4591, Superfamily II RNA helicase [DNA replication, recombination, ordinary in the substrata in the subs |                  |                      |                |               |                  |                 |                         |                      | contains ATP-GTP binding motif; similar to Saccharomyces cervisiae                                    |                  |                                              |                          |                          |                                  |                   |
| M225L        | 98791-96647        | 715            | 80,084        | 7.18  | 1                 |              | and repairj.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |                  | 1.26E-106            | 40%            | 55%<br>49%    |                  | 119568          |                         |                      | antiviral protein SKI2, corresponds to Swiss-Prot Accession Number<br>P35207                          |                  | 0.00E+00                                     | 49%                      | 67%                      | 3715                             |                   |
|              |                    |                |               |       | 2                 | COG1204      | COG1204, Superfamily II helicase [General function prediction only].<br>COG1202, Superfamily II helicase, archaea-specific [General function prediction only]                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | 206.78<br>130.08 | 1.49E-54<br>1.62E-31 | 32%<br>31%     | 52%           | 25403<br>25381   | 46430<br>231552 | 2                       |                      | DEAD/DEAH box helicase, putative PREDICTED: similar to ENSANGP0000020973                              | 318.55<br>311.23 | 4.90E-85<br>7.83E-83                         | 38%<br>38%               | 57%<br>58%               | 7454<br>12431                    | 91–567<br>123–578 |
|              |                    |                |               |       | 4                 | pfam00270    | DEAD, DEAD/DEAH box helicase. Members of this family include the<br>DEAD and DEAH box helicases. Helicases are involved in unwinding<br>nucleic acids. The DEAD box helicases are involved in various aspects of<br>RNA metabolism, including nuclear transcription, pre mRNA splicing,                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | 113.21           | 2.17E-26             | 21%            | 40%           | 10193            | 9206            | 4                       | CAG30992             | ! hypothetical protein                                                                                | 308.53           | 5.07E-82                                     | 37%                      | 56%                      | 12450                            | 127-598           |
|              |                    |                |               |       | 5                 | cmart00497   | ribosome biogenesis, nucleocytoplasmic transport, translation, RNA decav and organellar cene expression.  DEXDc, DEAD-like helicases superfamily;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | 101.07           | 1.09E-22             | 26%            | 43%           | 4 170            | 1190            | 5                       | EAL 22220            | GA17990-PA                                                                                            | 200 15           | 6.63E-82                                     | 37%                      | 55%                      | 12431                            | 57-513            |
|              |                    |                |               |       | 6                 |              | DEAUC, DEAU-like nelicases superraminy; .  COG1205, Distinct helicase family with a unique C-terminal domain including a metal-binding cysteine cluster [General function prediction                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | 93.54            | 1.09E-22             | 25%            | 41%           |                  | 75410           | 6                       |                      | GA10159-PA                                                                                            |                  | 1.48E-81                                     | 35%                      | 51%                      | 12541                            |                   |
|              |                    |                |               |       | 7                 |              | onlyl.  DEXDc, DEAD-like helicases superfamily. A diverse family of proteins                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | 81.29            | 1.00E-16             | 25%            | 44%           | 27157            | 1143            | 7                       |                      |                                                                                                       | 306.61           | 1.93E-81                                     | 37%                      | 55%                      | 12431                            | 152608            |
|              |                    |                |               |       | 8                 |              | involved in ATP-dependent RNA or DNA unwinding. This domain contains the ATP-binding region  Lhr, Lhr-like helicases [General function prediction only].                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 79.14            | 3.43E-16             | 25%            | 45%           |                  | 27358           | 8                       |                      | CG4152-PA<br>ENSANGP00000020973                                                                       |                  | 3.29E-81                                     | 36%                      | 55%                      | 12456                            |                   |
|              |                    |                |               |       | 9                 |              | DEXHc, DEXH-box helicases. A diverse family of proteins involved in<br>ATP-dependent DNA or RNA unwinding, needed in a variety of cellular                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | 72.42            | 3.94E-14             | 28%            | 45%           | 27157            | 1142            | 9                       | CAE45877             | hypothetical protein                                                                                  | 305.45           | 4.29E-81                                     | 36%                      | 56%                      | 12450                            | 140611            |
|              |                    |                |               |       | 10                | COG0513      | processes. The name derives from the sequence of the Walker B motif (motif II). This domain contains the ATP-binding recion SYMB, Superfamily II DNA and RNA helicases [DNA replication, recombination, and repair / Transcription / Translation, ribosomal                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 69.42            | 3.51E-13             | 20%            | 41%           | 16427            | 56425           | 10                      |                      | SKIV2L2 protein                                                                                       | 305.45           | 4.29E-81                                     | 36%                      | 56%                      | 12450                            | 141612            |
| M232R        |                    | 104            | 40.000        | 40.50 |                   |              | structure and biogenesisl.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |                  |                      |                |               |                  |                 |                         |                      |                                                                                                       |                  |                                              |                          |                          |                                  |                   |
| M232R        | 98893-99204        | 104            | 12,386        | 10.56 |                   | No Hit Found | Homo sperm syn, Homospermidine synthase. This family consists of                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |                  |                      |                |               |                  |                 |                         | NO HIT FOUND         | No Hit Found                                                                                          |                  |                                              |                          |                          |                                  |                   |
| M233L        | 10074199236        | 502            | 56,422        | 9.74  | 1                 | pfam06408    | several homospermidine synthase proteins (EC.2.5.1.44).<br>Homospermidine synthase (HSS) calalyses the synthesis of the<br>polyamine homospermidine from 2 mol putrescine in an NAD(+)-<br>denendent reaction<br>Saccharop, Gh. Saccharopine dehydrogenase. This family comprised of                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | 531.45           | 2.42E-152            | 38%            | 53%           | 27496            | 1470            | 1                       | NP_048585            | PBCV-1 homospermidine synthase                                                                        | 729.55           | 0.00E+00                                     | 69%                      | 83%                      | 21502                            | 36-518            |
|              |                    |                |               |       | 2                 | pfam03435    | three structural domains that can not be separated in the linear<br>sequence. In some organisms this enzyme is found as a bifunctional<br>polypeptide with lysine ketoglutarate reductase. The saccharopine                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 46.01            | 3.46E-06             | 22%            | 47%           | 30185            | 1137            | 2                       | ZP_00589757          | Homospermidine synthase                                                                               | 334.72           | 4.30E-90                                     | 38%                      | 57%                      | 27474                            | 5452              |
|              |                    |                |               |       | 3                 | COG1748      | riehvitronenase can also function as a sanchamoline refurdase. LYSS, Saccharopine dehydrogenase and related proteins (Amino acid transcort and metabolism).  E1_enzyme_family, Superfamily of activating enzymes (E1) of the ubiquitin-like proteins. This family includes classical ubiquitin-activating enzymes E1, ubiquitin-like ubiquitin-activating enzymes E1, ubiquitin-like ubiquitin-like proteins. This family includes classical ubiquitin-activating enzymes E1, ubiquitin-like ubiquitin | 46.05            | 3.71E-06             | 24%            | 49%           | 28-269           | 2184            | 3                       | AAM05046             | homospermidine synthase                                                                               | 305.83           | 2.14E-81                                     | 37%                      | 57%                      | 27474                            | 13-459            |
|              |                    |                |               |       | 4                 | cd01483      | mechanistic homologes, like MoeB, Thiff and others. The common<br>reaction mechanism catalyzed by MoeB, ThiF and the E1 enzymes<br>begins with a nucleophilic attack of the C-terminal carboxylate of MoaD,<br>ThiS and ubiquitin, respectively, on the alpha-phosphate of an ATP<br>molecule bound at the active site of the activating enzymes, leading to                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | 35.02            | 0.007331             | 20%            | 37%           | 30-130           | 2121            | 4                       | AAM29862             | homospermidine synthase                                                                               | 301.60           | 4.04E-80                                     | 36%                      | 57%                      | 27474                            | 6452              |
|              |                    |                |               |       |                   |              | the formation of a high-energy acyladenylate intermediate and<br>subsequently to the formation of a thiocarboxylate at the C termini of<br>MoaD and ThiS                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |                  |                      |                |               |                  |                 |                         | 70.00                |                                                                                                       | 05               | 0.07                                         |                          |                          | oz :-                            |                   |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                  |                      |                |               |                  |                 | 6                       | ZP 00584810          | Homospermidine synthase Homospermidine synthase                                                       | 252.68           | 2.07E-76<br>2.15E-65<br>1.82E-64             | 36%<br>33%               | 55%<br>52%<br>54%        | 27474<br>30474<br>26474          | 14446             |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                  |                      |                |               |                  |                 | 8                       | AAU28555<br>BAE51070 | hypothetical protein<br>homospermidine synthase<br>Homospermidine synthase<br>homospermidine synthase | 248.83<br>234.57 | 1.82E-64<br>3.10E-64<br>6.06E-60<br>4.34E-58 | 34%<br>34%<br>32%<br>34% | 54%<br>54%<br>52%<br>52% | 26474<br>26474<br>33474<br>77478 | 13-446<br>34-462  |
| M237L        | 101090100764       | 109            | 12,662        | 5.20  |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                  |                      |                |               |                  |                 |                         |                      | No Hit Found                                                                                          |                  |                                              |                          |                          |                                  |                   |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                  |                      |                |               |                  |                 |                         |                      |                                                                                                       |                  |                                              |                          |                          |                                  |                   |

| Gene<br>Name | Genome<br>Position | A.A.<br>length | Peptide<br>Mw | pl    | CDD Hit<br>Number | COGs         | COG Definition                                                                                                                                                                                                                                                                                                                                                 | Bit<br>Score | E-value   | %<br>Identity P | %<br>ositive | Query<br>from-to | Hit<br>from-to | BLASTp<br>Hit<br>Number | Hit<br>Accession           | BLASTp Definition                                                                                                                                               | Bit<br>Score               | E-value <sub>I</sub>             | %<br>dentity P    | %<br>ositive      | Query H<br>from-to      | Hit from-<br>to            |
|--------------|--------------------|----------------|---------------|-------|-------------------|--------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|-----------|-----------------|--------------|------------------|----------------|-------------------------|----------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------|----------------------------------|-------------------|-------------------|-------------------------|----------------------------|
| M238R        | 101294101650       | 119            | 13,221        | 10.74 |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                |              |           |                 |              |                  |                | 1                       | NP_048582 A                | 234L                                                                                                                                                            | 93.20                      | 2.38E-18                         | 43%               | 57%               | 1-118                   | 1107                       |
| M240L        | 101915101646       | 90             | 11,406        | 10.51 |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                |              |           |                 |              |                  |                | 1                       | NP_048581 A                | 233R                                                                                                                                                            | 85.50                      | 4.98E-16                         | 47%               | 70%               | 586                     | 1191                       |
| M241L        | 102200101940       | 87             | 10,153        | 4.84  | 1                 | COG0695      | GrxC, Glutaredoxin and related proteins [Posttranslational modification, protein turnover, chaperones].                                                                                                                                                                                                                                                        | 43.03        | 3.20E-05  | 31%             | 53%          | 372              | 4-72           | 1                       | CAE58431 H                 | hypothetical protein CBG01566                                                                                                                                   | 48.91                      | 5.22E-05                         | 30%               | 52%               | 483                     | 31114                      |
| M242L        | 102804102235       | 190            | 20,890        | 10.14 |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                |              |           |                 |              |                  |                | 1                       | NP 048578 A                | 230R                                                                                                                                                            | 157.15                     | 2.38E-37                         | 40%               | 65%               | 2-190                   | 3194                       |
| M243R        | 102829103062       | 78             | 8,683         | 6.49  |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                |              |           |                 |              |                  |                |                         | No Hit Found N             | lo Hit Found                                                                                                                                                    |                            |                                  |                   |                   |                         |                            |
| M246L        | 103535103296       | 80             | 9,502         | 7.85  |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                |              |           |                 |              |                  |                |                         | No Hit Found N             | lo Hit Found                                                                                                                                                    |                            |                                  |                   |                   |                         |                            |
| M247R        | 103586104335       | 250            | 28,319        | 8.21  |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                |              |           |                 |              |                  |                | 1                       | NP 048357 A                | .9R                                                                                                                                                             |                            | 1.46E-61                         | 64%               | 79%               | 79249                   | 3173                       |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                |              |           |                 |              |                  |                | 2                       | NP_049005 A                | imilar to Chlorella virus PBCV-1 ORF A450R, corresponds to GenBank<br>ccession Number U42580<br>imilar to PBCV-1 ORF A275R, encoded by GenBank Accession Number | 103.61<br>99.75            | 5.60E-21                         | 32%               | 47%<br>46%        | 6-248                   | 11-253                     |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                |              |           |                 |              |                  |                |                         | U                          | 142580                                                                                                                                                          |                            | 8.09E-20<br>1.21E-15             | 30%               | 46%<br>44%        | 6-248<br>9-226          | 5246<br>8217               |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                |              |           |                 |              |                  |                | 5                       | NP_048629 <sup>Si</sup>    | 79R imilar to PBCV-1 ORF A79R, corresponds to GenBank Accession lumber U17055                                                                                   |                            | 2.69E-15                         | 27%               | 46%               | 9-247                   | 11-248                     |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                |              |           |                 |              |                  |                | 7                       | AAU06304 h                 | ypothetical protein A275R                                                                                                                                       | 58.15                      | 3.08E-11<br>2.70E-07             | 26%<br>25%        | 44%<br>44%        | 9-233<br>81-247         | 11-227<br>1167             |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                |              |           |                 |              |                  |                | 8                       | AAU06301 h                 | ypothetical protein A275R                                                                                                                                       |                            | 6.01E-07                         | 25%               | 44%               | 81247                   | 1167                       |
| M249R        | 104467104880       | 138            | 15,412        | 9.27  |                   | No Hit Found | AceF_Pyrivate/2-oxodutarate dehydrogenase complex                                                                                                                                                                                                                                                                                                              |              |           |                 |              |                  |                | 1                       | NP_048923 A                | .567L                                                                                                                                                           | 77.41                      | 1.34E-13                         | 38%               | 52%               | 12127                   | 10-123                     |
| M253L        | 106765104888       | 626            | 67,332        | 8.25  | 1                 |              | dihydrolipoamide acyltransferase (E2) component, and related enzymes                                                                                                                                                                                                                                                                                           | 41.13        | 0.000118  | 20%             | 34%          | 336438           | 77176          | 1                       | NP_048921 A                | 565R                                                                                                                                                            | 261.54                     | 6.10E-68                         | 41%               | 59%               | 25336                   | 26-336                     |
|              |                    |                |               |       | 2                 | COG3883      | COG3883, Uncharacterized protein conserved in bacteria [Function unknown].                                                                                                                                                                                                                                                                                     | 40.32        | 0.000214  | 21%             | 47%          | 12176            | 5170           |                         |                            | NA binding protein                                                                                                                                              | 259.23                     | 3.03E-67                         | 41%               | 59%               | 25336                   | 26-336                     |
|              |                    |                |               |       | 3                 | pfam03276    | Gag_spuma, Spumavirus gag protein                                                                                                                                                                                                                                                                                                                              | 39.68        | 0.000275  | 27%             | 40%          | 333433           | 178273         | 3                       | NP_048741 L                | ys-, Pro-rich, PAPK (10x); similar to wheat Pro-, Lys-rich protein,<br>orresponds to GenBank Accession Number X52472                                            | 63.93                      | 1.87E-08                         | 45%               | 53%               | 251312                  | 2999                       |
| M258R        | 106836108620       | 595            | 64,077        | 10.10 | 1                 | COG3064      | ToIA, Membrane protein involved in colicin uptake [Cell envelope biogenesis. outer membrane].                                                                                                                                                                                                                                                                  | 44.39        | 1.17E-05  | 24%             | 37%          | 111305           | 56253          | 1                       | NP_048917 si<br>G          | imilar to Chlorella virus CVK2 DNA binding protein, corresponds to<br>senBank Accession Number D78305                                                           | 208.38                     | 5.74E-52                         | 45%               | 63%               | 378594                  | 417647                     |
|              |                    |                |               |       | 2                 | pfam06519    | ToIA, ToIA protein. This family consists of several bacterial ToIA proteins<br>as well as two eukaryotic proteins of unknown function. Tol proteins are<br>involved in the translocation of group A colicins. Colicins are bacterial<br>protein toxins, which are active against Escherichia coli and other related                                            | 43.96        | 1.46E-05  | 24%             | 40%          | 145339           | 85269          | 2                       | BAA11342 D                 | INA binding protein                                                                                                                                             | 204.91                     | 6.35E-51                         | 44%               | 62%               | 378594                  | 416-646                    |
|              |                    |                |               |       |                   |              | species (See pfam01024). TolA is anchored to the cytoplasmic membrane by a single membrane by a single membrane spanning segment near the N-terminus, leaving most of the protein exposed to the periplasm Myosin Lail. T, Myosin tail. The myosin molecule is a multi-subunit complex made up of two heavy chains and four light chains it is a               |              |           |                 |              |                  |                |                         |                            |                                                                                                                                                                 |                            |                                  |                   |                   |                         |                            |
|              |                    |                |               |       | 3                 | pfam01576    | fundamental contractile protein found in all eukaryote cell types. This family consists of the colled-coil myosin heavy chain tail region. The colled-coil is composed of the tail from two molecules of myosin. These can then assemble into the macromolecular thick filament. The colled-coil recion provides the structural blackbone the thick filament.  | 42.99        | 2.99E-05  | 24%             | 42%          | 88-273           | 379550         | 3                       | BAA83789 al                | Iginate lyase                                                                                                                                                   | 95.90                      | 4.15E-18                         | 36%               | 51%               | 410574                  | 153-317                    |
|              |                    |                |               |       | 4                 | COG0845      | AcrA, Membrane-fusion protein [Cell envelope biogenesis, outer<br>membrane].                                                                                                                                                                                                                                                                                   | 42.03        | 5.35E-05  | 16%             | 32%          | 37-257           | 2216           | 4                       | NP_048562 P                | BCV-1 alginate lyase                                                                                                                                            | 94.36                      | 1.21E-17                         | 35%               | 50%               | 410574                  | 141-305                    |
|              |                    |                |               |       | 5                 |              | myosin-like domain [Function unknown]                                                                                                                                                                                                                                                                                                                          | 39.73        |           | 19%             | 35%          | 102297           |                | 5                       | BAB19127 v                 |                                                                                                                                                                 | 93.20                      | 2.69E-17                         | 35%               | 51%               |                         | 169-333                    |
| M263R        | 108657109310       | 218            | 24,292        | 10.15 | 6                 | No Hit Found | CAP, CAP protein                                                                                                                                                                                                                                                                                                                                               | 39.21        | 0.000407  | 25%             | 39%          | 252383           | 192330         | 6                       | NP_048915 A                | nnamed protein product                                                                                                                                          | 55.07<br>73.94             | 8.12E-06<br>3.64E-12             | 30%<br>27%        | 42%<br>43%        | 81277<br>1176           | 567-776<br>1-185           |
| M264I        | 109732109313       | 140            | 14.802        | 4.77  |                   | nfam00692    | dUTPase, dUTPase. dUTPase hydrolyses dUTP to dUMP and                                                                                                                                                                                                                                                                                                          | 131.51       | 6.05E-32  | 52%             | 69%          | 11130            | 2129           | 1                       | ΔΔW51452 de                | eoxyuridine triphosphatase                                                                                                                                      | 218.39                     | 4.99E-56                         | 76%               | 88%               | 2140                    | 3141                       |
| MZO4E        | 100702 100010      | 140            | 14,002        | 4.77  | . 2               |              | nvronhosnhate  Dod, Deoxycytidine deaminase [Nucleotide transport and metabolism].                                                                                                                                                                                                                                                                             |              | 1.33E-08  | 29%             | 48%          |                  | 74158          |                         | ND OLOGON SI               | imilar to tomato dUTP pyrophosphatase corresponds to GenBank                                                                                                    |                            | 6.51E-56                         | 75%               | 89%               | 2-140                   | 3141                       |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                |              |           |                 |              |                  |                | 3                       | AAW51453 de<br>EAL88001 di | ccesssion Number S40549<br>eoxyuridine triphosphatase                                                                                                           | 215.70<br>172.94           | 3.23E-55<br>2.40E-42             | 75%<br>63%        | 87%<br>77%        | 2-140<br>4-140          | 3141<br>114250             |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                |              |           |                 |              |                  |                | 5                       | XP 657875 h                | U I Pase<br>ypothetical protein AN0271.2<br>LTP diphosphatase/ hydrolase                                                                                        | 172.17                     | 4.10E-42<br>2.03E-41             | 63%<br>62%        | 78%<br>76%        | 4-140<br>4-139<br>3-140 | 71-206<br>29-166           |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                |              |           |                 |              |                  |                | 7                       | AAB22611 de                | eoxyuridine trinhosohatase: dUTPase: P18                                                                                                                        | 169.86<br>167.93<br>166.78 | 7.73E-41                         | 62%<br>62%<br>59% | 76%<br>73%<br>74% | 3-140<br>3-140<br>2-140 | 29166<br>32169<br>37175    |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                |              |           |                 |              |                  |                | 9                       | XP 469212 ni               | UTP diphosphatase<br>utative deoxyuridine triphosphatase<br>nnamed protein product                                                                              | 166.78<br>166.39           | 1.72E-40<br>1.72E-40<br>2.25E-40 | 61%<br>61%        | 75%<br>76%        | 5-140<br>4-139          | 37-175<br>90-225<br>58-193 |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                |              |           |                 |              |                  |                | 10                      |                            | imilar to Sulfolobus TATA-binding protein, corresponds to GenBank                                                                                               |                            |                                  |                   |                   |                         |                            |
| M266R        | 109853110671       | 273            | 30,421        | 8.38  | 1                 | pfam00352    | TBP, Transcription factor TFIID (or TATA-binding protein, TBP)                                                                                                                                                                                                                                                                                                 | 37.03        | 0.002008  | 24%             | 48%          | 142237           | 3-85           | 1                       | NP_048908 A<br>YP_142807 T | ccession Number S55311&aoos: ATA-box binding protein (TBP)                                                                                                      |                            | 9.41E-44<br>8.61E-05             | 41%<br>25%        | 59%<br>41%        | 25268<br>75252          | 13-268<br>138-346          |
| M269L        | 111880110681       | 400            | 44,555        | 5.69  | 1                 | pfam04451    | Capsid_Iridovir, Iridovirus major capsid protein. This family includes the major capsid protein of iridoviruses, chlorella virus and Spodoplera assovirus, which are all dsDNA viruses with no RNA stage. This is the most abundant structural protein and can account for up to 45% of virion protein. In Chlorella virus MT325 the major capsid protein is a | 363.48       | 9.75E-102 | 40%             | 57%          | 1396             | 2443           | 1                       | NP_048914 Si<br>A          | imilar to Chlorella virus PBCV-1 ORF A11L, corresponds to GenBank ccession Number U42580                                                                        | 510.38                     | 4.21E-143                        | 60%               | 76%               | 1-400                   | 1400                       |
|              |                    |                |               |       |                   |              | 14414 611 87901                                                                                                                                                                                                                                                                                                                                                |              |           |                 |              |                  |                | 2                       | NP 048359 co               | ontains aminoacyl-tRNA synthetase class-II signature imilar to PBCV-1 major capsid protein, corresponds to Swiss-Prot                                           |                            | 5.19E-109<br>2.12E-86            | 49%<br>40%        | 67%<br>62%        | 2400<br>2398            | 3403<br>4399               |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                |              |           |                 |              |                  |                | 4                       | AAC27492 m                 | imilar to PBCV-1 major capsid protein, corresponds to Swiss-Prot<br>ccession Number P30328<br>najor capsid protein Vp49                                         | 276.17                     | 1.34E-72                         | 38%               | 54%               | 2-400                   | 3432                       |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                |              |           |                 |              |                  |                | 5<br>6                  |                            | najor capsid protein MCP1 BCV-1 major capsid protein Vp54, corresponds to GenBank Accession                                                                     | 271.17<br>270.78           | 4.30E-71<br>5.61E-71             | 39%<br>38%        | 54%<br>54%        | 2400                    | 3437<br>3437               |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                |              |           |                 |              |                  |                | 7                       | BAA76600 m                 | naior cansid protein                                                                                                                                            | 267.70                     | 4.75E-70                         | 37%               | 54%               | 2400                    | 3436                       |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                |              |           |                 |              |                  |                | 8<br>9                  | 1M4X C c                   | najor capsid protein Vp54<br>Chain C, Pbcv-1 Virus Capsid, Quasi-Atomic Model                                                                                   | 261.92<br>240.74           | 2.61E-68<br>6.22E-62             | 37%<br>37%        | 53%<br>53%        | 2400<br>32400           | 3437<br>9413               |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                |              |           |                 |              |                  |                | 10                      | 1M3Y_D C                   | chain D, The Structure Of Major Capsid Protein Of A Large, Lipid containing Dna Virus                                                                           | 240.74                     | 6.22E-62                         | 37%               | 53%               | 32400                   | 9-413                      |
| M272L        | 113304111928       | 459            | 51,475        | 8.82  | 1                 | COG0553      | HepA, Superfamily II DNA/RNA helicases, SNF2 family [Transcription / DNA reolication. recombination. and reoair]. SNF2_N, SNF2 family N-terminal domain. This domain is found in                                                                                                                                                                               | 158.77       | 4.66E-40  | 27%             | 43%          | 8438             | 339851         | 1                       | NP_048904 <sup>Si</sup>    | imilar to Caenorhabditis transcription activator, corresponds to Swiss-<br>rot Accession Number P41877                                                          | 451.44                     | 2.79E-125                        | 51%               | 71%               | 8-457                   | 7458                       |
|              |                    |                |               |       | 2                 | pfam00176    | proteins involved in a variety of processes including transcription regulation (e.g., SNF2, STH1, brahma, MOT1), DNA repair (e.g., ERCC6, RAD16, RAD5), DNA recombination (e.g., RAD54), and chromatin unwinding (e.g., ISWI) as well as a variety of other proteins with little functional information (e.g., Ideatar FEI 11).                                | 126.16       | 2.44E-30  | 24%             | 44%          | 11-262           | 1287           | 2                       | BAC15031 h                 | elicase (swi/snf family)                                                                                                                                        | 168.32                     | 4.71E-40                         | 30%               | 51%               | 8-422                   | 6061027                    |
|              |                    |                |               |       | 3                 |              | SSL2, DNA or RNA helicases of superfamily II [Transcription / DNA                                                                                                                                                                                                                                                                                              | 81.30        | 8.96E-17  | 22%             | 38%          | 7393             | 36375          | 3                       | AAP98807 sv                | wt/snf helicase                                                                                                                                                 | 162.16                     | 3.38E-38                         | 29%               | 47%               | 4437                    | 6981154                    |
|              |                    |                |               |       | 4                 |              | DEXDc, DEAD-like helicases superfamily; . DEXHc DEXH-hox helicases A diverse family of proteins involved in                                                                                                                                                                                                                                                    | 78.73        | 5.89E-16  | 19%             | 36%          | 8193             | 9199           | 4                       | CAG35724 pt                | robable helicase                                                                                                                                                | 161.38                     | 5.76E-38                         | 28%               | 47%               | 2450                    | 9381397                    |
|              |                    |                |               |       | 5                 | cd00269      | ATP-dependent DNA or RNA unwinding, needed in a variety of cellular processes. The name derives from the sequence of the Walker B motif (motif II). This domain contains the ATP-binding region                                                                                                                                                                | 73.58        | 2.00E-14  | 25%             | 49%          | 29164            | 1143           | 5                       | CAC96916 lin               | n1685                                                                                                                                                           | 161.00                     | 7.52E-38                         | 29%               | 47%               | 4443                    | 6121058                    |

| Name Position length Mw P Number COGS Coordinated Score Levelue Identity Positive from-to from-to Number Accession  HELICS, Helicase superfamily containing associated with  FIXTH. DEAD, and DEAD by containing upont initiation feature 4.5                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                                                                                                                                                     | Query Hit from-                                                         |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------|
| DEXDc-, DEAD-, and DEAH-box proteins, yeast initiation factor 4A,                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | Score Score Identity Positive fro                                                                                                                   | om-to to                                                                |
| Ski2p, and Hepatitis Cvirus N5 seliciases; this domain is found in a wide variety of helicases and helicase indeed profilers; may not be an divided profilers and the helicase. To a consider the helicase of the helicase of 23.42 2.2E-14 26% 44% 275-402 3-130 6 NP_300906 SWUSNF fan superfamilies at repearal excercioning or organization of their signature.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | imily helicase_2 159.84 1.88E-37 29% 46%                                                                                                            | 4437 6981154                                                            |
| molfs; all helicases share the ability to unwind nucleic acid duplexes with<br>a distinct directional polarity; they utilize the free energy from uncleoside<br>triphosphate hydrolysis to fuel their translocation along DNA, unwinding<br>the duries in the noncess.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                                                                                                                                                     |                                                                         |
| DEXDC_DEAD-like helicases superfamily. A diverse family of proteins                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | icase/SNF2 family domain protein 159.84 1.68E-37 29% 48% 12 family 159.07 2.86E-37 28% 47%                                                          | 3422 8561294<br>4437 6121053                                            |
| contains the ATP-binding revious.  Helicase C. Helicase conserved C-terminal domain. Tribis domain.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 112 tamily 139.07 2.00E-37 20% 47%                                                                                                                  | 4-437 612-1053                                                          |
| 9 pfam00271 family is found in a wide variety of helicases and helicase related 59.46 3.23E-10 26% 47% 316-395 1-78 9 NP_296465 helicase, Snt proteins. It may be that this is not an automonously folding unit, but an internal part of the helicase.  DEAD, DEAD/DEAD DEAH box helicase. Members of this family include the                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | 1/2 family 159.07 2.86E-37 28% 47%                                                                                                                  | 2437 7111169                                                            |
| DEAD and DEAH box helicases are involved in unwinding 10 pfam00270 rundles acids. The DEAD box helicases are involved in unwinding 40 pfam00270 rundles acids. The DEAD box helicases are involved in various aspects of 4,03 6,79E-06 20% 39% 8–138 11–142 10 ZP_00231434 helicase, Snf. RNA metabolism, including under transcription, por mRNA splicing, 41 pfam00270 rundles control splicing, and the property of the pro | 159.07 2.86E-37 28% 47%                                                                                                                             | 4-437 6121053                                                           |
| M278R 113482—114282 267 29,095 4.11 No Hit Found No Hit Found                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | nd                                                                                                                                                  |                                                                         |
| M282R 114658-114915 86 9,836 10.85 No Ht Found 1 NP_048875 A519L                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 98.60 5.76E-20 60% 72%                                                                                                                              | 780 477                                                                 |
| M283R 114950-116107 386 44,286 9.76 1 COG1231 (Monoamine oxidase [Amino acid transport and metabolism]. 65.77 4.64E-12 18% 30% 1-384 1-439 1 BAA83788 orf1                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | 545.43 1.12E-153 68% 80%                                                                                                                            | 8386 19397                                                              |
| 2 COG1233, Phytoene dehydrogenase and related proteins (Secondary 47 30 1 45E-06 38% 48% 11_63 7_50 2 NP_04864 similar to b                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | bovine monoamine oxidase, corresponds to Swiss-Prot 534.64 1.98E-150 67% 79% Number P21398                                                          | 8-386 16-394                                                            |
| M288L 117169—116114 352 40,610 10.08 No Hit Found No Hit Found                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |                                                                                                                                                     |                                                                         |
| M289R 117311118414 368 40,590 10.22 No HR Found 1 NP_048562 PBCV-1 alain                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |                                                                                                                                                     | 61-366 5-320                                                            |
| 2 BAB19127 vAL-1 3 BAA83789 alginate lyasv                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | se 407.91 2.62E-112 56% 65%                                                                                                                         | 1-366 1-348<br>1-366 1-332                                              |
| 4 BA5415 kypothecial 5 Np_04897 smiles to C                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | Nils II Company Day 1 Company                                                                                                                       | 125361 5247<br>133367 409648                                            |
| 6 BAA11342 DNA-binding                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | g protein 72.02 3.40E-11 25% 42% 16                                                                                                                 | 166367 430647                                                           |
| M294R 118438—119088 217 23,981 8,96 No Hit Found No Hit Found                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                                                                                                                                                     |                                                                         |
| M298R 119184-119654 157 18,097 5.79 No Hit Found 1 NP_048561 A214L                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |                                                                                                                                                     | 22-154 3-135                                                            |
| M299R 119736-120182 149 16,648 4.54 No Hit Found 1 NP_048560 A213L                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |                                                                                                                                                     | 1-126 1-125                                                             |
| MODULE 121199-120253 322 36,45% 3.97 I COS2/1 domain [General function prediction only]. 49.71 2.01E-07 29% 31% 141313 3404007 NO FILE Found                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | nd .                                                                                                                                                |                                                                         |
| M307L 122464—121349 372 41,792 5.47 1 pfam027s4 domain. These psytocoal-dependent decarboxyslase acting on ornithine, 196.63 1.77E-51 38% 57% 25-260 4246 1 NP_048554 PBCV-1 argin fold.  2 C000019 LysA. Disminoprimelate decarboxyslase (Amino acid transport and 192.42 3.35E-50 28% 45% 19-372 29-394 2 CAE02644 ornithine dec                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |                                                                                                                                                     | 1-372 1-372                                                             |
| metabolismi. Orn DAP Ang. deC, Pyridoxal-dependent decarboxylase, C-terminal                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |                                                                                                                                                     |                                                                         |
| 3 pfam00278 sheet domain. These pyridoxal-dependent decarboxylases act on                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |                                                                                                                                                     | 1-364 48-418                                                            |
| 4 COG1166 Spar, Aguine teclariox/pare (sperimune duosymmens) primino duu 54.10 1.47E-08 23% 40% 45-270 118-374 4 BAA83427 omitime dec                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |                                                                                                                                                     | 1-364 47-417<br>1-364 50-420                                            |
| 6 AAL83709 publishe omil 7 AAV1802C omiline dec                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | carboxylase 260.77 5.20E-68 38% 57%                                                                                                                 | 1364 50420<br>1364 48418                                                |
| 8 NP_571876 omithine dec<br>9 AAQ14852 omithine dec                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | carboxylase 258.84 1.98E-67 38% 56%                                                                                                                 | 2-372 21-409<br>1-364 48-418                                            |
| M308L 123252122548 235 25,935 11.57 No Hit Found 1 NP_048552 2095R                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | carboxylase 258.84 1.98E-67 39% 57% 178.72 1.22E-43 47% 60%                                                                                         | 1-364 47-417<br>1-228 1-205                                             |
| M310L 123954-123295 220 24,037 4.83 No HH Found 1 1 Pp.04850 Agoing                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 269.63 4.61E-71 63% 74%                                                                                                                             | 1-220 1-213                                                             |
| M312R 124001-124336 112 12,026 4.52 No Hit Found 1 N P 048549 A202L                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                                                                                                                                                     | 3112 2112                                                               |
| DSPc, Dual specificity phosphatases (DSP); Ser/Thr and Try protein  M313L 124897124343 185 20.826 10.87 1 cd00127 phosphatases. Structurally similar to tyrosine-specific phosphatases but 132.67 2.68E-32 35% 53% 26–159 4–139 1 NP 048659 similar to hu.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | numan protein Tyr-phosphatase, corresponds to GenBank 242.66 4.08E-63 66% 85%                                                                       |                                                                         |
| wm a snalower active sinc ent an a distinctive active sinc signature molif. HOxXxxxx. Characterized as VHz. or Cd25-like. DSPc, Dual specificity phosphatase, catalytic domain. Ser/Thr and Tyr                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | Number U2/193                                                                                                                                       | 6-168 7169<br>24164 159300                                              |
| to fail of systems-specinic prospirateses, except for a recognition rection.  PREFIXETE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | D: similar to Dual specificity protein phosphatase 16 (Mitopen-                                                                                     |                                                                         |
| 3 smart00195 DSPc, Dual specificity phosphatase, catalytic domain; 122.69 2.95E-29 34% 56% 24-161 2-139 3 XP_543810 activated print MNP-7) MNP-7                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | rotein kinase phosphatase 7) (MAP kinase phosphatase 7) 84.73 1.42E-15 34% 53% 2                                                                    | 24163 159299                                                            |
| 4 COG2463 CDC14, Predicted protein-tyrosine phosphatase [Signal transduction 52.40 3.94E-08 23% 44% 58-181 65-180 4 AAH42101 DUSP16 prot merbranisms! 5 smart00404 PTPC_motify Protein tyrosine phosphatase, catalytic domain motif; 36.17 0.003842 22% 48% 99-147 36-90 5 AAH31643 Unknown (pn                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                                                                                                                                     | 24-163 159-299<br>24-163 110-250                                        |
| 6 AAJ0226 Dual specific<br>7 AAJ0225 Dual specific                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | city phosphatase 16 83.57 3.17E-15 34% 53% 2 city phosphatase 16 83.57 3.17E-15 34% 53% 2                                                           | 24163 159299<br>24163 159299                                            |
| 9 BAB21791 KIA41700 pm                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | orotein 83.57 3.17E-15 34% 53% 2                                                                                                                    | 24163 268408<br>24163 184324                                            |
| 10 AAH59232 Dusp16 prote<br>M314L 125280125006 85 10.248 3.73 No Hit Found 1 NP 048680 A306L                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |                                                                                                                                                     | 24162 159298<br>3485 3586                                               |
| M314L 125260-125006 85 10,248 3.73 No Hit Found 1 NP_048660 A306L<br>M315L 125720-125385 112 13,007 4.87 No Hit Found 1 NP_048683 A308L                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                                                                                                                                                     | 3485 3586<br>4398 2079                                                  |
| M317L 126306-125903 168 18,118 8.10 No Hit Found 1 NP 048665 A310L                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 182.57 3.87E-45 53% 69%                                                                                                                             | 1-168 1-170                                                             |
| 1 11 00000 70102                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | kDa translation peptide 340.50 2.56E-92 71% 82%                                                                                                     | 1235 1230                                                               |
| M319. 127075-128353 241 26,796 5.32 No Hit Found 1 NP_048667 PBCV-1.33M                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | 84.34 3.30E-15 28% 43%<br>57.77 3.31E-07 26% 40%                                                                                                    | 5-236 10-259<br>9-213 10-222<br>1-222 13-247                            |
| M319L 127075128353 241 26,796 5.32 No Hit Found 1 NP_048667 PBCV-1 33kt 2 NP_077561 ExV-1-76 3 AAR29866 FirrV-1:1 4 AAR2885 FirrV-1-B10                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                                                                                                                                                     |                                                                         |
| 2 NP_07561 ExV-1-76<br>3 APA270561 FirV-1-11                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | nd                                                                                                                                                  |                                                                         |
| 2 NP_077561 EsV-1-76<br>3 AAR26966 FirrV-11<br>4 AAR2685 FirrV-1-810                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | nd 132.88 2.72E-30 83% 87%                                                                                                                          | 279 580                                                                 |
| 2 NP_077561 EsV1-176 3 AAR299696 FirtV-1-1 4 AAR2985 FirtV-1-810 M322L 127493—127242 84 9.374 8.18 No Hit Found No Hit Found                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | 132.88 2.72E-30 83% 87%  I membrane protein 91.28 5.46E-17 26% 47% 12 protein, pudative 90.89 7.13E-17 25% 46% 12 protein 78.57 3.6EE-13 24% 52% 17 | 2-79 580<br>129364 18256<br>129364 18256<br>125357 2231<br>118365 14270 |

| Gene<br>Name<br>M328R           | Genome<br>Position<br>129391129729                                                                           | A.A.<br>length                                              | Peptide<br>Mw<br>11,846 |       | CDD Hit<br>Number | COGs No Hit Found    | COG Definition                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | Bit<br>Score   | E-value              | %<br>Identity F | %<br>Positive | Query<br>from-to | Hit<br>from-to       | BLASTp<br>Hit<br>Number | Hit<br>Accession<br>NP_048677 A321R                             | BLASTp Definition                                                               | Bit<br>Score<br>62.00             | E-value lo                                                                       | %<br>lentity Po                               | %<br>ositive                                  | Query I<br>from-to<br>4-95                                  | Hit from-<br>to                                              |
|---------------------------------|--------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------|-------------------------|-------|-------------------|----------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------|----------------------|-----------------|---------------|------------------|----------------------|-------------------------|-----------------------------------------------------------------|---------------------------------------------------------------------------------|-----------------------------------|----------------------------------------------------------------------------------|-----------------------------------------------|-----------------------------------------------|-------------------------------------------------------------|--------------------------------------------------------------|
|                                 | 130346129849                                                                                                 | 166                                                         |                         | 5.50  |                   | No Hit Found         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |                |                      |                 |               |                  |                      | 1                       | NP 048678 A322L                                                 |                                                                                 |                                   | 5.95E-27                                                                         | 38%                                           | 57%                                           | 1159                                                        | 1170                                                         |
| M332L                           | 131644130379                                                                                                 | 422                                                         | 46,792                  | 5.07  | 1                 | pfam01298            | Lipoprotein_5, Transferrin binding protein-like solute binding protein. This family of proteins are distantly related to other families of solute                                                                                                                                                                                                                                                                                                                                                                                                                                           | 34.97          | 0.006974             | 19%             | 37%           | 21               | 16 47167             | 1                       | NP_048680 A324L                                                 |                                                                                 | 396.74                            | 7.28E-109                                                                        | 52%                                           | 62%                                           | 11422                                                       | 13-453                                                       |
|                                 |                                                                                                              |                                                             |                         |       |                   |                      | hinding proteins                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                |                      |                 |               |                  |                      | 3                       | NP_077588 EsV-1-103                                             |                                                                                 | 76.64                             | 4.24E-16<br>1.67E-12<br>2.33E-06                                                 | 29%<br>28%<br>30%                             | 47%<br>50%<br>52%                             | 46254<br>74256<br>140250                                    | 58-270<br>192-365<br>225-333                                 |
| M334R                           | 131667132263                                                                                                 | 199                                                         | 23,258                  | 10.51 |                   | No Hit Found         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |                |                      |                 |               |                  |                      |                         | No Hit Found No Hit Found                                       |                                                                                 |                                   |                                                                                  |                                               |                                               |                                                             |                                                              |
| M337L                           | 133396132341                                                                                                 | 352                                                         | 40,349                  | 9.76  |                   | No Hit Found         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |                |                      |                 |               |                  |                      | 1                       | NP_048684 A328L                                                 |                                                                                 | 266.54                            | 8.78E-70                                                                         | 38%                                           | 60%                                           | 1352                                                        | 1350                                                         |
| M341R                           | 133727134449                                                                                                 | 241                                                         | 26,780                  | 5.31  |                   | No Hit Found         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |                |                      |                 |               |                  |                      | 1<br>2<br>3             | AAR26966 FirrV-1-I1                                             | translation peptide                                                             | 342.04<br>84.34<br>58.54<br>57.00 | 8.79E-93<br>3.30E-15<br>1.94E-07<br>5.65E-07                                     | 71%<br>28%<br>26%<br>26%                      | 82%<br>43%<br>40%<br>40%                      | 1-235<br>5-236<br>9-213<br>1-222                            | 1-230<br>10-259<br>10-222<br>13-247                          |
| Leu<br>Phe<br>Arg<br>Gly<br>Asn | 134513134585<br>134610134693<br>134718134790<br>134814134886<br>134890134960<br>134984135057<br>135080135151 | 73 bs<br>84 bs<br>73 bs<br>73 bs<br>71 bs<br>74 bs<br>72 bs |                         |       |                   |                      | anticodon TAT anticodon TAA anticodon TAA anticodon TCA anticodon TCT anticodon TCC anticodon GTT anticodon GTT                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |                |                      |                 |               |                  |                      |                         |                                                                 |                                                                                 |                                   |                                                                                  |                                               |                                               |                                                             |                                                              |
| Tyr<br>Intron<br>Lys            | 135174135259<br>135211135223<br>135262135334<br>135496135567                                                 | 72 bs<br>86 bs<br>13 bs<br>73 bs<br>72 bs                   |                         |       |                   |                      | anticodon GTA Intron (135211-135223) anticodon CTT anticodon CGT                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                |                      |                 |               |                  |                      |                         |                                                                 |                                                                                 |                                   |                                                                                  |                                               |                                               |                                                             |                                                              |
|                                 | 136141135620                                                                                                 |                                                             | 20,192                  | 6.76  |                   | No Hit Found         | anicodon CO1                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                |                      |                 |               |                  |                      | 1                       | NP 048693 A337L                                                 |                                                                                 | 66.63                             | 3.39E-10                                                                         | 48%                                           | 59%                                           | 100173                                                      | 578                                                          |
| M346L                           | 136605136189                                                                                                 | 139                                                         | 15,952                  | 4.56  |                   | No Hit Found         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |                |                      |                 |               |                  |                      |                         | No Hit Found No Hit Found                                       |                                                                                 |                                   |                                                                                  |                                               |                                               |                                                             |                                                              |
| M348L                           | 138420136744                                                                                                 | 559                                                         | 60,603                  | 9.29  |                   | No Hit Found         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |                |                      |                 |               |                  |                      | 1                       | NP_048699 A342L                                                 |                                                                                 | 734.95                            | 0.00E+00                                                                         | 68%                                           | 78%                                           | 16550                                                       | 37-561                                                       |
| M354R                           | 138534139847                                                                                                 | 438                                                         | 48,898                  | 4.48  | 1                 | COG3534              | AbfA, Alpha-L-arabinofuranosidase [Carbohydrate transport and metabolism].                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | 41.05          | 0.00011              | 26%             | 41%           | 5724             | 7 66256              | 1                       | CAD86595 cellulase precur                                       | sor                                                                             | 54.68                             | 7.16E-06                                                                         | 25%                                           | 36%                                           | 54290                                                       | 282516                                                       |
| M357L                           | 140281139901                                                                                                 | 127                                                         | 14,450                  | 10.21 |                   | No Hit Found         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |                |                      |                 |               |                  |                      | 1<br>2<br>3             | BAA22201 URF14.2                                                |                                                                                 | 147.90<br>147.52                  | 6.26E-35<br>8.18E-35<br>1.07E-34<br>2.38E-34                                     | 66%<br>66%<br>65%<br>64%                      | 78%<br>79%<br>78%<br>78%                      | 2-108<br>2-108<br>2-108<br>2-108                            | 3109<br>3109<br>3109<br>3109                                 |
| M358R                           | 140365140835                                                                                                 | 157                                                         | 18,719                  | 9.88  |                   | No Hit Found         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |                |                      |                 |               |                  |                      | 5                       |                                                                 |                                                                                 | 112.46                            | 3.81E-24<br>1.54E-36                                                             | 73%                                           | 85%<br>76%                                    | 169                                                         | 73–141<br>1–120                                              |
| M359L                           | 140726141079                                                                                                 | 118                                                         | 38,475                  | 8.32  | 1                 |                      | Cyt_CS_DNA_methylase, Cytosine-CS specific DNA methylases; Methyl transfer reactions play an important role in many aspects of biodynose-specific DNA methylases are found both in prokaryotes and eukaryotes. DNA methylases net found both in prokaryotes and eukaryotes. DNA methylasen, or the covalent addition of a methyl group to cytosine within the context of the CpG dirudectide, has profound effects on the mammalian genome. These effects include transcriptional repression via inhibition of transcription factor brinding or the recruitment.                            | 160.08         | 1.74E-40             | 29%             | 43%           | 525              | 59 4271              | 1                       | NP_048873 M.CviAII cytosin                                      | e DNA methyltransferase                                                         | 450.28                            | 4.09E-125                                                                        | 62%                                           | 74%                                           | 1-340                                                       | 1342                                                         |
|                                 |                                                                                                              |                                                             |                         |       |                   |                      | of methyl-binding proteins and their associated chromatin remodeling factors, X chromosome inactivation, imprinting and the suppression of parasitic DNA sequences. DNA methylation is also essential for proper embryonic development and is an important player in both DNA repair and nanome stability.                                                                                                                                                                                                                                                                                  |                |                      |                 |               |                  |                      |                         |                                                                 |                                                                                 |                                   |                                                                                  |                                               |                                               |                                                             |                                                              |
|                                 |                                                                                                              |                                                             |                         |       | 2                 | pfam00145<br>COG0270 | DNA methylase, C-5 cytosine-specific DNA methylase Dcm, Site-specific DNA methylase [DNA replication, recombination, and                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |                | 2.29E-39<br>4.34E-25 | 28%<br>27%      | 43%<br>44%    |                  | 36 4289<br>31 3197   |                         | NP_048886 M.CviAIV cytosi AAV84097 CviPII m5C DNA               |                                                                                 |                                   | 4.24E-122<br>2.51E-82                                                            | 61%<br>47%                                    | 76%<br>61%                                    | 1-328<br>2-340                                              | 2332<br>16357                                                |
|                                 |                                                                                                              |                                                             |                         |       |                   |                      | reoairl.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |                |                      |                 |               |                  |                      | 4<br>5<br>7<br>8<br>9   | AAC64006 cytosine methylt<br>AAC55063 cytosine methylt          | transferase<br>transferase<br>.CviAV cytosine DNA methyltransferase<br>ethylase | 266.16<br>265.00<br>100.14        | 2.69E-76<br>1.09E-69<br>2.44E-69<br>1.04E-19<br>9.73E-18<br>2.83E-17<br>6.30E-17 | 44%<br>40%<br>40%<br>32%<br>35%<br>33%<br>31% | 58%<br>56%<br>56%<br>48%<br>51%<br>44%<br>48% | 2-342<br>5-342<br>5-342<br>2-162<br>5-155<br>5-222<br>2-180 | 3-358<br>6-362<br>6-362<br>6-192<br>6-165<br>4-226<br>72-249 |
| M361R                           | 141459141671                                                                                                 | 71                                                          | 7,643                   | 11.29 |                   | No Hit Found         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |                |                      |                 |               |                  |                      |                         | NP 048874 a518R                                                 |                                                                                 | 112.46                            | 3.76E-24                                                                         | 76%                                           | 81%                                           | 171                                                         | 171                                                          |
| M362L                           | 142574141936                                                                                                 | 213                                                         | 23,843                  | 3.72  |                   | No Hit Found         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |                |                      |                 |               |                  |                      | 1 2                     | NP_048709 Asp/Glu rich; DA<br>YP 142843 unknown                 | AEDDDIYxxET (2X) negative charge cluster                                        |                                   | 1.29E-59<br>2.39E-05                                                             | 58%<br>32%                                    | 66%<br>58%                                    | 1213<br>134213                                              | 1207<br>140221                                               |
| M363L                           | 143695142673                                                                                                 | 341                                                         | 38,820                  | 3.99  | 1                 | COG5271              | MDN1, AAA ATPase containing von Willebrand factor type A (vWA) domain [General function prediction only].                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | 47.01          | 1.67E-06             | 24%             | 50%           | 171              | 52 3504085           | 1                       | NP_048714 A357L                                                 |                                                                                 | 161.77                            | 2.89E-38                                                                         | 51%                                           | 71%                                           | 160320                                                      | 86249                                                        |
| M367L                           | 144396143740                                                                                                 | 219                                                         | 25,122                  | 4.82  |                   | No Hit Found         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |                |                      |                 |               |                  |                      | 1                       | NP_048479 A131L                                                 |                                                                                 | 66.63                             | 5.87E-10                                                                         | 36%                                           | 50%                                           | 88218                                                       | 11–135                                                       |
|                                 |                                                                                                              |                                                             |                         |       |                   |                      | GIY-YIG_Clerm, GIYX[10-1]YIG stamily or class I homing endomucleases Cherminus (GIY-YIG_Clerm). Homing endomucleases Promote the mobility of introo or intellin by recognizing and cleaving a homologous allele that lacks the sequence. They catalyze a double-stand break in the DNA near the insertion size of that element to facilitate homing at that site. Class I homing endonucleases are sorted into four families based on the presence of these motifs in their respective N-                                                                                                   |                |                      |                 |               |                  |                      |                         |                                                                 |                                                                                 |                                   |                                                                                  |                                               |                                               |                                                             |                                                              |
| M370R                           | 144446145246                                                                                                 | 267                                                         | 30,020                  | 10.18 | 1                 | cd00283              | termin: LAGLIBADG, His-Cya box, HNY, and GNY-YIG. This CD contains several but not all members of the GIV-YIG family. The C-terminus of GIY-YIG is a DNA-binding domain which is separated from the N-terminus by a long, flexible linker. The DNA-binding domain consists of a minor-grove binding alpha-helix, and a helix-tum-helix. Some also contain a zin-flight (e. E-Tell y Minor is not required to DNA-binding or contain a zin-flight (e. E-Tell y Minor is not required to DNA-binding or domain to cleave the homing site at a fixed distance from the intron insertion, site. | 67.33          | 1.43E-12             | 48%             | 63%           | 13426            | 7 1113               |                         | NP_048671 A315L                                                 |                                                                                 | 244.97                            | 1.75E-63                                                                         | 47%                                           | 64%                                           | 1-267                                                       | 1-240                                                        |
|                                 |                                                                                                              |                                                             |                         |       | 2                 | smart00497           | IENR1, Intron encoded nuclease repeat motif; Repeat of unknown function, but possibly DNA-binding via helix-turn-helix motif (Ponting, unpublished).                                                                                                                                                                                                                                                                                                                                                                                                                                        | 45.51          | 5.54E-06             | 40%             | 60%           | 21626            | 67 1-51              | 2                       | NP_049007 similar to Chlore<br>Accession Numl                   | ella virus PBCV-1 ORF A315L, corresponds to GenBank<br>ber U42580               | 226.87                            | 4.93E-58                                                                         | 47%                                           | 62%                                           | 1-267                                                       | 1223                                                         |
|                                 |                                                                                                              |                                                             |                         |       | 3                 |                      | GIYc, GIY-YIG type nucleases (URI domain); .  NUMOD1, NUMOD1 domain                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                | 5.99E-05<br>0.003294 | 34%<br>41%      | 55%<br>69%    | 19<br>2162       |                      | 3                       | NP 048641 PBCV-1 33kd po<br>NP 048851 similar to PBC            | V-1 ORF A315L, corresponds to GenBank Accession                                 | 197.59<br>164.47                  | 3.20E-49<br>3.00E-39                                                             | 42%<br>38%                                    | 60%<br>54%                                    | 5267<br>1258                                                | 7247<br>1212                                                 |
|                                 |                                                                                                              |                                                             |                         |       | *                 | piaii107403          | Nones (, NONE) i dellant.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | JO. 17         | 0.003284             | ₹170            | 0970          | 2 1024           | ·· 1-33              | 5                       | NP_899393 SegD                                                  |                                                                                 | 80.49                             | 5.71E-14                                                                         | 31%                                           | 47%                                           | 1213                                                        | 1198                                                         |
|                                 |                                                                                                              |                                                             |                         |       |                   |                      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |                |                      |                 |               |                  |                      | 6<br>7<br>8             | AAK09365 intron encoded I<br>AAC49248 ORF211<br>AAC49244 ORF301 | BMOI                                                                            | 73.94<br>72.40<br>72.40           | 5.34E-12<br>1.55E-11<br>1.55E-11                                                 | 29%<br>33%<br>30%                             | 43%<br>46%<br>43%                             | 4-265<br>96-252<br>9-192                                    | 5266<br>77193<br>80240                                       |
|                                 |                                                                                                              |                                                             |                         |       |                   |                      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |                |                      |                 |               |                  |                      | 9<br>10                 | YP 293795 putative endonu<br>CAC51107 putative GIY-YIO          | 3 endonuclease                                                                  | 72.02                             | 2.03E-11<br>2.65E-11                                                             | 36%<br>35%                                    | 55%<br>54%                                    | 2109<br>148251                                              | 3111<br>105218                                               |
| M372R                           | 145289148681                                                                                                 | 1131                                                        | 124,388                 | 10.87 | 1                 |                      | DUF566, Family of unknown function (DUF566). Family of related proteins that is plant specific                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                | 1.30E-05             | 23%             |               |                  | 30 44169             | 1                       | NP_048720 similar to chicke<br>Number P02845                    | en vitellogenin II, corresponds to Swiss-Prot Accession                         |                                   | 0.00E+00                                                                         | 61%                                           |                                               | 3531029                                                     | 1651                                                         |
|                                 |                                                                                                              |                                                             |                         |       | 3                 |                      | MAP65 ASE1, Microtubule associated protein (MAP65/ASE1 family)<br>HELICc, helicase superfamily c-terminal domain; .                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | 38.86<br>37.52 | 0.000503<br>0.001249 | 24%<br>31%      | 41%<br>57%    |                  | 31 410559<br>91 1980 | 3                       | NP 048717 A360R<br>NP_048718 A361R                              |                                                                                 | 251.91<br>133.27                  | 9.52E-65<br>4.94E-29                                                             | 59%<br>72%                                    | 71%<br>84%                                    | 14226<br>249332                                             | 27–239<br>2–85                                               |

| Gene<br>Name | Genome<br>Position           | A.A.<br>length | Peptide<br>Mw | pl    | CDD Hit<br>Number | COGs         | COG Definition                                                                                                                                                                                                                                                                                                                                                          | Bit<br>Score | E-value  | %<br>Identity Po | %<br>sitive | Query<br>from-to | Hit<br>from-to | BLASTp<br>Hit | Hit<br>Accessio       | on BLASTp Definition                                                                                                                                                                      | Bit<br>Score            | E-value lo                       | %<br>dentity Po   |                   | Query H                    | lit from-<br>to            |
|--------------|------------------------------|----------------|---------------|-------|-------------------|--------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|----------|------------------|-------------|------------------|----------------|---------------|-----------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------|----------------------------------|-------------------|-------------------|----------------------------|----------------------------|
|              | 148717149253                 | 179            |               | 4.78  |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                         | 000.0        |          |                  | 00          |                  |                | 1             | A487                  | 79 translation elongation factor EF-3 homolog - Chlorella virus CVK2                                                                                                                      | 61.62                   | 1.17E-08                         | 39%               | 58%               | 75168                      | 84171                      |
|              |                              |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                         |              |          |                  |             |                  |                | 2             |                       |                                                                                                                                                                                           | 54.68                   | 1.43E-06                         | 41%               | 58%               | 85162                      | 54129                      |
|              | 149243149830                 | 196            |               | 10.92 |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                         |              |          |                  |             |                  |                |               |                       | nd No Hit Found                                                                                                                                                                           |                         |                                  |                   |                   |                            |                            |
|              | 150512149901<br>150719151156 | 146            | 23,518        | 10.92 |                   | No Hit Found | 5 GIYc, GIY-YIG type nucleases (URI domain); .                                                                                                                                                                                                                                                                                                                          | 35.44        | 0.005785 | 21%              | 38%         | 76172            | 183            |               | NP_0487               |                                                                                                                                                                                           |                         | 1.04E-39<br>8.11E-19             | 44%<br>37%        | 61%<br>51%        | 12203                      | 5203                       |
| M379R        | 150/19151156                 | 146            | 17,153        | 4.36  |                   | NO HIT FOUND |                                                                                                                                                                                                                                                                                                                                                                         |              |          |                  |             |                  |                | 1             | NP_04884              | 42 A486L                                                                                                                                                                                  | 94.74                   | 8.11E-19                         | 3/%               | 51%               | 1146                       | 1152                       |
| M381R        | 151222152637                 | 472            | 53,700        | 6.38  | 1                 | pfam0445     | Capsid_Indowir, Indowirus major capsid protein. This family includes the<br>major capsid protein of indowiruses, othorella visus and Spodopiera<br>accovirus, which are all dsDNA viruses with no RNA stage. This is the<br>most abundant structural protein and can account for up to 45% of virion<br>protein. In Chiorella virus MT325 the major capsid protein is a | 178.20       | 6.16E-46 | 29%              | 45%         | 81-470           | 3422           | 1             | NP_04874              | 40 similar to PBCV-1 major capsid protein, corresponds to Swiss-Prot<br>Accession Number P30328                                                                                           | 291.97                  | 2.94E-77                         | 53%               | 72%               | 148405                     | 1257                       |
|              |                              |                |               |       |                   |              | alvoorotein                                                                                                                                                                                                                                                                                                                                                             |              |          |                  |             |                  |                | 2             |                       | 35 hypothetical major capsid protein<br>01 major capsid protein MCP1                                                                                                                      | 163.31<br>139.81        | 1.58E-38<br>1.87E-31             | 30%<br>30%        | 48%<br>45%        | 81446<br>81438             | 3401<br>3380               |
|              |                              |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                         |              |          |                  |             |                  |                | 4<br>5        | BAA2219               | 98 major capsid protein Vp54  PBCV-1 major capsid protein Vp54, corresponds to GenBank Accession                                                                                          | 137.89<br>137.89        | 7.09E-31<br>7.09E-31             | 29%<br>29%        | 42%<br>42%        | 81469<br>81469             | 3410<br>3410               |
|              |                              |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                         |              |          |                  |             |                  |                | 6             | BAA766                | 00 major capsid protein                                                                                                                                                                   | 137.50                  | 9.26E-31                         | 29%               | 43%               | 81469                      | 3409                       |
|              |                              |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                         |              |          |                  |             |                  |                | 8             | 1M3Y_                 | 92 major capsid protein Vp49  D Chain D, The Structure Of Major Capsid Protein Of A Large, Lipid                                                                                          | 135.96<br>115.93        | 2.69E-30<br>2.89E-24             | 30%<br>28%        | 44%<br>40%        | 81438<br>103469            | 3371<br>1386               |
|              |                              |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                         |              |          |                  |             |                  |                | 9             | 1M4X                  | Containino. Dna Virus<br>C Chain C, Pbov-1 Virus Capsid, Quasi-Atomic Model<br>Lys., Pro-rich, PAPK (10x); similar to wheat Pro-, Lys-rich protein,                                       |                         | 3.77E-24                         | 28%               | 40%               | 103469                     | 1386                       |
|              |                              |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                         |              |          |                  |             |                  |                | 10            | NP_04874              | corresponds to GenBank Accession Number X52472                                                                                                                                            | 95.52                   | 4.04E-18                         | 62%               | 75%               | 401472                     | 385456                     |
| M385L        | 154232152652                 | 527            | 59,190        | 11.06 |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                         |              |          |                  |             |                  |                | 1             |                       | 41 similar to Chlamydia histone-like protein, corresponds to GenBank<br>Accession Number D71563                                                                                           | 150.21                  | 1.59E-34                         | 30%               | 47%               | 61391                      | 3296                       |
|              |                              |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                         |              |          |                  |             |                  |                | 2             | _                     | 36 similar to PBCV-1 ORF A34R, corresponds to GenBank Accession<br>Number U17055                                                                                                          | 117.09                  | 1.49E-24                         | 46%               | 55%               |                            | 414568                     |
|              |                              |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                         |              |          |                  |             |                  |                | 3             |                       | on similar to hovine cylicin L corresponds to Swiss-Prot Accession Number                                                                                                                 | 114.39<br>95.52         | 9.69E-24<br>4.66E-18             | 48%<br>41%        | 65%<br>52%        | 403527<br>7161             | 1126<br>465609             |
|              |                              |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                         |              |          |                  |             |                  |                | 5             | NP_04903              | P35662 similar to Chlorella virus PBCV-1 ORF A282L, corresponds to GenBank                                                                                                                |                         | 3.02E-09                         | 72%               | 84%               | 58101                      | 244                        |
| M388R        | 154334155038                 | 235            | 27.144        | 4.58  |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                         |              |          |                  |             |                  |                | 1             |                       | Accession Number U42580 36 hypothetical protein PSSM4 035                                                                                                                                 | 55 45                   | 1.56E-06                         | 26%               | 42%               | 6205                       | 306511                     |
|              |                              |                |               |       |                   |              | ATP_bind_3, PP-loop family. This family of proteins belongs to the PP-                                                                                                                                                                                                                                                                                                  |              |          |                  |             |                  |                |               |                       |                                                                                                                                                                                           |                         |                                  |                   |                   |                            |                            |
| M389R        | 155125156621                 | 499            | 57,656        | 8.49  | 1                 | pfam0117     | Ionn sunerfamily PP-ATPase, N-terminal domain of predicted ATPase of the PP-loop faimly implicated in cell cycle control [Cell division and chromosome partitioning]. This is a sunfamily of Adenine nucleotide alpha bydrolases.                                                                                                                                       | 124.59       | 7.26E-30 | 27%              | 47%         | 196399           | 1204           | 1             | NP_0489               | 10 similar to MesJ cell cycle protein                                                                                                                                                     | 525.78                  | 1.30E-147                        | 52%               | 67%               | 1-494                      | 1497                       |
|              |                              |                |               |       | 2                 |              | 2 superfamily Adeninosine nucleotide alpha hydrolases superfamily<br>includes N type ATP PPases and ATP sulphurylases. It forms a<br>apha/beta/apha fold which binds to Adenosine group. This domain has a<br>stronoly conserved motif SGGXD at the N terminus.                                                                                                         | 123.03       | 2.62E-29 | 29%              | 48%         | 196381           | 1185           | 2             | T180                  | 59 hypothetical protein A557L - Chlorella virus PBCV-1                                                                                                                                    | 104.38                  | 9.38E-21                         | 51%               | 65%               | 26122                      | 14–112                     |
|              |                              |                |               |       | 3                 |              | <ul> <li>MesJ, Predicted ATPase of the PP-loop superfamily implicated in cell<br/>cycle control (Cell division and chromosome partitioning).</li> <li>Alpha_ANH_like_II, This is a subfamily of Adenine nucleotide alpha<br/>bytrolaces superfamily. Adeninosing purjectible alpha bytrolaces</li> </ul>                                                                | 104.54       | 8.67E-24 | 25%              | 44%         | 181389           | 6218           | 3             | CAG395                | 32 conserved hypothetical protein                                                                                                                                                         | 87.04                   | 1.55E-15                         | 23%               | 47%               | 197495                     | 14-320                     |
|              |                              |                |               |       | 4                 |              | superfamily includes N type ATP PPases and ATP sulphurylases. It forms a apha/beta/apha fold which binds to Adenosine group. This subfamily of proteins is predicted to bind ATP. This domainhas a strongly conserved motif SGGKD at the N terminus.                                                                                                                    | 66.04        | 3.30E-12 | 23%              | 41%         | 196367           | 1180           | 4             | CAI801                | 46 conserved hypothetical protein                                                                                                                                                         | 86.27                   | 2.64E-15                         | 23%               | 46%               | 197495                     | 14-320                     |
|              |                              |                |               |       | 5                 | pfam0604     | DUF924, Bacterial protein of unknown function (DUF924). This family consists of several invoothetical bacterial proteins of unknown function.                                                                                                                                                                                                                           | 46.44        | 2.53E-06 | 23%              | 40%         | 2105             | 4117           | 5             | BAB566                | 71 conserved hypothetical protein                                                                                                                                                         | 85.89                   | 3.45E-15                         | 23%               | 46%               | 197495                     | 14320                      |
|              |                              |                |               |       | 6                 | cd01712      | Thil, Thil is required for thiazole synthesis in the thiamine biosynthesis<br>2 pathway. It belongs to the Adenosine Nucleotide Hydrolysis sucerfamily<br>and predicted to hind to Adenosine pucleotide                                                                                                                                                                 | 36.73        | 0.002192 | 23%              | 41%         | 201356           | 6153           | 6             | AAW376                | 65 tRNA(IIe)-lysidine synthetase                                                                                                                                                          | 85.89                   | 3.45E-15                         | 23%               | 46%               | 197495                     | 14-320                     |
|              |                              |                |               |       | 7                 |              | CysH, 3'-phosphoadenosine 5'-phosphosulfate<br>sulfotransferase (PAPS reductase)/FAD synthetase and related enzymes<br>[Amino acid transport and metabolism / Coenzyme metabolism].                                                                                                                                                                                     | 35.02        | 0.008227 | 20%              | 38%         | 178364           | 23199          | 7             | YP_19320              | 01 putative cell cycle                                                                                                                                                                    | 84.73                   | 7.69E-15                         | 25%               | 45%               | 196440                     | 23272                      |
|              |                              |                |               |       | 8                 | COG030       | Thil, Thiamine biosynthesis ATP pyrophosphatase [Coenzyme metabolism].                                                                                                                                                                                                                                                                                                  | 34.87        | 0.008691 | 24%              | 46%         | 196356           | 177-329        | 9             | CAG4224<br>ZP_0057629 | 41 conserved hypothetical protein                                                                                                                                                         |                         | 1.00E-14<br>1.31E-14<br>1.11E-13 | 26%<br>22%<br>27% | 47%<br>46%<br>51% | 192419<br>197495<br>196394 | 28260<br>14320<br>20217    |
| M395L        | 156892156632                 | 87             | 9,642         | 10.44 |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                         |              |          |                  |             |                  |                |               | No Hit Four           | nd No Hit Found                                                                                                                                                                           |                         |                                  |                   |                   |                            |                            |
| M397R        | 157055157816                 | 254            | 28,999        | 8.32  |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                         |              |          |                  |             |                  |                | 1             | NP_0483               | 57 A9R<br>07 similar to PBCV-1 ORF A275R, encoded by GenBank Accession Number                                                                                                             | 215.31                  | 1.36E-54                         | 57%               | 78%               | 89254                      | 8173                       |
|              |                              |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                         |              |          |                  |             |                  |                | 2             | NP_04880              | of similar to Pack 1 OKF A278, encoded by GenBank Accession Number 143590  similar to Chlorella virus PBCV-1 ORF A450R, corresponds to GenBank                                            | 98.98                   | 1.42E-19                         | 25%               | 45%               | 2-253                      | 1246                       |
|              |                              |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                         |              |          |                  |             |                  |                | 3             |                       | Accession Number LI42580  Similar to PBCV-1 ORF A79R, corresponds to GenBank Accession Number U17055                                                                                      | 95.13                   | 2.06E-18<br>5.98E-18             | 27%<br>25%        | 44%<br>44%        | 5253<br>2252               | 10-253<br>4-248            |
|              |                              |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                         |              |          |                  |             |                  |                | 5             | NP_04852              | 25 A177R                                                                                                                                                                                  | 92.82                   | 1.02E-17                         | 26%               | 47%               | 2-241                      | 4230                       |
|              |                              |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                         |              |          |                  |             |                  |                | 6<br>7        | AAU063                | 04 hypothetical protein A275R                                                                                                                                                             | 83.19<br>79.72          | 8.09E-15<br>8.94E-14             | 25%<br>30%        | 46%<br>50%        | 2232<br>89252              | 1218<br>4167               |
|              |                              |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                         |              |          |                  |             |                  |                | 8<br>9        | AAU063                | 01 hypothetical protein A275R<br>02 hypothetical protein A275R                                                                                                                            | 78.57<br>66.63          | 1.99E-13<br>7.83E-10             | 30%<br>32%        | 50%<br>52%        | 89252<br>121253            | 4167<br>1132               |
| M399L        | 158394157846                 | 183            | 20,781        | 7.72  | . 1               | pfam02940    | mRNA_triPase, mRNA capping enzyme, beta chain. The beta chain of mRNA capping enzyme has triphosphatase activity. The function of the capping enzyme also depends on the guanylytimansferase activity conferred by the alcha chain (see ofam01331).                                                                                                                     | 41.15        | 9.92E-05 | 34%              | 55%         | 94156            | 191256         | 1             | NP_04880              | 06 PBCV-1 RNA triphosphatase                                                                                                                                                              | 184.11                  | 1.67E-45                         | 54%               | 69%               | 5183                       | 11–193                     |
| M401R        | 158497159168                 | 224            | 25,101        | 7.96  | . 1               | smart00702   | P4Hc, Prolyl 4-hydroxylase alpha subunit homologues. Mammalian enzymes catalyse hydroxylation of collagen, for example. Prokaryotic enzymes might catalyse hydroxylation of antibiotic peptides. These are 2-                                                                                                                                                           | 116.71       | 2.03E-27 | 29%              | 44%         | 34222            | 1178           | 1             | NP_04843              | 33 PBCV-1 prolyl 4-hydroxylase                                                                                                                                                            | 227.64                  | 2.09E-58                         | 55%               | 71%               | 32224                      | 48-241                     |
|              |                              |                |               |       |                   |              | oxoglutarate-dependent dioxygenases, requiring 2-oxoglutarate and<br>dioxygena so consibstates and ferrous iron as a cnfactor<br>2OG-Fell_Oxy, 2OG-Fe(II) oxygenase superfamily. This family contains<br>members of the 2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase<br>superfamily. This family includes the C-terminal of prolyl 4-hydroxylase                 |              |          |                  |             |                  |                |               |                       |                                                                                                                                                                                           |                         |                                  |                   |                   |                            |                            |
|              |                              |                |               |       | 2                 | pfam0317     | alpha subunit. The holoenzyme has the activity EC:1.14.1.2 catalysing the reaction: Procollagen L-proline + 2-oxoglutarate + O2 <=> procollagen trans -4-hydroxy-L-proline + succinate + CO2. The full enzyme consists of a alpha2 beta2 complex with the alpha subunit contributing most of the parts of the active site. The family also includes                     | 42.02        | 5.20E-05 | 27%              | 39%         | 119223           | 2–96           | 2             | AAZ623                | 10 Procollagen-proline,2-oxoglutarate-4-dioxygenase                                                                                                                                       | 93.20                   | 6.16E-18                         | 34%               | 49%               | 34222                      | 90-274                     |
|              |                              |                |               |       |                   |              | lvsvl hvdrolases. isopenicillin svnthases and AlkB                                                                                                                                                                                                                                                                                                                      |              |          |                  |             |                  |                | 3             | CAD1652               | 21 HYPOTHETICAL PROTEIN                                                                                                                                                                   |                         | 3.99E-17<br>6.81E-17             | 33%<br>33%        | 47%<br>47%        | 34222<br>34222             | 97–281<br>97–281           |
|              |                              |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                         |              |          |                  |             |                  |                | 5             | ZP 0059909            | 67 Prolyl 4-hydroxylase alpha subunit 97 Procollagen-proline,2-oxoglutarate-4-dioxygenase                                                                                                 | 87.43                   | 6.81E-17<br>3.38E-16<br>9.83E-16 | 33%<br>32%<br>30% | 47%<br>46%<br>46% | 34222<br>32222<br>34222    | 97–281<br>99–285<br>73–257 |
|              |                              |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                         |              |          |                  |             |                  |                | 7             | ZP 0050829            | 85 hypothetical protein BdolA 01003928<br>97 Procollagen-proline,2-oxoglutarate-4-dioxygenase                                                                                             | 85.89<br>83.57<br>83.19 | 4.88E-15                         | 31%               | 46%<br>46%<br>44% | 34222                      | 55237                      |
|              |                              |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                         |              |          |                  |             |                  |                | 9             | AAF085                | 47 prolyl 4-hydroxylase 83 unknown protein 40 bysythetical protein Real/03000555                                                                                                          | 83.19<br>78.57<br>78.57 | 6.37E-15<br>1.57E-13<br>1.57E-13 | 29%<br>32%<br>29% | 44%<br>43%<br>40% | 30-222<br>34-221<br>34-222 | 56-237<br>45-223<br>87-269 |
| M403R        | 159213159530                 | 106            | 11,844        | 10.09 | 1                 | pfam0008     | Thioredoxin, Thioredoxin. Thioredoxins are small enzymes that participate in redox reactions, via the reversible oxidation of an active                                                                                                                                                                                                                                 | 61.78        | 6.44E-11 | 29%              | 48%         | 29105            | 33108          |               |                       | 40 hypothetical protein Rqel02000555<br>contains cytochrome C family heme-binding site signature; similar to<br>05 maize protein disulphide isomerase, correspond to Swiss-Prot Accession |                         | 1.57E-13<br>4.40E-28             | 29%<br>56%        | 72%               | 499                        | 6101                       |
|              |                              |                |               |       | 2                 |              | centre distrince bond. Some members with only the active site are not separated from the noise  COG3118, Thioredoxin domain-containing protein [Posttranslational                                                                                                                                                                                                       |              | 0.000665 | 20%              | 40%         |                  | 39125          |               |                       | Number P52588  42 transglutaminase                                                                                                                                                        |                         | 8.34E-11                         | 35%               | 58%               | 12105                      | 8106                       |
|              |                              |                |               |       | 2                 |              | modification. protein turnover. chaperonesl.                                                                                                                                                                                                                                                                                                                            | 30.35        | 0.00000  | 2070             | 40,0        | 1499             | 30123          | 3             |                       | 87 protein disulfide isomerase ER-60                                                                                                                                                      |                         | 1.42E-10                         | 36%               | 57%               |                            | 43–123                     |

| Gene<br>Name | Genome<br>Position | A.A.<br>length | Peptide<br>Mw | pl    | CDD Hit<br>Number | COGs         | COG Definition                                                                                                                                                                                                                                                                                                                                                                                                                        | Bit<br>Score | E-value  | %<br>Identity F |     | Query<br>from-to | riit  | BLASTp<br>Hit<br>Number | Hit<br>Accession        | n BLASTp Definition                                                                                                                         | Bit<br>Score                     | E-value lo                                   | %<br>dentity Po          | %<br>ositive             | Query<br>from-to         | Hit from-<br>to                 |
|--------------|--------------------|----------------|---------------|-------|-------------------|--------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|----------|-----------------|-----|------------------|-------|-------------------------|-------------------------|---------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------|----------------------------------------------|--------------------------|--------------------------|--------------------------|---------------------------------|
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                       |              |          |                 |     |                  |       | 4                       |                         | 5 Hypothetical protein zgc:77086                                                                                                            |                                  | 3.17E-10                                     | 36%                      | 56%                      | 23105                    | 45125                           |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                       |              |          |                 |     |                  |       | 5                       |                         | 2 PREDICTED: similar to Protein disulfide isomerase A6 precursor (Protein disulfide isomerase P5) (Thioredoxin domain containing protein 7) |                                  | 3.17E-10                                     | 45%<br>45%               | 62%                      | 29102                    | 60-134                          |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                       |              |          |                 |     |                  |       | 7                       | AAC24752                | 2 hypothetical protein<br>2 transqlutaminase precursor                                                                                      | 64.70                            | 3.17E-10<br>9.22E-10                         | 35%                      | 56%                      | 29102<br>12105           | 60-134<br>37-135                |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                       |              |          |                 |     |                  |       | 9<br>10                 | AAH44524                | 9 unnamed protein product<br>4 Sb:cb825 protein<br>3 protein disulfide isomerase 1                                                          | 63.54                            | 9.22E-10<br>2.05E-09<br>3.50E-09             | 45%<br>37%<br>36%        | 62%<br>56%<br>58%        | 2997<br>23105<br>2399    | 55-124<br>43-123<br>46-120      |
|              |                    |                |               |       |                   |              | AarF, Predicted unusual protein kinase [General function prediction                                                                                                                                                                                                                                                                                                                                                                   |              |          |                 |     |                  |       |                         |                         | 2 similar to Clostridium pasteurianum ORF, corresponds to GenBank                                                                           |                                  |                                              |                          |                          |                          |                                 |
| M404R        | 159584160972       | 463            | 52,177        | 6.01  | 1                 | COG0661      | ABC1, ABC1 family. This family includes ABC1 from yeast and AarF                                                                                                                                                                                                                                                                                                                                                                      | 175.14       | 5.17E-45 | 27%             | 43% | 24385            | 37423 | 1                       | NP_048802               | Accession Number Z28353                                                                                                                     | 620.16                           | 4.58E-176                                    | 64%                      | 84%                      | 13463                    | 11-462                          |
|              |                    |                |               |       |                   |              | from E. coli. These proteins have a nuclear or mitochondrial subcellular                                                                                                                                                                                                                                                                                                                                                              |              |          |                 |     |                  |       |                         |                         |                                                                                                                                             |                                  |                                              |                          |                          |                          |                                 |
|              |                    |                |               |       | 2                 | pfam03109    | not clear, however yeast ABC1 suppresses a cytochrome b mRNA translation defect and is essential for the electron transfer in the bc1 complex and E. coil Aarf is required for ubiquinone production. It has been suggested that members of the ABC1 family are novel chaperonins. These proteins are unrelated to the ABC transporter proteins:                                                                                      | 114.54       | 9.83E-27 | 41%             | 58% | 96-209           | 7119  | 2                       | ZP_00673531             | 1 ABC-1                                                                                                                                     | 138.27                           | 5.28E-31                                     | 24%                      | 46%                      | 49448                    | 70–501                          |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                       |              |          |                 |     |                  |       | 3                       | BAD78862<br>ABA22988    | 2 hypothetical protein<br>8 ABC-1                                                                                                           |                                  | 7.63E-30<br>9.97E-30                         | 24%<br>24%               | 46%<br>46%               | 49448<br>37448           | 83515<br>71516                  |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                       |              |          |                 |     |                  |       | 5<br>6                  | BAB76214                | 3 488aa long conserved hypothetical protein<br>4 alr4515                                                                                    | 134.04<br>133.27                 | 9.97E-30<br>1.70E-29                         | 24%<br>23%               | 46%<br>46%               | 12457<br>37448           | 9459<br>82527                   |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                       |              |          |                 |     |                  |       | 7<br>8                  | BAC92119<br>ZP 01006074 | 4 Predicted protein kinase                                                                                                                  | 130.18<br>128.64                 | 1.44E-28<br>4.19E-28                         | 25%<br>26%               | 46%<br>48%               | 54394<br>36392           | 102470<br>48428                 |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                       |              |          |                 |     |                  |       | 9<br>10                 | AAY80974<br>ZP_00112431 | universally conserved protein     COG0661: Predicted unusual protein kinase                                                                 | 127.87<br>127.49                 | 7.14E-28<br>9.33E-28                         | 26%<br>23%               | 45%<br>45%               | 12384<br>49448           | 9399<br>78511                   |
| M407R        | 161042161380       | 113            | 13,229        | 5.00  |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                       |              |          |                 |     |                  |       | 1                       | NP_048801               | 1 A444L                                                                                                                                     | 71.63                            | 7.35E-12                                     | 35%                      | 56%                      | 4112                     | 299                             |
| M409R        | 161649163055       | 469            | 50,358        | 10.79 | 1                 | COG0810      | TonB, Periplasmic protein TonB, links inner and outer membranes [Cell envelope biogenesis, outer membranel.                                                                                                                                                                                                                                                                                                                           | 60.16        | 2.18E-10 | 32%             | 36% | 66194            | 36161 |                         | No Hit Found            | d No Hit Found                                                                                                                              |                                  |                                              |                          |                          |                          |                                 |
| M411L        | 163494163072       | 141            | 15,822        | 4.42  |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                       |              |          |                 |     |                  |       | 1                       | NP_048798               | 8 A441L                                                                                                                                     | 125.95                           | 3.35E-28                                     | 46%                      | 61%                      | 6141                     | 2137                            |
|              |                    |                |               |       |                   |              | ANK, ankyrin repeats; ankyrin repeats mediate protein-protein interactions in very diverse families of proteins. The number of ANK                                                                                                                                                                                                                                                                                                    |              |          |                 |     |                  |       |                         |                         |                                                                                                                                             |                                  |                                              |                          |                          |                          |                                 |
| M413L        | 164611163613       | 333            | 36,902        | 6.59  | 1                 | cd00204      | repeats in a protein can range from 2 to over 20 (ankyrins, for example).  ANK repeats may occur in combinations with other types of domains. The                                                                                                                                                                                                                                                                                     | 135.60       | 4.30E-33 | 59%             | 79% | 62188            | 1126  | 1                       | EAL29245                | 5 GA14074-PA                                                                                                                                | 197.59                           | 4.62E-49                                     | 41%                      | 59%                      | 9270                     | 170434                          |
|              |                    |                |               |       |                   |              | structural repeat unit contains two antiparallel helices and a beta-hairpin,<br>repeats are stacked in a superhelical arrangement; this alignment<br>contains 4 consecutive repeats                                                                                                                                                                                                                                                   |              |          |                 |     |                  |       |                         |                         |                                                                                                                                             |                                  |                                              |                          |                          |                          |                                 |
|              |                    |                |               |       | 2                 |              | Arp, FOG: Ankyrin repeat [General function prediction only].  Ank, Ankyrin repeat. There's no clear separation between noise                                                                                                                                                                                                                                                                                                          | 91.49        | 7.70E-20 | 34%             | 50% | 46233            | 17212 | 2                       | AAM11327                | 7 GH01626p                                                                                                                                  | 196.05                           | 1.35E-48                                     | 42%                      | 59%                      | 9270                     | 25289                           |
|              |                    |                |               |       | 3                 | pfam00023    | and signal on the HMM search Ankyrin repeats generally consist of a<br>beta, alpha, alpha, beta order of secondary structures. The repeats<br>associate to form a higher order structure.                                                                                                                                                                                                                                             | 51.22        | 1.13E-07 | 56%             | 81% | 6799             | 1-33  | 3                       | XP_681288               | 8 hypothetical protein AN8019.2                                                                                                             | 196.05                           | 1.35E-48                                     | 43%                      | 60%                      | 10270                    | 8011064                         |
|              |                    |                |               |       | 4                 | smart00248   | ANK, ankyrin repeats; Ankyrin repeats are about 33 amino acids long<br>and occur in at least four consecutive copies. They are involved in protein<br>protein interactions. The core of the repeat seems to be an helix-loop-                                                                                                                                                                                                         | 42.64        | 4.08E-05 | 55%             | 76% | 6796             | 1-30  | 4                       | AAN12046                | 6 CG7462-PC, isoform C                                                                                                                      | 196.05                           | 1.35E-48                                     | 42%                      | 59%                      | 9270                     | 175439                          |
|              |                    |                |               |       |                   |              | helix structure                                                                                                                                                                                                                                                                                                                                                                                                                       |              |          |                 |     |                  |       | 5                       | AAF50525                | 5 CG7462-PB, isoform B                                                                                                                      | 196.05                           | 1.35E-48                                     | 42%                      | 59%                      | 9270                     | 175439                          |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                       |              |          |                 |     |                  |       | 6<br>7                  | EAA03765                | 8 RE55168p<br>5 ENSANGP00000006233                                                                                                          |                                  | 1.35E-48<br>1.14E-47                         | 42%<br>42%               | 59%<br>59%               | 9270<br>9270             | 325589<br>175439                |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                       |              |          |                 |     |                  |       | 9                       | ZP 00374082             | 4 hypothetical protein AN1130.2 2 ankyrin 1, erythrocyte splice form 1 8 PREDICTED: similar to CG7462-PB, isoform B                         | 190.66<br>189.89<br>187.96       | 5.65E-47<br>9.64E-47<br>3.66E-46             | 43%<br>42%<br>40%        | 56%<br>59%<br>59%        | 13273<br>25274<br>9270   | 577-840<br>81-333<br>180-444    |
| M416R        | 164684165577       | 298            | 33,175        | 5.33  |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                       |              |          |                 |     |                  |       |                         | NP_048800               |                                                                                                                                             |                                  | 1.68E-44                                     | 33%                      | 54%                      | 6-297                    | 8-306                           |
|              |                    |                |               |       |                   |              | Trypan_PARP, Procyclic acidic repetitive protein (PARP). This family                                                                                                                                                                                                                                                                                                                                                                  |              |          |                 |     |                  |       |                         |                         |                                                                                                                                             |                                  |                                              |                          |                          |                          |                                 |
| M417L        | 167392165593       | 600            | 64,526        | 8.50  | 1                 | pfam05887    | consists of several Trypanosoma brucei procyclic acidic repetitive protein<br>(PARP) like sequences. The procyclic acidic repetitive protein (parp)<br>genes of Trypanosoma brucei encode a small family of abundant surface<br>proteins whose expression is restricted to the procyclic form of the<br>parasite. They are found at two unlinked loci, parpA and parpB;<br>transcription of thoth loc is is developmentably remulated | 59.22        | 4.27E-10 | 50%             | 52% | 53115            | 60122 | 1                       | BAE02830                | 0 surface protein                                                                                                                           | 65.08                            | 7.94E-09                                     | 45%                      | 56%                      | 181255                   | 9971071                         |
| M421L        | 167746167417       | 110            | 12,636        | 6.50  |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                       |              |          |                 |     |                  |       | 1                       | NP_048796               | 6 A439R                                                                                                                                     | 124.79                           | 7.39E-28                                     | 61%                      | 82%                      | 19103                    | 21-105                          |
| M423R        | 167767168024       | 86             | 9,971         | 10.79 | 1                 |              | GrxC, Glutaredoxin and related proteins [Posttranslational modification, protein turnover, chaperones].                                                                                                                                                                                                                                                                                                                               | 51.89        | 5.60E-08 | 34%             | 55% | 372              | 5-72  | 1                       |                         | 5 similar to E. coli glutaredoxin, corresponds to Swiss-Prot Accession Number P37687                                                        | 91.28                            | 9.20E-18                                     | 55%                      | 81%                      | 372                      | 574                             |
|              |                    |                |               |       | 2                 | pfam00462    | Glutaredoxin, Glutaredoxin                                                                                                                                                                                                                                                                                                                                                                                                            | 38.70        | 0.000524 | 25%             | 49% | 172              | 2-73  | 2<br>3<br>4<br>5        | AAL51366<br>AAN30771    | 8 probable peroxiredoxin/qlutaredoxin family protein 6 GLUTAREDOXIN 1 qlutaredoxin 3 2 Glutaredoxin                                         | 50.83<br>50.45<br>50.45<br>49.29 | 1.38E-05<br>1.80E-05<br>1.80E-05<br>4.01E-05 | 38%<br>32%<br>32%<br>31% | 57%<br>57%<br>57%<br>59% | 174<br>170<br>170<br>371 | 173–241<br>8–73<br>4–69<br>8–72 |
|              |                    |                |               |       |                   |              | dNK, Deoxyribonucleoside kinase (dNK) catalyzes the phosphorylation of deoxyribonucleosides to yield corresponding monophosphates                                                                                                                                                                                                                                                                                                     |              |          |                 |     |                  |       |                         |                         |                                                                                                                                             |                                  |                                              |                          |                          |                          |                                 |
| M425R        | 168038168595       | 186            | 21,497        | 4.90  | 1                 | cd01673      | (dNMPs). This family consists of various deoxynucleoside kinases including deoxyribo-cyldine (EC 2.7.1.74), guanosine (EC 2.7.1.13), adenosine (EC 2.7.1.16), and thymidine (EC 2.7.1.21) kinases. They are key enzymes in the salvage of deoxyribonucleosides originating from extra- or intra-glular breakdown of DNA.                                                                                                              | 101.08       | 8.98E-23 | 31%             | 52% | 3165             | 1178  | 1                       | NP_048773               | 3 contains ATP/GTP-binding site motif A; similar to Bacillus subtilus 24.1 kDa protein, corresponds to Swiss-Prot Accession Number P37530   | 190.27                           | 2.44E-47                                     | 50%                      | 72%                      | 3178                     | 2180                            |
|              |                    |                |               |       | 2                 | COG1428      | COG1428, Deoxynucleoside kinases [Nucleotide transport and metabolism].                                                                                                                                                                                                                                                                                                                                                               | 91.09        | 9.10E-20 | 31%             | 51% | 1181             | 4210  | 2                       | NP_149606               | 6 143R                                                                                                                                      | 88.20                            | 1.31E-16                                     | 31%                      | 51%                      | 1152                     | 1157                            |
|              |                    |                |               |       | 3                 |              | dNK, Deoxynucleoside kinase. This family consists of various deoxynucleoside kinases cytidine EC:2.7.1.74, guanosine EC:2.7.1.113, adenosine EC:2.7.1.76 and thymidine kinase EC:2.7.1.21 (which also                                                                                                                                                                                                                                 | 77.00        | 1.29E-15 | 32%             | 49% | FO 400           | 4 400 |                         | VD 070500               |                                                                                                                                             | 78.57                            | 1 005 10                                     | 31%                      | 56%                      | 3152                     | 4161                            |
|              |                    |                |               |       | 3                 |              | phosphorylates deoxyundine and deoxycytosine.) I hese enzymes<br>cataliyse the production of deoxynucleotide 5':-monophosphate<br>from a deoxynucleoside. Using ATP and yielding ADP in the process                                                                                                                                                                                                                                   |              |          |                 |     |                  | 1138  |                         |                         | 6 deoxynucleoside kinases                                                                                                                   |                                  | 1.03E-13                                     |                          |                          |                          |                                 |
|              |                    |                |               |       | 4                 | COG0125      | Tmk, Thymidylate kinase [Nucleotide transport and metabolism].  TMPK, Thymidine monophosphate kinase (TMPK), also known as                                                                                                                                                                                                                                                                                                            | 54.12        | 1.48E-08 | 22%             | 46% | 4184             | 6204  | 4                       | CAG42289                | 9 putative deoxyadenosine kinase protein                                                                                                    | 72.79                            | 5.68E-12                                     | 29%                      | 47%                      | 3181                     | 11-211                          |
|              |                    |                |               |       | 5                 | cd01672      | thymidylate kinase, catalyzes the phosphorylation of thymidine<br>monophosphate (TMP) to thymidine diphosphate (TDP) utilizing ATP as<br>its preferred phophoryl donor. TMPK represents the rate-limiting step in<br>either de novo or salvage biosynthesis of thymidine triphosphate (TTP).                                                                                                                                          | 49.50        | 3.19E-07 | 19%             | 43% | 3181             | 2198  | 5                       | CAI80195                | 5 deoxypurine kinase subunit                                                                                                                | 72.79                            | 5.68E-12                                     | 29%                      | 48%                      | 3181                     | 11–211                          |
|              |                    |                |               |       |                   |              | NK, Nucleoside/nucleotide kinase (NK) is a protein superfamily consisting of multiple families of enzymes that share structural similarity                                                                                                                                                                                                                                                                                            |              |          |                 |     |                  |       |                         |                         |                                                                                                                                             |                                  |                                              |                          |                          |                          |                                 |
|              |                    |                |               |       | 6                 | cd02019      | and are functionally related to the catalysis of the reversible phosphate<br>group transfer from nucleoside triphosphates to nucleosides/nucleotides,                                                                                                                                                                                                                                                                                 | 41.73        | 7.78E-05 | 21%             | 39% | 3159             | 1165  | 6                       | AAW37713                | 3 deoxynucleoside kinase family protein                                                                                                     | 72.02                            | 9.68E-12                                     | 32%                      | 50%                      | 3153                     | 11-172                          |
|              |                    |                |               |       |                   |              | nucleoside monophosphates, or sugars. Members of this family play a<br>wide variety of essential roles in nucleotide metabolism, the biosynthesis<br>of coenzymes and aromatic compounds, as well as the metabolism of                                                                                                                                                                                                                |              |          |                 |     |                  |       |                         |                         |                                                                                                                                             |                                  |                                              |                          |                          |                          |                                 |
|              |                    |                |               |       | 7                 | pfam02223    | sugar and sulfate  Thymidylate kin, Thymidylate kinase  NDUO42, NADH:Ubiquinone oxioreductase, 42 kDa (NDUO42) is a                                                                                                                                                                                                                                                                                                                   | 41.42        | 9.97E-05 | 20%             | 38% | 6177             | 1186  | 7                       | XP_770504               | 4 deoxypurine kinase subunit                                                                                                                | 71.63                            | 1.26E-11                                     | 33%                      | 51%                      | 4155                     | 33192                           |
|              |                    |                |               |       |                   |              | family of proteins that are highly similar to deoxyribonucleoside kinases                                                                                                                                                                                                                                                                                                                                                             |              |          |                 |     |                  |       |                         |                         |                                                                                                                                             |                                  |                                              |                          |                          |                          |                                 |
|              |                    |                |               |       | 8                 | cd02030      | of NADH: Ubiquinone oxioreductase (complex I), a multi-protein complex located in the simulti-protein complex located in the inner mitochondrial membrane. The main function of the complex is to transport electrons from NADH to ubiquinone, which is                                                                                                                                                                               | 39.85        | 0.000245 | 25%             | 45% | 3153             | 1184  | 8                       | CAC84481                | 1 thymidine kinase                                                                                                                          | 71.63                            | 1.26E-11                                     | 25%                      | 46%                      | 4182                     | 13208                           |
|              |                    |                |               |       | 9                 | COG0572      | complex is to tansport electrons from the body and accompanied by the translocation of protons from the mitochondrial matrix to the inter membrane snace  Udk, Uridine kinase [Nucleotide transport and metabolism].                                                                                                                                                                                                                  | 36.38        | 0.002876 | 22%             | 44% | 3159             | 10179 | 9                       | YP_187798               | B deoxynucleoside kinase family protein                                                                                                     | 71.63                            | 1.26E-11                                     | 30%                      | 48%                      | 3177                     | 11-197                          |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                       |              |          |                 |     |                  |       | 10                      | NP_078725               | 5 Deoxynucleoside kinase                                                                                                                    | 69.71                            | 4.81E-11                                     | 29%                      | 52%                      | 3152                     | 4161                            |

| Gene<br>Name | Genome<br>Position           | A.A.<br>length | Peptide<br>Mw | pl    | CDD Hit<br>Number | COGs         | COG Definition                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | Bit<br>Score | E-value   | %<br>Identity F | %<br>ositive | Query<br>from-to | Hit<br>from-to | BLASTp<br>Hit<br>Number         | Hit<br>Accession                                                                 | BLASTp Definition                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | Bit<br>Score                                                | E-value                                                                                      | %<br>dentity P                                       | %<br>ositive                                         | Query<br>from-to                                                   | Hit from-<br>to                                                    |
|--------------|------------------------------|----------------|---------------|-------|-------------------|--------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|-----------|-----------------|--------------|------------------|----------------|---------------------------------|----------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------|----------------------------------------------------------------------------------------------|------------------------------------------------------|------------------------------------------------------|--------------------------------------------------------------------|--------------------------------------------------------------------|
| M427R        | 168662169717                 | 352            | 39,681        | 8.71  |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |           |                 |              |                  |                | 1<br>2<br>3<br>4<br>5<br>6<br>7 | NP_048920<br>NP_077492<br>NP_077600<br>BAD61685                                  | , similar to E. coli ribonucleoside-triphosphate reductase, corresponds to SAIS-M. Archesian Number P78013 AS-M. SAIS-M. SAIS- | 261.92<br>254.22                                            | 1.22E-71<br>2.16E-68<br>4.51E-66<br>1.32E-33<br>6.18E-15<br>1.80E-06<br>9.89E-05             | 43%<br>40%<br>39%<br>35%<br>30%<br>27%<br>26%        | 61%<br>59%<br>58%<br>47%<br>42%<br>36%<br>35%        | 9-352<br>1-352<br>10-352<br>4-199<br>4-166<br>4-212<br>4-194       | 8-356<br>1-347<br>9-350<br>5-200<br>146-303<br>109-277<br>319-490  |
| M430L        | 170803169724                 | 360            | 40,028        | 8.34  | 1                 | COG1910      | COG1910, Periplasmic molybdate-binding protein/domain [Inorganic ion transport and metabolism].                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 35.25        | 0.006527  | 35%             | 49%          | 4083             | 87130          | 1                               | NP_048774                                                                        | \$ A417L                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 212.62                                                      | 1.56E-53                                                                                     | 37%                                                  | 57%                                                  | 1-348                                                              | 1372                                                               |
| M435L        | 171179170838                 | 114            | 13,067        | 10.50 |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |           |                 |              |                  |                | 1                               | NP_048777                                                                        | 7 A420L                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | 70.09                                                       | 2.20E-11                                                                                     | 48%                                                  | 65%                                                  | 45114                                                              | 170                                                                |
|              | 171162171413                 | -              | 10,358        | 11.31 |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |           |                 |              |                  |                |                                 |                                                                                  | d No Hit Found                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |                                                             |                                                                                              |                                                      |                                                      |                                                                    |                                                                    |
|              | 171644171402                 | 81             |               | 7.94  |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |           |                 |              |                  |                |                                 |                                                                                  | d No Hit Found                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |                                                             |                                                                                              |                                                      |                                                      |                                                                    |                                                                    |
|              | 171671172138<br>172162172521 | 156            |               | 4.92  |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |           |                 |              |                  |                | 1                               | NP 048780                                                                        | 0 A423R<br>d No Hit Found                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | 82.42                                                       | 4.24E-15                                                                                     | 36%                                                  | 61%                                                  | 36154                                                              | 33–155                                                             |
|              | 172564172911                 | 116            |               | 4.84  |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |           |                 |              |                  |                | 1                               | NP 048783                                                                        |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 95.90                                                       | 3.72E-19                                                                                     | 42%                                                  | 64%                                                  | 6116                                                               | 4114                                                               |
|              | 173265172918                 |                | 13,195        | 5.80  |                   |              | Thioredoxin, Thioredoxin. Thioredoxins are small enzymes that participate in redox reactions, via the reversible oxidation of an active centre disulfide bond. Some members with only the active site are not                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | 43.29        | 2.50E-05  | 25%             | 49%          | 2593             | 2596           |                                 |                                                                                  | contains thioredoxin active site-like sequence; similar to Synechocystis<br>4 thioredoxin-like protein, corresponds to Swiss-Prot Accession Number<br>P52232                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |                                                             | 2.33E-21                                                                                     | 44%                                                  | 65%                                                  | 2-115                                                              | 7119                                                               |
|              |                              |                |               |       |                   |              | separated from the noise                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |              |           |                 |              |                  |                | 3<br>4<br>5<br>6<br>7<br>8<br>9 | AAC71342<br>AAP04832<br>CAH63538<br>CAH59450<br>BAE56042<br>AAH84521<br>CAG25528 | OCOGOSES: Thick-disulfide isomerase and thioredoxins Private Communication Private Communication Part of Communication Part of Communication Private Comm  | 51.60<br>51.60<br>50.83<br>50.45<br>50.45<br>50.45<br>49.29 | 2.77E-06<br>8.05E-06<br>8.05E-06<br>1.37E-05<br>1.79E-05<br>1.79E-05<br>1.79E-05<br>4.00E-05 | 31%<br>29%<br>29%<br>30%<br>28%<br>30%<br>28%<br>28% | 54%<br>51%<br>55%<br>53%<br>51%<br>51%<br>55%<br>53% | 17-105<br>4-103<br>4-102<br>4-102<br>3-95<br>2-90<br>4-90<br>10-83 | 21-112<br>2-102<br>1-99<br>1-99<br>12-106<br>1-92<br>5-91<br>11-86 |
|              |                              |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |           |                 |              |                  |                |                                 | NP_395913                                                                        | contains thioradovin active site like sequences similar to Synachocyclic                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |                                                             | 4.00E-05                                                                                     | 29%                                                  | 51%                                                  | 11101                                                              | 23114                                                              |
| M448L        | 173669173334                 | 112            | 12,692        | 5.76  | 1                 | COG3118      | COG3118, Thioredoxin domain-containing protein [Posttranslational modification, protein turnover, chaperones].                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 40.66        | 0.00015   | 38%             | 56%          | 5589             | 81115          | 1                               | NP_048784                                                                        | 4 thioredoxin-like protein, corresponds to Swiss-Prot Accession Number<br>P52232 contains thioredoxin active site-like sequence; similar to Synechocystis                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | 53.91                                                       | 1.59E-06                                                                                     | 28%                                                  | 53%                                                  | 1-105                                                              | 7110                                                               |
| M449L        | 174017173706                 | 104            | 11,732        | 8.10  |                   | No Hit Found | ANK. ankvrin repeats: ankvrin repeats mediate protein-protein                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |              |           |                 |              |                  |                | 1                               | NP_048784                                                                        | 4 thioredoxin-like protein, corresponds to Swiss-Prot Accession Number<br>P52232                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 52.37                                                       | 4.60E-06                                                                                     | 32%                                                  | 52%                                                  | 295                                                                | 9106                                                               |
| M451L        | 175368174043                 | 442            | 51,066        | 5.24  | 1                 | cd00204      | ANY, anyon's repeats, misyn's rejease institute to the composition of | 44.69        | 8.95E-06  | 23%             | 41%          | 157287           | 1-126          | 1                               | NP_048786                                                                        |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                                             | 6.47E-55                                                                                     | 33%                                                  | 51%                                                  |                                                                    | 34–466                                                             |
|              |                              |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |           |                 |              |                  |                | 2<br>3<br>4                     | BAB03143                                                                         | 5 protein binding<br>3 ankyrin-like protein<br>2 protein binding                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 55.84<br>53.53<br>51.60                                     | 3.26E-06<br>1.62E-05<br>6.14E-05                                                             | 25%<br>23%<br>22%                                    | 44%<br>43%<br>43%                                    | 51-237<br>49-253<br>49-237                                         | 61-250<br>599-805<br>114-305                                       |
|              | 175405175728                 | 108            | ,             | 9.58  |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |           |                 |              |                  |                |                                 |                                                                                  | d No Hit Found                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |                                                             |                                                                                              |                                                      |                                                      |                                                                    |                                                                    |
|              | 175778176056                 | 93             |               | 10.74 |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |           |                 |              |                  |                |                                 | NP_048957                                                                        |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 91.28                                                       | 8.97E-18                                                                                     | 52%                                                  | 65%                                                  | 11–91                                                              | 12101                                                              |
| M457R        | 176074177201                 | 376            | 42,427        | 10.07 |                   | No Hit Found | GIY-YIG Cterm, GIYX[10-11)YIG tamily of class I homing endonucleases C-terminus (GIY-YIG Cterm). Homing endonucleases                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |              |           |                 |              |                  |                | 1 2                             | NP 048579<br>NP_048983                                                           | contains ATP(GTP-binding motif A     jamillar to Chlorella virus PBCV-1 ORF A231L, corresponds to GenBank     Accession Number U42580                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |                                                             | 3.71E-138<br>1.43E-81                                                                        | 61%<br>44%                                           | 76%<br>62%                                           | 2-376<br>10-336                                                    | 3–383<br>104–428                                                   |
| M460R        | 177267178019                 | 251            | 28,857        | 10.08 | . 1               | cd00283      | promote the mobility of intrin or intenii by recognizing and cleaving a homologous allele that lacks the sequence. They calabyze a double-strand break in the DNA near the insertion sits of that element to facilitate homing at that site. Classa homing endouncleases are sorted into four families based on the presence of these motifs in their respective North Control of the presence of these motifs in their respective North Control of CHY-VIG family. The Certainnis of CHY-VIG family. The DNA-binding domain which is separated from the N-terminus by a long-fine solid provide bring alpha-helix, and a helix-turn-helix. Some also contain a zinc finger (e. I-Ted) which is not required for DNA binding or calalysis, but is a component of the linker and direct the calalytic calalysis, but is a component of the linker and direct the calalytic.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | 39.98        | 0.000211  | 44%             | 57%          | 124–173          | 1165           |                                 |                                                                                  | 7 putative GIY-YIG endonuclease                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 61.62                                                       | 2.46E-08                                                                                     | 28%                                                  | 45%                                                  | 39-212                                                             | 14–187                                                             |
|              |                              |                |               |       | 2                 | smart00465   | GIY-, GIY-YIG type nucleases (URI domain); . GIY-YIG, GIY-YIG datalytic domain. This domain called GIY-YIG is found in the amino terminal region of excinuclease abc subunit c (uvrC), bacteriophage T4 endonucleases segA, segB, segC, segD and segE; it is                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | 36.98        | 0.001988  | 23%             | 43%          | 26-112           | 1-78           | 2                               | . AAU16837                                                                       | 7 GIY-YIG catalytic domain containing protein; possible intron encoded endonuclease                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 57.00                                                       | 6.06E-07                                                                                     | 30%                                                  | 51%                                                  | 26149                                                              | 2-136                                                              |
|              |                              |                |               |       | 3                 | pfam01541    | also found in putative endonucleases encoded by group l introns of fungi<br>and phage. The structure of I-TevI a GIY-YIG endonuclease, reveals a<br>novel alphabeta-fold with a central three-stranded antiparallel beta-sheet<br>flanked by three helices. The most conserved and putative catalytic<br>residues are located on a shallow, concave surface and include a metal                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 36.68        | 0.002641  | 26%             | 43%          | 29115            | 4-85           | 3                               | CAA38813                                                                         | 3 GIY ND1 i4 grp IB protein b                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | 51.22                                                       | 3.33E-05                                                                                     | 30%                                                  | 45%                                                  | 14167                                                              | 54-220                                                             |
|              |                              |                |               |       |                   |              | coordination site                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |              |           |                 |              |                  |                | 4                               | NP_074961                                                                        | 1 orf261                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 51.22                                                       | 3.33E-05                                                                                     | 30%                                                  | 45%                                                  | 14167                                                              | 53-219                                                             |
| M463L        | 179395178097                 | 433            | 47,651        | 6.92  | 1                 | pfam04451    | Capsid_Iridovir, Iridovirus major capsid protein. This family includes the<br>major capsid protein of iridoviruses, chlorella virus and Spoopdera<br>accovirus, which are all dLDNA viruses with no RNA stage. This is the<br>protein. In Chlorella virus MT325 the major capsid protein is a<br>"hornoration".                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 478.65       | 1.87E-136 | 54%             | 65%          | 1429             | 1443           |                                 |                                                                                  | 2 major capsid protein Vp49                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 763.45                                                      | 0.00E+00                                                                                     | 88%                                                  | 89%                                                  | 1-433                                                              | 1-432                                                              |
|              |                              |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |           |                 |              |                  |                | 2                               | NP_048787                                                                        | 7 PBCV-1 major capsid protein Vp54, corresponds to GenBank Accession<br>Number M85052<br>B major capsid protein Vp54                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |                                                             | 0.00E+00<br>0.00E+00                                                                         | 75%<br>74%                                           | 80%<br>80%                                           | 1433                                                               | 1437                                                               |
|              |                              |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |           |                 |              |                  |                | 3<br>4<br>5                     | BAA7660                                                                          | 5 major capsid protein Vp54  1 major capsid protein MCP1  D major capsid protein                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 643.27                                                      | 0.00E+00<br>0.00E+00<br>0.00E+00                                                             | 74%<br>74%<br>73%                                    | 80%<br>79%                                           | 1-433                                                              | 1-437<br>1-437<br>1-436                                            |
|              |                              |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |           |                 |              |                  |                | 6                               | 1M4X C                                                                           | Chain C. Phov-1 Virus Cansid. Quasi-Atomic Model                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 603.98                                                      | 3.12E-171<br>3.12E-171                                                                       | 73%                                                  | 79%<br>79%                                           | 25-433                                                             | 1-413                                                              |
|              |                              |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |           |                 |              |                  |                | 7<br>8                          | BAE06835                                                                         | Chain D, The Structure Of Major Capsid Protein Of A Large, Lipid Containino. Dna Virus 5 hypothetical major capsid protein                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | 316.62                                                      | 9.94E-85                                                                                     | 41%                                                  | 54%                                                  | 1433                                                               | 1440                                                               |
|              |                              |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |           |                 |              |                  |                | 9<br>10                         | NP 048359                                                                        | O contains aminoacyl-IRNA synthetase class-II signature imiliar to Chlorella virus PBCV-1 ORF A11L, corresponds to GenBank Accession Number U42580                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |                                                             | 4.34E-64<br>1.65E-63                                                                         | 37%<br>35%                                           | 52%<br>51%                                           | 1-433<br>3-433                                                     | 1403<br>2400                                                       |

| Gene<br>Name | Genome<br>Position           | A.A.<br>length | Peptide<br>Mw | pl           | CDD Hit<br>Number | COGs         | COG Definition                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | Bit<br>Score | E-value  | %<br>Identity F | %          | Query<br>from to | Hit<br>from to | BLASTp<br>Hit     | Hit<br>Accession                   | BLASTp Definition                                                                                                                                                       | Bit<br>Score                              | E-value                                                  | %<br>Identity P                 | %                               |                                                | Hit from-<br>to                                  |
|--------------|------------------------------|----------------|---------------|--------------|-------------------|--------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|----------|-----------------|------------|------------------|----------------|-------------------|------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------|----------------------------------------------------------|---------------------------------|---------------------------------|------------------------------------------------|--------------------------------------------------|
| Numb         | ·                            | iong           |               |              |                   |              | GIY-YIG_Cterm, GIYX(10-11)YIG family of class I homing endonucleases Cherminus (GIY-YIG_Cterm). Homing endonucleases promote the mobility of intron or inten by recognizing and cleaving a homologous aflete that lacks the sequence. They catalyze a double-stand break in the DNA near the insertion site of that element to facilitate homing at that site. Class I homing endonucleases are sorted into four families based on the orseen cell freshee molfis in their respective N-                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 50010        |          |                 | 001470     |                  |                | Number            | 7.00000.0                          |                                                                                                                                                                         | 555.5                                     |                                                          |                                 | 00.1110                         |                                                | .0                                               |
| M465L        | 180410179550                 | 287            | 32,625        | 8.99         | 1                 | cd0028       | tamiles based on the Speakene or meast mosts in their respectives. If a termin LACIUDAD, His-Cys box, NH, and GHY-NT, This CD contains several but not at members of the GHY-NTG family. The C-terminus of the GHY-NTG family. The C-terminus of a terminus by a long fields finiter. The DNA-binding domain consists of a minor-groove binding sights-helix, and a helix-turn-helix. Some also contain a zinc fining (i.e. I-lev) which is not required for DNA binding or catalysis, but is a component of the linker and directs the catalytic ordination to the control of the catalytic ordination to | 68.10        | 7.64E-13 | 57%             | 73%        | 116179           | 1-64           | 1                 | NP_048671                          | 1 A315L                                                                                                                                                                 | 237.65                                    | 3.16E-61                                                 | 46%                             | 60%                             | 1283                                           | 1-240                                            |
|              |                              |                |               |              | 2                 | smart00465   | insertion site GIYc, GIY-YIG type nucleases (URI domain); IENR1. Intron encoded nuclease repeat motif: Repeat of unknown                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 43.52        | 1.96E-05 | 32%             | 54%        | 190              | 1-83           |                   |                                    | 1 PBCV-1 33kd peptide                                                                                                                                                   | 221.86                                    | 1.79E-56                                                 | 44%                             | 61%                             | 3285                                           | 6249                                             |
|              |                              |                |               |              | 3                 | smart00497   | Function, but possibly DNA-binding via helik-turn-heim molif repeat of usin/tuwn function, but possibly DNA-binding via helik-turn-heim molif (Porting, GIY-YIG, Gaty-YIG catalytic domain. This domain called GIY-YIG is toud in the amin terminal region of excinuclease abe subunit c (urvC), bacteriophage T4 endonucleases segA, segB, segC, segD and segE; till sale found in the amin terminal region of excinucleases of this given the segond of the segond of the segOrgen of the se | 41.66        | 8.38E-05 | 37%             | 58%        | 232285           | 1-53           |                   |                                    | s similar to Chlorella virus PBCV-1 ORF A315L, corresponds to GenBank Accession Number U42580                                                                           | 192.20                                    | 1.52E-47                                                 | 40%                             | 52%                             | 1286                                           | 1-226                                            |
|              |                              |                |               |              | 4                 | pfam0154     | also found in putative endonucleases encoded by group I introns of fungi<br>I and phage. The structure of I-reVi a GIY-YIG endonuclease, reveals a<br>novel alpha/beta-fold with a central three-stranded antiparallel beta-sheet<br>flanked by three helices. The most conserved and putative catalytic<br>residues are located on a shallow, concave surface and include a metal<br>concrinations and                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | 37.06        | 0.001784 | 30%             | 56%        | 159              | 164            | 4                 | NP_048851                          | similar to PBCV-1 ORF A315L, corresponds to GenBank Accession<br>Number M74440                                                                                          | 137.50                                    | 4.45E-31                                                 | 35%                             | 48%                             | 1-286                                          | 1211                                             |
|              |                              |                |               |              |                   |              | 144 EUROPE AUG                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |              |          |                 |            |                  |                | 5                 | AAK09365<br>NP 899393              | 5 intron encoded Bmol<br>3 SegD                                                                                                                                         | 82.42<br>79.34                            | 1.70E-14<br>1.44E-13                                     | 31%<br>30%                      | 45%<br>47%                      | 4210<br>1219                                   | 5217<br>1221                                     |
|              |                              |                |               |              |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |          |                 |            |                  |                | 7                 | AAC49244                           |                                                                                                                                                                         | 75.87<br>73.17                            | 1.59E-12<br>1.03E-11                                     | 31%<br>25%                      | 46%<br>45%                      | 4173<br>4254                                   | 76-239<br>5215                                   |
|              |                              |                |               |              |                   |              | Glycos_transf_2, Glycosyl transferase. Diverse family, transferring sugar                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |              |          |                 |            |                  |                | 10                | YP 293795                          | 5 putative endonuclease<br>5 unnamed protein product                                                                                                                    | 71.63<br>71.25                            | 3.00E-11                                                 | 36%<br>30%                      | 54%<br>45%                      | 2–112<br>13–183                                |                                                  |
| M467R        | 180530183076                 | 849            | 95,588        | 6.42         |                   |              | 5 from UDP-glucose, UDP-N-acetyl- galactosamine, GDP-mannose or<br>CDP-abequose, to a range of substrates including cellulose, dolichol                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |              | 3.41E-07 | 30%             | 43%        | 257384           |                |                   | NP_048462                          |                                                                                                                                                                         |                                           | 0.00E+00                                                 | 64%                             | 77%                             | 374848                                         | 5477                                             |
|              |                              |                |               |              | 2                 |              | amosonate and reichoic acids  Woak, Glycosyltransferases involved in cell wall biogenesis [Cell envelope biogenesis, outer membrane].                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 47.44        | 1.35E-06 | 18%             | 35%        | 252526           |                |                   | NP_048459                          |                                                                                                                                                                         |                                           | 3.58E-138                                                | 58%                             | 76%                             | 1373                                           |                                                  |
|              |                              |                |               |              | 3                 | COG121       | 5 COG1215, Glycosyltransferases, probably involved in cell wall biogenesis ICell envelope biogenesis. outer membranel.  6 COG1216, Predicted glycosyltransferases [General function prediction                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | 44.15        | 1.51E-05 | 25%<br>26%      | 38%<br>46% |                  | 53167          | 3                 |                                    | 5 possible glycosyltransferase                                                                                                                                          |                                           | 2.66E-16                                                 | 32%<br>30%                      | 46%<br>45%                      | 255427                                         | 6189                                             |
|              |                              |                |               |              | 4                 | COG121       | 6 noted Exostosin, Exostosin family. The EXT family is a family of tumour<br>suppressor genes. Mutations of EXT1 on 82,24.1, EXT2 on 11p11-13,<br>and EXT3 on 10p have been associated with the autosomal dominant<br>disorder known as hereditary multiple exostoses (HME). This is the most<br>common known skeletal dysplassia. The chromosomal locations of other                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 38.96        | 0.000491 | 26%             | 46%        | 253370           | 3116           | 4                 | CAG34747                           | 7 hypothetical protein                                                                                                                                                  | 89.74                                     | 4.53E-16                                                 | 30%                             | 45%                             | 1-226                                          | 1-225                                            |
|              |                              |                |               |              | 5                 | pfam03016    | 6 common known skeletal dysplasia. The chromosomal locations of other<br>EXT genes suggest association with other forms of neoplasia. EXT1 and<br>EXT2 have both been shown to encode a heparan sulphate polymerase<br>with both D-glucuronyl (GIcla, and N-acetyl-Djucosaminoglycus)<br>(GicNAC) transferase activities. The nature of the defect in heparan<br>sulphate historythesis in IMLET is unchazer.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | 35.03        | 0.007743 | 23%             | 43%        | 658819           | 177-325        | 5                 | ZP_00202013                        | 3 COG0463: Glycosyltransferases involved in cell wall biogenesis                                                                                                        | 86.27                                     | 5.01E-15                                                 | 32%                             | 49%                             | 255427                                         | 48-231                                           |
|              |                              |                |               |              |                   |              | Chinada in Compact in Salar C 10 acai                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |              |          |                 |            |                  |                | 8                 | AAU03775<br>ZP 00340129            | 1 unknown 7 Coll'21e: Predicted qlycosylfransferases 5 conserved hypothetical protein 9 hypothetical protein RakaH01000503 5 COGO457: PGG: TPR repeat                   | 81.65<br>79.34<br>58.15<br>57.00<br>55.45 | 6.13E-13<br>1.46E-06                                     | 26%<br>29%<br>26%<br>30%<br>29% | 40%<br>44%<br>50%<br>51%<br>50% | 255496<br>255428<br>59207<br>59207<br>59207    | 6-259<br>29-215<br>103-250<br>107-254<br>107-254 |
| M472R        | 183130183585                 | 152            | 16,607        | 8.18         |                   | No Hit Found | i                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |              |          |                 |            |                  |                | 1                 | NP_048789                          | 9 Lys-rich                                                                                                                                                              | 114.01                                    | 1.30E-24                                                 | 46%                             | 67%                             | 1109                                           | 1109                                             |
| M475R        | 183617183964                 | 116            | 12,729        | 11.93        |                   | No Hit Found | i                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |              |          |                 |            |                  |                | 1 2               | NP_048927<br>NP_048792             |                                                                                                                                                                         | 157.15<br>56.61                           |                                                          | 79%<br>61%                      | 89%<br>70%                      | 1101<br>144                                    | 1101<br>447                                      |
| M476L        | 184825184220                 | 202            | 22,832        | 10.53        | 1                 | pfam0585     | MC1, Non-histone chromosomal protein MC1. This family consists of<br>4 archaeal chromosomal protein MC1 sequences which protect DNA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 48.51        | 6.22E-07 | 37%             | 53%        | 106169           | 2-70           | 1                 | NP_048794                          | similar to Methanothrix chromosomal protein MC1A, corresponds to<br>Swiss-Prot Accession Number P15251                                                                  | 141.35                                    | 1.57E-32                                                 | 70%                             | 80%                             | 101200                                         | 1100                                             |
|              |                              |                |               |              |                   |              | adainst thermal denaturation                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |              |          |                 |            |                  |                | 3                 | P15249                             | 9 Chromosomal protein MC1a<br>0 Chromosomal protein MC1b                                                                                                                | 54.68<br>51.22                            | 2.13E-05                                                 | 42%<br>45%                      | 50%<br>60%                      | 108191<br>108167                               | 484<br>462                                       |
| M478L        | 185179184862                 | 106            | 12,215        | 4.88         |                   | No Hit Found | 1                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |              |          |                 |            |                  |                | 1<br>2<br>3<br>4  | NP 077527                          | 5 A408L                                                                                                                                                                 | 113.62<br>78.57<br>50.83<br>49.68         | 6.17E-14<br>1.38E-05                                     | 48%<br>33%<br>30%<br>33%        | 74%<br>61%<br>50%<br>51%        | 4-105<br>3-102<br>3-101<br>3-103               | 8110<br>40139<br>6110<br>5110                    |
| M479R        | 185275185766                 | 164            | 18.674        | 5.08         |                   | No Hit Found | 1                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |              |          |                 |            |                  |                | 1                 | NP_048768                          |                                                                                                                                                                         | 79.34                                     | 4.24E-14                                                 | 37%                             | 56%                             | 1131                                           | 1133                                             |
| M482R        | 185802186338                 | 179            | 20,331        | 7.46         |                   | No Hit Found | 1                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |              |          |                 |            |                  |                |                   | NP_048769                          |                                                                                                                                                                         | 178.72                                    | 6.55E-44                                                 | 49%                             | 65%                             | 1178                                           | 1179                                             |
| M484L        | 187082186345                 | 246            | 26,921        | 9.59         |                   | No Hit Found | 1                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |              |          |                 |            |                  |                | 1                 | NP_048770                          | Gln-rich, QQQM(4x); similar to human transcription factor TFIID, corresponds to Swiss-Prot Accession Number P20226                                                      | 150.21                                    | 5.05E-35                                                 | 37%                             | 50%                             | 2245                                           | 3244                                             |
| M485R        | 187170187862                 | 231            | 25,816        | 4.81         |                   | No Hit Found | 1                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |              |          |                 |            |                  |                | 1                 | NP_048765                          | 5 A408L                                                                                                                                                                 | 209.92<br>64.31                           | 4.76E-53<br>3.23E-09                                     | 51%<br>26%                      | 75%<br>49%                      | 1187                                           | 37-233                                           |
|              |                              |                |               |              |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |          |                 |            |                  |                | 2<br>3<br>4       | NP_048767<br>AAR26867              | 7 A410L<br>7 FirrV-1-A43                                                                                                                                                | 59.69<br>58.54                            | 3.23E-09<br>7.96E-08<br>1.77E-07                         | 26%<br>29%<br>24%               | 53%<br>53%                      | 5143<br>497<br>4130                            | 7108                                             |
| M488L        | 188632187886                 | 249            | 26,091        | 12.88        |                   | No Hit Found | 1                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |              |          |                 |            |                  |                | 1                 | ZP_00283849                        | OCOG4991: Uncharacterized protein with a bacterial SH3 domain homologue collagen triple helix repeat:antifreeze protein, type I                                         | 98.21<br>85.89                            | 2.33E-19<br>1.20E-15                                     | 36%<br>37%                      | 61%<br>47%                      | 125235<br>125235                               |                                                  |
|              |                              |                |               |              |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |          |                 |            |                  |                | 3                 | AAK14819                           | e collagen triple helix repeat:antifreeze protein, type I<br>9 hsp70-like protein<br>0 procollagen type I alpha 2 chain                                                 | 85.89<br>85.11<br>84.34                   | 2.04E-15<br>3.49E-15                                     | 52%                             | 52%<br>45%                      | 125235<br>138225<br>126235                     | 257-344                                          |
|              |                              |                |               |              |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |          |                 |            |                  |                | 5                 | AAB96638                           | procollagen type i aipna 2 chain<br>8 precollagen D<br>1 precollagen-D                                                                                                  | 83.57<br>83.19                            | 5.95E-15                                                 | 38%<br>37%<br>37%               | 45%<br>45%                      | 125235                                         | 354464                                           |
|              |                              |                |               |              |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |          |                 |            |                  |                | 7<br>8<br>9<br>10 | CAE29034<br>NP_031761<br>NP_892013 | i precoivagen-U  Collagen triple helix repeat:Antifreeze protein, type I  procollagen, type IV, alpha 4  collagen, type I, alpha 2  hsp70 interacting protein, putative | 83.19<br>83.19<br>82.80<br>82.80<br>82.42 | 7.77E-15<br>7.77E-15<br>1.01E-14<br>1.01E-14<br>1.33E-14 | 36%<br>40%<br>38%<br>51%        | 45%<br>45%<br>44%<br>44%<br>51% | 125235<br>123235<br>126235<br>126235<br>134225 | 6121<br>457573                                   |
| M491R        | 188838189983                 | 382            | 43,963        | 8.00         |                   | No Hit Found | 1                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |              |          |                 |            |                  |                |                   |                                    | d No Hit Found                                                                                                                                                          |                                           |                                                          |                                 |                                 |                                                |                                                  |
| M495R        | 190020190646                 | 209            | 23,349        | 8.61         |                   | No Hit Found | 1                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |              |          |                 |            |                  |                | 1                 | NP_048764                          | 4 A407L                                                                                                                                                                 | 309.30                                    | 4.72E-83                                                 | 70%                             | 82%                             | 1209                                           | 1209                                             |
| M496L        | 191314190700                 | 205            | 22,858        | 8.17         |                   | No Hit Found | 1                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |              |          |                 |            |                  |                |                   |                                    | 7 contains Gln-rich, neutral zinc metallopeptidase, zinc binding region signature                                                                                       |                                           |                                                          | 63%                             | 79%                             | 1167                                           |                                                  |
|              |                              |                |               |              |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |          |                 |            |                  |                |                   |                                    | 9 metal-dependent hydrolase                                                                                                                                             |                                           | 1.64E-08                                                 | 29%                             | 47%                             | 4181                                           |                                                  |
|              | 191377191910<br>191937192380 | 178<br>148     |               | 8.62<br>8.39 |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |          |                 |            |                  |                |                   | NP_048879                          |                                                                                                                                                                         | 230.72<br>130.95                          | 1.43E-59<br>1.04E-29                                     | 74%<br>52%                      | 83%<br>63%                      | 27175<br>18145                                 | 20168<br>1127                                    |
| INIDUTR 1    | 101937192380                 | 148            | 10,447        | 8.39         |                   | NO DIE FOUND | •                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |              |          |                 |            |                  |                | 1                 | NF_048882                          | L MOZOR                                                                                                                                                                 | 130.95                                    | 1.04E-29                                                 | 52%                             | 63%                             | 16145                                          | 1-12/                                            |

| Gene<br>Name | Genome<br>Position           | A.A.   | Peptide<br>Mw     | pl    | CDD Hit | COGs         | COG Definition                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | Bit<br>Score | E-value  | %<br>Identity Po | %       | Query   | Hit     | BLASTp<br>Hit | Hit<br>Accessi    | ion BLASTp Definition                                                                                                                                    | Bit<br>Score               | E-value lo                        | %<br>dantitu Da   | %                 | Query                        | Hit from-                    |
|--------------|------------------------------|--------|-------------------|-------|---------|--------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|----------|------------------|---------|---------|---------|---------------|-------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------|-----------------------------------|-------------------|-------------------|------------------------------|------------------------------|
| Name         | Position                     | iengtn | ww                | ·     | Number  |              | GIY-YIG Clerm, GIYX(10-11)YIG family of class I homing endonucleases C-terminus (GIY-YIG_Cterm). Homing endonucleases promote the mobility of intron or intein by recognizing and cleaving a homologous allele that lacks the sequence. They catalyze a double-                                                                                                                                                                                                                                                               | Score        |          | identity P       | ositive | rrom-to | Trom-to | Number        | Accessi           | ion                                                                                                                                                      | Score                      | 10                                | ientity Po        | SITIVE            | trom-to                      | το                           |
| M502L        | 193151192417                 | 245    | 27,605            | 9.83  | 1       | cd0028       | strand break in the DNA near the insertion site of that element to facilitate homing at that site. Class I homing endonucleases are sorted into four families based on the presence of these motifs in their respective N-3 termin: LAGLIDADG, His-Cys box, RNH, and GIY-YIG. This CD contains several but not all members of the GIY-VIG family The Carpinius of                                                                                                                                                             | 57.70        | 1.00E-09 | 43%              | 55%     | 140241  | 16113   | 1             | NP_048            | 3671 A315L                                                                                                                                               | 245.74                     | 8.76E-64                          | 53%               | 69%               | 1-241                        | 1-240                        |
|              |                              |        |                   |       |         |              | several but not all members of the GIV-YIG family. The C-terminus of<br>GIV-YIG is a DNA-binding domain which is separated from the N-<br>terminus by a long, flexible linker. The DNA-binding domain consists of a<br>minor-groove binding aligna-helic, and a helick-turn-helic. Some also<br>contain a zinc finger (ie. I-Terly which is not required for DNA binding or<br>catalysis, but is a component of the linker and directs the catalytic<br>domain to cleave the horning site at a fixed distance from the infron |              |          |                  |         |         |         |               |                   |                                                                                                                                                          |                            |                                   |                   |                   |                              |                              |
|              |                              |        |                   |       | 2       | smart00465   | insertion eite 5 GIYc, GIY-YIG type nucleases (URI domain); .                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | 44.68        | 9.16E-06 | 34%              | 51%     | 190     | 1-83    | 2             | NP_049            | similar to Chlorella virus PBCV-1 ORF A315L, corresponds to GenBank<br>Accession Number L142580                                                          | 229.95                     | 4.97E-59                          | 49%               | 66%               | 1244                         | 1226                         |
|              |                              |        |                   |       | 3       | smart00497   | IENR1, Intron encoded nuclease repeat motif; Repeat of unknown function, but possibly DNA-binding via helix-turn-helix motif (Ponting,                                                                                                                                                                                                                                                                                                                                                                                        | 40.50        | 0.000189 | 40%              | 58%     | 190243  | 1-53    | 3             | NP_048            | 3641 PBCV-1 33kd peptide                                                                                                                                 | 175.64                     | 1.11E-42                          | 42%               | 60%               | 3-241                        | 6247                         |
|              |                              |        |                   |       |         |              | unnuthished) GIY-YIG, GIY-YIG catalytic domain. This domain called GIY-YIG is found in the amino terminal region of excinuclease abc subunit c (uvrC), bacteriophage T4 endonucleases segA, segB, segC, segD and segE; it is                                                                                                                                                                                                                                                                                                  |              |          |                  |         |         |         |               |                   |                                                                                                                                                          |                            |                                   |                   |                   |                              |                              |
|              |                              |        |                   |       | 4       | pfam0154     | also found in putative endonucleases encoded by group l introns of fungi<br>and phage. The structure of I-revi a GIY-YIC endonuclease, reveals a<br>novel alpha/beta-fold with a central three-stranded antiparallel beta-sheet<br>flanked by three helices. The most conserved and putative catalotic<br>residues are located on a shallow, concave surface and include a metal<br>coordination side.                                                                                                                        | 40.15        | 0.00022  | 27%              | 45%     | 187     | 1–89    | 4             |                   | similar to PBCV-1 ORF A315L, corresponds to GenBank Accession<br>Number M74440                                                                           | 165.24                     | 1.50E-39                          | 45%               | 62%               | 1196                         | 1179                         |
|              |                              |        |                   |       |         |              | coordination site                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |              |          |                  |         |         |         | 5<br>6        | YP 293            | 3393 SegD<br>3795 putative endonuclease                                                                                                                  | 95.90<br>75.10             | 1.12E-18<br>2.05E-12              | 36%<br>38%        | 53%<br>58%        | 1190<br>2107                 | 1184<br>3111                 |
|              |                              |        |                   |       |         |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |          |                  |         |         |         | 7<br>8        | CAA38             | 3365 intron encoded Bmol<br>3804 GIY COII i1 grp IB protein                                                                                              | 73.94<br>68.94             | 4.56E-12<br>1.47E-10              | 31%<br>29%        | 48%<br>47%        | 4216<br>4222                 | 5242<br>74286                |
|              |                              |        |                   |       |         |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |          |                  |         |         |         | 9<br>10       | AAC49             | 9244 ORF301<br>9588 group I intron GIY-YIG endonuclease                                                                                                  | 67.01<br>65.47             | 5.57E-10<br>1.62E-09              | 28%<br>29%        | 43%<br>47%        | 4179<br>4212                 | 76236<br>5215                |
| M504R        | 193217193495                 | 93     | 11,055            | 10.73 |         | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |          |                  |         |         |         | 1             | NP_048            | 3883 A527R                                                                                                                                               | 76.64                      | 2.29E-13                          | 60%               | 74%               | 669                          | 569                          |
| M506L        | 193756193547                 | 70     | 7,977             | 10.35 |         | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |          |                  |         |         |         |               | No Hit For        | ound No Hit Found                                                                                                                                        |                            |                                   |                   |                   |                              |                              |
| M507L        | 194020193784                 | 79     | 8,652             | 9.93  |         | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |          |                  |         |         |         | 1             | NP 048            | 888 A532L                                                                                                                                                | 95.13                      | 6.29E-19                          | 60%               | 77%               | 173                          | 174                          |
| M508R        | 194269195849                 | 527    | 57,118            | 4.98  |         | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |          |                  |         |         |         | 1 2           | AAA66             | 8889 A533R<br>8400 unknown protein                                                                                                                       |                            | 1.78E-118<br>6.75E-118            | 58%<br>58%        | 72%<br>71%        | 1362<br>1362                 | 1365<br>1365                 |
|              |                              |        |                   |       |         |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |          |                  |         |         |         | 3<br>4        | NP_077            | 3890 a534R<br>1576 EsV-1-91                                                                                                                              |                            | 3.21E-35<br>3.46E-05              | 76%<br>28%        | 87%<br>47%        | 438526<br>3161               | 16104<br>8158                |
|              | 196067195852                 |        |                   | 4.28  |         | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |          |                  |         |         |         |               | NP_048            |                                                                                                                                                          |                            | 3.41E-09                          | 45%               | 66%               | 172                          | 171                          |
|              | 196357196121                 | 79     |                   | 10.91 |         | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |          |                  |         |         |         | 1             |                   | 3892 A536L                                                                                                                                               | 62.77                      | 3.46E-09                          | 41%               | 55%               | 179                          | 173                          |
|              | 196755197237<br>201296197730 | 161    | 17,937<br>120,260 | 9.41  |         | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |          |                  |         |         |         |               |                   | ound No Hit Found                                                                                                                                        | 047.05                     | 6 95F-175                         | 34%               | 47%               | 821180                       | 11157                        |
| M518L 2      | 201296197730                 | 1189   | 120,260           | 4.65  |         | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |          |                  |         |         |         | 2             | AAA66             | 1896 A540L<br>\$404 unknown protein<br>1302 outer membrane protein                                                                                       | 617.85<br>140.20<br>125.18 | 6.95E-175<br>4.25E-31<br>1.41F-26 | 34%<br>41%<br>28% |                   | 821180<br>10041180<br>383917 | 11157<br>70273<br>196765     |
|              |                              |        |                   |       |         |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |          |                  |         |         |         | 4             | AAX44             | outer membrane protein<br>1675 possible T4-like proximal tail fiber<br>1161 Hep Hag                                                                      | 122.09<br>116.70           | 1.20E-25<br>5.03E-24              | 26%<br>23%        | 38%<br>35%        | 342910<br>132913             | 180-765<br>180-734<br>15-776 |
|              |                              |        |                   |       |         |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |          |                  |         |         |         | 6             | ZP 00532          | 2602 Hep Hag 297 hypothetical protein AsucDRAFT 1054                                                                                                     | 100.91<br>94.74            | 2.86E-19<br>2.05E-17              | 23%<br>21%        | 36%<br>33%        | 75669<br>16903               | 307865<br>2821213            |
|              |                              |        |                   |       |         |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |          |                  |         |         |         | 8             | CAH36             | 1974 outer membrane protein<br>1964 putative membrane protein                                                                                            | 94.36<br>92.82             | 2.67E-17<br>7.78E-17              | 21%<br>23%        | 33%<br>34%        | 39903<br>2811098             | 55-890<br>81-736             |
|              |                              |        |                   |       |         |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |          |                  |         |         |         | 10            | AAU49             | 9476 haemaqluttinin family protein                                                                                                                       | 92.82                      | 7.78E-17                          | 23%               | 34%               | 2811098                      | 81-736                       |
| M522L :      | 203211201367                 | 615    | 68,418            | 8.61  |         | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |          |                  |         |         |         | 1             | NP_048            | contains Pro-rich Px motif, PAPK (19X); similar to Arabidopsis anter-<br>specific Pro-rich protein, corresponds to Swiss-Prot Accession Number<br>P40602 | 61.23                      | 1.18E-07                          | 20%               | 35%               | 1-295                        | 1266                         |
| M525L 2      | 204063203305                 | 253    | 28,489            | 7.70  |         | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |          |                  |         |         |         | 2             | NP 048            | ones similar to Chlorella virus PBCV-1 ORF A450R, corresponds to GenBank                                                                                 | 219.16<br>126.33           | 9.37E-56<br>8.26E-28              | 63%<br>32%        | 79%<br>49%        | 86248<br>2249                | 11173<br>8255                |
|              |                              |        |                   |       |         |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |          |                  |         |         |         | 3             | NP_048            | Accession Number U42580<br>similar to PBCV-1 ORF A275R, encoded by GenBank Accession Number<br>U42580                                                    | 112.85                     | 9.45E-24                          | 31%               | 48%               | 1-247                        | 1-246                        |
|              |                              |        |                   |       |         |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |          |                  |         |         |         | 4             | NP_048            | 3629 Similar to PBCV-1 ORF A79R, corresponds to GenBank Accession                                                                                        | 103.61                     | 5.74E-21                          | 27%               | 48%               | 1246                         | 4248                         |
|              |                              |        |                   |       |         |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |          |                  |         |         |         | 5             | NP_048            | 8427 A79R<br>8525 A177R                                                                                                                                  | 97.06<br>91.28             | 5.37E-19<br>2.95E-17              | 30%<br>25%        | 46%<br>45%        | 1-226<br>1-250               | 1218<br>4245                 |
|              |                              |        |                   |       |         |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |          |                  |         |         |         | 7             | NP_049            | similar to Chlorella virus PBCV-1 ORF A450R, corresponds to GenBank<br>Accession Number U42580                                                           | 63.54                      | 6.58E-09<br>6.16E-07              | 50%<br>28%        | 68%<br>47%        | 158                          | 154<br>1134                  |
|              |                              |        |                   |       |         |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |          |                  |         |         |         | 9<br>10       | AAU06<br>AAU06    | hypothetical protein A275R<br>hypothetical protein A275R                                                                                                 | 57.00<br>55.84             | 6.16E-07<br>1.37E-06              | 26%<br>26%        | 46%<br>46%        | 86246<br>86246               | 7167<br>7167                 |
| M526L 2      | 207559204266                 | 1098   | 110,869           | 6.64  |         | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |          |                  |         |         |         | 1 2           | ZP_00950          | 8896 A540L<br>0302 outer membrane protein                                                                                                                | 147.13                     | 5.99E-165<br>3.19E-33             | 33%<br>25%        | 47%<br>39%        | 821097<br>29599              | 11164<br>197773              |
|              |                              |        |                   |       |         |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |          |                  |         |         |         | 3             | AAA66<br>ZP_00533 | 8404 unknown protein<br>8161 Hep_Hag                                                                                                                     | 136.35<br>129.41           | 5.64E-30<br>6.89E-28              | 33%<br>23%        | 50%<br>37%        | 8851097<br>12710             | 32-280<br>54-755             |
|              |                              |        |                   |       |         |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |          |                  |         |         |         | 5<br>6        | CAH36             | 1675 possible T4-like proximal tail fiber<br>1064 putative membrane protein                                                                              | 125.56<br>122.48           | 9.95E-27<br>8.42E-26              | 24%<br>25%        | 37%<br>40%        | 296967<br>360988             | 14-705<br>77-680             |
|              |                              |        |                   |       |         |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |          |                  |         |         |         | 8             | ZP 00488          | 3391 COG5295: Autotransporter adhesin<br>3067 COG5295: Autotransporter adhesin                                                                           | 122.48<br>121.32<br>120.94 | 8.42E-26<br>1.88E-25<br>2.45E-25  | 25%<br>25%        | 40%<br>40%<br>40% | 360988<br>360988<br>360988   | 61-664<br>61-664<br>77-680   |
|              |                              |        |                   |       |         |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |              |          |                  |         |         |         | 10            | ZP_00445          | 9476 haemaqluttinin family protein<br>997 COG5295: Autotransporter adhesin                                                                               | 120.94                     | 2.45E-25<br>2.45E-25              | 25%<br>25%        | 40%               | 360988                       | 61-664                       |
|              |                              |        |                   |       |         |              | deoxycytidylate_deaminase, Deoxycytidylate deaminase domain.<br>Deoxycytidylate deaminase catalyzes the deamination of dCMP to                                                                                                                                                                                                                                                                                                                                                                                                |              |          |                  |         |         |         |               |                   |                                                                                                                                                          |                            |                                   |                   |                   |                              |                              |
| M530L :      | 208054207623                 | 144    | 16,139            | 5.38  | 1       | cd0128       | dUMP, providing the nucleotide substrate for thymidylate synthase. The 3 enzyme binds Zn++, which is required for catalytic activity. The activity of the enzyme is allosterically regulated by the ratio of dCTP to dTTP not only in eukaryotic cells but also in T-even phage-infected Escherichia                                                                                                                                                                                                                          | 130.40       | 1.48E-31 | 43%              | 57%     | 17-135  | 11131   | 1             | NP_048            | similar to Vibrio fischeri dCMP deaminase, corresponds to Swiss-Prot<br>Accession Number P33968                                                          | 205.68                     | 3.29E-52                          | 66%               | 82%               | 5-144                        | 3142                         |
|              |                              |        |                   |       | 2       | COG213       | coli with dCTP acting as an activator and dTTP as an inhibitor                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 100.20       | 3 26F-25 | 34%              | 53%     | 2 444   | 5148    | 2             | VP 222            | 8954 decovered dela decompose                                                                                                                            | 122.00                     | 4 78F-27                          | 50%               | 65%               | 9-139                        | 110-237                      |
|              |                              |        |                   |       | 3       |              | metabolism1.  dCMP_cyt_deam, Cytidine and deoxycytidylate deaminase zinc-binding                                                                                                                                                                                                                                                                                                                                                                                                                                              |              | 1.05E-18 | 42%              | 57%     |         | 9100    |               |                   | 9954 deoxycytidylate deaminase<br>2863 COG2131: Deoxycytidylate deaminase                                                                                |                            | 4.78E-27<br>1.88E-23              | 43%               | 63%               | 21141                        | 20-139                       |
|              |                              |        |                   |       | Ü       |              | recion<br>cytidine_deaminase-like, Cytidine and deoxycytidylate deaminase zinc-<br>binding region. The family contains cytidine deaminases, nucleoside                                                                                                                                                                                                                                                                                                                                                                        |              |          |                  |         |         |         |               |                   |                                                                                                                                                          |                            |                                   |                   |                   |                              |                              |
|              |                              |        |                   |       | 4       | cd0078i      | deaminases, deoxycytidylate deaminases and riboflavin deaminases.<br><sup>3</sup> Also included are the apoBec family of mRNA editing enzymes. All<br>members are Zn dependent. The zinc ion in the active site plays a central<br>role in the proposed catalytic mechanism, activating a water molecule to<br>form a hydroxide ion that performs a nucleophilic attack on the substrate.                                                                                                                                     | 62.72        | 3.26E-11 | 33%              | 52%     | 21–113  | 1392    | 4             | BAE51             | 1501 Deoxycytidylate deaminase                                                                                                                           | 108.23                     | 7.15E-23                          | 42%               | 63%               | 21141                        | 20-139                       |
|              |                              |        |                   |       |         |              | om a mysrossocion mar periornio a nuoreoprino arrack on me substrate                                                                                                                                                                                                                                                                                                                                                                                                                                                          |              |          |                  |         |         |         |               |                   |                                                                                                                                                          |                            |                                   |                   |                   |                              |                              |

| Gene<br>Name | Genome<br>Position | A.A.<br>length | Peptide<br>Mw | pl    | CDD Hit<br>Number | COGs         | COG Definition                                                                                                                                                                                                                    | Bit<br>Score | E-value               | %<br>dentity Po | %<br>ositive | Query<br>from-to | Hit<br>from-to  | BLASTp<br>Hit<br>Number | Ac     | Hit BLAST                                                                                  | Definition                                 | Bit<br>Score   | E-value                          | %<br>Identity P   |                   | Query<br>from-to       | Hit from-<br>to         |
|--------------|--------------------|----------------|---------------|-------|-------------------|--------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|-----------------------|-----------------|--------------|------------------|-----------------|-------------------------|--------|--------------------------------------------------------------------------------------------|--------------------------------------------|----------------|----------------------------------|-------------------|-------------------|------------------------|-------------------------|
|              |                    |                |               |       |                   |              | Riboflavin_deaminase-reductase, Riboflavin-specific deaminase.<br>Riboflavin biosynthesis protein RibD                                                                                                                            |              |                       |                 |              |                  |                 | Number                  |        |                                                                                            |                                            |                |                                  |                   |                   |                        |                         |
|              |                    |                |               |       |                   |              | (Diaminohydroxyphosphoribosylaminopyrimidine deaminase) catalyzes<br>the deamination of 2,5-diamino-6-ribosylamino-4(3H)-pyrimidinone                                                                                             |              |                       |                 |              |                  |                 | _                       | _      |                                                                                            |                                            |                |                                  |                   |                   |                        |                         |
|              |                    |                |               |       | 5                 | cd01284      | 5' phosphate, which is an intermediate step in the biosynthesis of<br>riboflavin. The ribG gene of Bacillus subtilis and the ribD gene of E. coli<br>are bifunctional and contain this deaminase domain and a reductase           | 59.87        | 2.74E-10              | 38%             | 55%          | 23113            | 1693            | 5                       | , C    | CAE77204 dCMP deaminase                                                                    |                                            | 105.92         | 3.55E-22                         | 40%               | 56%               | 1142                   | 5147                    |
|              |                    |                |               |       | 6                 | COG0117      | domain which catalyzes the subsequent reduction of the ribosyl side characteristics. RibD, Pyrimidine deaminase [Coenzyme metabolism].                                                                                            | 56.02        | 3.75E-09              | 32%             | 54%          | 20113            | 22100           | 6                       | S YI   | P_424377 probable deoxycytidylate deaminas                                                 | se                                         | 103.61         | 1.76E-21                         | 41%               | 57%               | 9142                   | 6140                    |
|              |                    |                |               |       |                   |              | nucleoside_deaminase, Nucleoside deaminases include adenosine,<br>guanine and cytosine deaminases. These enzymes are Zn dependent                                                                                                 |              |                       |                 |              |                  |                 |                         |        |                                                                                            |                                            |                |                                  |                   |                   |                        |                         |
|              |                    |                |               |       |                   |              | and catalyze the deamination of nucleosides. The zinc ion in the active<br>site plays a central role in the proposed catalytic mechanism, activating a<br>water molecule to form a hydroxide ion that performs a nucleophilic     |              |                       |                 |              |                  |                 |                         |        |                                                                                            |                                            |                |                                  |                   |                   |                        |                         |
|              |                    |                |               |       |                   |              | attack on the substrate. The functional enzyme is a homodimer. Cytosine                                                                                                                                                           |              |                       |                 |              |                  |                 |                         |        |                                                                                            |                                            |                |                                  |                   |                   |                        |                         |
|              |                    |                |               |       | 7                 | cd01285      | and is a member of the pyrimidine salvage pathway. Cytosine deaminase is found in bacteria and fungi but is not present in mammals; for this                                                                                      | 55.27        | 6.23E-09              | 33%             | 57%          | 28120            | 19100           | 7                       | 7 IP_0 | 01016265 hypothetical protein LOC549019                                                    |                                            | 102.83         | 3.00E-21                         | 37%               | 57%               | 2140                   | 22-162                  |
|              |                    |                |               |       |                   |              | reason, the enzyme is currently of interest for antimicrobial drug design<br>and gene therapy applications against tumors. Some members of this                                                                                   |              |                       |                 |              |                  |                 |                         |        |                                                                                            |                                            |                |                                  |                   |                   |                        |                         |
|              |                    |                |               |       |                   |              | family are tRNA-specific adenosine deaminases that generate inosine at<br>the first position of their anticodon (position 34) of specific tRNAs; this<br>modification is thought to enlarge the codon recognition capacity during |              |                       |                 |              |                  |                 |                         |        |                                                                                            |                                            |                |                                  |                   |                   |                        |                         |
|              |                    |                |               |       |                   |              | protein synthesis. Other members of the family are guanine deaminases which deaminate quanine to vanthine as part of the utilization of quanine : CumB, Cytosine/adenosine deaminases [Nucleotide transport and                   |              |                       |                 |              |                  |                 |                         |        |                                                                                            |                                            |                |                                  |                   |                   |                        |                         |
|              |                    |                |               |       | 8                 | COG0590      | CumB, Cytosine/adenosine deaminases [Nucleotide transport and metabolism / Translation. ribosomal structure and biogenesis].                                                                                                      | 51.11        | 1.20E-07              | 41%             | 54%          | 28113            | 30104           |                         |        | .00527356 Cytidine/deoxycytidylate deaminasi<br>P_781375 PREDICTED: similar to Deoxycytid  |                                            |                | 1.14E-20<br>1.95E-20             | 46%<br>38%        | 68%<br>54%        | 21113<br>1139          | 22-115<br>60-200        |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                   |              |                       |                 |              |                  |                 | 10                      | X      | P_849027 PREDICTED: similar to Deoxycytic isoform 1                                        | dylate deaminase (dCMP deaminase)          |                | 2.54E-20                         | 38%               | 54%               | 2-139                  | 11-150                  |
| M531L 2      | 109217208123       | 365            | 41,273        | 8.49  | )                 | No Hit Found |                                                                                                                                                                                                                                   |              |                       |                 |              |                  |                 | 1                       | l Ni   | P 048502 A1541                                                                             |                                            | 295.05         | 2.44E-78                         | 44%               | 59%               | 24364                  | 6-347                   |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                   |              |                       |                 |              |                  |                 | 2                       | 2 NI   | P_048920 similar to Chlorella virus PBCV-1 C<br>Accession Number U42580                    |                                            | 264.23         | 4.61E-69                         | 43%               | 56%               | 15365                  | 1351                    |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                   |              |                       |                 |              |                  |                 | 3                       |        | Swiss-Prot Accession Number P28                                                            | phosphate reductase, corresponds to<br>903 | 260.38         | 6.66E-68                         | 39%               | 57%               | 24365                  | 8357                    |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                   |              |                       |                 |              |                  |                 | 5<br>6                  | 5 E    | P_077492 EsV-1-7<br>EAA73437 hypothetical protein FG03969.1<br>P_077600 EsV-1-115          |                                            |                | 3.58E-21<br>2.49E-06<br>1.24E-05 | 37%<br>26%<br>26% | 47%<br>36%<br>34% | 19196<br>55230<br>5196 | 5190<br>316477<br>86289 |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                   |              |                       |                 |              |                  |                 |                         |        |                                                                                            |                                            |                |                                  |                   |                   |                        |                         |
| M535L 2      | 11856209244        | 871            | 95,334        | 6.52  | 2 1               |              | 4 MgtA, Cation transport ATPase [Inorganic ion transport and metabolism].                                                                                                                                                         |              | 1.49E-153<br>2.00E-62 | 30%<br>27%      | 51%<br>48%   |                  | 33898<br>152678 |                         |        | P_483341 putative calcium-transporting ATPa<br>P_636219 hypothetical protein DDB0188438    | ise 8, plasma membrane-type                |                | 7.17E-150<br>3.56E-149           | 36%<br>35%        | 55%<br>55%        | 34849                  | 1721053<br>64977        |
|              |                    |                |               |       | 3                 |              | <ul> <li>ZntA, Cation transport ATPase [Inorganic ion transport and metabolism].</li> <li>E1-E2 ATPase, E1-E2 ATPase</li> </ul>                                                                                                   |              | 4.77E-40              | 33%             | 55%          | 101334           |                 |                         |        | P_194719 ACA10; calcium-transporting ATPa                                                  | se/ calmodulin binding                     |                | 1.50E-147                        | 36%               | 55%               |                        | 1591035                 |
|              |                    |                |               |       | 4                 | COG2216      | KdpB, High-affinity K+ transport system, ATPase chain B [Inorganic ion transport and metaholism] Hydrolase, haloacid dehalogenase-like hydrolase. This family are                                                                 | 158.44       | 5.43E-40              | 28%             | 49%          | 48699            | 20609           | 4                       | , C    | CAD67616 calcium-dependent ATPase                                                          |                                            | 521.93         | 3.68E-146                        | 37%               | 55%               | 25849                  | 1401015                 |
|              |                    |                |               |       |                   |              | structurally different from the alpha/ beta hydrolase family (pfam00561).  This family includes L-2-haloacid dehalogenase, epoxide hydrolases and                                                                                 |              |                       |                 |              |                  |                 |                         |        |                                                                                            |                                            |                |                                  |                   |                   |                        |                         |
|              |                    |                |               |       | 5                 | pfam00702    | <sup>2</sup> phosphatases. The structure of the family consists of two domains. One<br>is an inserted four helix bundle, which is the least well conserved region                                                                 | 64.23        | 1.31E-11              | 20%             | 40%          | 449631           | 31197           | 5                       | 5 C    | CAD67615 putative P-type II calcium ATPase                                                 |                                            | 512.30         | 2.92E-143                        | 38%               | 56%               | 40850                  | 1531017                 |
|              |                    |                |               |       |                   |              | of the alignment. The rest of the fold is composed of the core alpha/beta domain  Cation_ATPase_C, Cation transporting ATPase, C-terminus. Members                                                                                |              |                       |                 |              |                  |                 |                         |        |                                                                                            |                                            |                |                                  |                   |                   |                        |                         |
|              |                    |                |               |       | 6                 | pfam00689    | of this families are involved in Na+/K+, H+/K+, Ca++ and Mg++<br>transport                                                                                                                                                        | 56.40        | 2.54E-09              | 21%             | 38%          | 726850           | 1147            |                         |        | P_473800 OSJNBb0015N08.12                                                                  |                                            | 510.38         | 1.11E-142                        | 35%               | 55%               | 31848                  | 1581040                 |
|              |                    |                |               |       | 7                 | pfam00690    | Cation_ATPase_N, Cation transporter/ATPase, N-terminus. Members of this families are involved in Na+/K+, H+/K+, Ca++ and Mg++ transport.                                                                                          | 47.13        | 1.67E-06              | 33%             | 52%          | 2380             | 2179            | 7                       | 7 NI   | P_851200 ACA8 (AUTOINHIBITED CA2+<br>transporting ATPase/ calmodulin bi                    | -ATPASE, ISOFORM 8); calcium-<br>inding    | 509.22         | 2.47E-142                        | 36%               | 55%               | 35849                  | 1591032                 |
|              |                    |                |               |       | 8                 |              | 7 COG4087, Soluble P-type ATPase [General function prediction only]. 1 Cof, Predicted hydrolases of the HAD superfamily [General function                                                                                         | 46.50        | 2.67E-06              | 30%             | 53%          | 530644           |                 | 8                       |        | EAL90415 P-type calcium ATPase, putative                                                   |                                            |                | 4.22E-142                        | 35%               | 54%               |                        | 2751191                 |
|              |                    |                |               |       | 10                | COG0561      | prediction onlv1. SerB, Phosphoserine phosphatase [Amino acid transport and                                                                                                                                                       | 45.87        | 3.81E-06<br>0.000187  | 32%<br>25%      | 47%<br>40%   | 594660<br>488637 | 197263<br>49195 | 10                      |        | CAB43665 Ca2+-transporting ATPase-like pro<br>AAU44048 putative P-type ATPase              | tein                                       |                | 1.23E-141<br>4.66E-141           | 35%<br>35%        | 53%<br>55%        |                        | 1591059<br>121994       |
|              |                    |                |               |       |                   |              | metabolisml.  S TKc. Serine/Threonine protein kinases, catalytic domain.                                                                                                                                                          |              |                       |                 |              |                  |                 |                         |        | 7,                                                                                         |                                            |                |                                  |                   |                   |                        |                         |
| MEASE S      | 12005212820        | 272            | 30.497        | 8.76  | 5 1               | cd00190      | Phosphotransferases of the serine or threonine-specific kinase subfamily.  The enzymatic activity of these protein kinases is controlled by                                                                                       | 152 02       | 1.28E-38              | 29%             | 51%          | 15265            | 1256            |                         | ı NI   | P 048643 similar to PBCV-1 serine/threoni                                                  | ine protein kinase, corresponds to         | 134.81         | 2.64E-30                         | 32%               | 49%               | 3265                   | 5279                    |
| 1110-1011 2  | 12000 212020       | 2,2            | 00,401        | 0.70  |                   | 0000100      | phosphorylation of specific residues in the activation segment of the<br>catalytic domain, sometimes combined with reversible conformational                                                                                      | 100.02       | 1.202.00              | 2010            | 0170         | 10 200           | . 200           |                         |        | GenBank Accession Number U146                                                              | 60                                         | 104.01         | 2.042 00                         | 0270              | 4070              | 0 200                  | 0 2.0                   |
|              |                    |                |               |       | 2                 | smart00220   | chances in the C-terminal autoreculatory tail  S_TKc, Serine/Threonine protein kinases, catalytic domain;  Phosohotransferases. Serine or threonine-specific kinase subfamily                                                     | 151.52       | 5.62E-38              | 28%             | 49%          | 16265            | 1256            | 2                       | 2 0    | CAG59101 unnamed protein product                                                           |                                            | 102.45         | 1.45E-20                         | 29%               | 50%               | 20267                  | 49-309                  |
|              |                    |                |               |       | 3                 |              | Pkinase, Protein kinase domain SPS1, Serine/threonine protein kinase [General function prediction only /                                                                                                                          | 148.51       | 5.37E-37              | 29%             | 48%          | 16265            |                 | 3                       |        | P_666308 calcium-dependent protein kinase                                                  |                                            | 94.36          | 3.95E-18                         | 30%               | 50%               | 22270                  | 191-445                 |
|              |                    |                |               |       | 4                 | COG0515      | 5 Signal transduction mechanisms / Transcription / DNA replication,<br>recombination. and reoairl.<br>TyrKc, Tyrosine kinase, catalytic domain. Phosphotransferases; tyrosine-                                                    | 108.32       | 6.85E-25              | 25%             | 44%          | 16270            | 2283            | 4                       |        | AAS47705 calcium-dependent protein kinase 1                                                | 1                                          | 94.36          | 3.95E-18                         | 30%               | 50%               | 22270                  | 191445                  |
|              |                    |                |               |       |                   |              | specific kinase subfamily. Enzymes with TyrKc domains belong to an<br>extensive family of proteins which share a conserved catalytic core                                                                                         |              |                       |                 |              |                  |                 |                         |        |                                                                                            |                                            |                |                                  |                   |                   |                        |                         |
|              |                    |                |               |       | 5                 | cd00192      | 2 common to both serine/threonine and tyrosine protein kinases. Enzymatic<br>activity of tyrosine protein kinases is controlled by phosphorylation of                                                                             | 82.54        | 3.41E-17              | 27%             | 44%          | 13205            | 6207            | 5                       | S XI   | P_504770 hypothetical protein                                                              |                                            | 94.36          | 3.95E-18                         | 28%               | 48%               | 20265                  | 103350                  |
|              |                    |                |               |       |                   |              | specific tyrosine residues in the activation segment of the catalytic<br>domain or a C-terminal tyrosine (tail) residue with reversible<br>conformational chances                                                                 |              |                       |                 |              |                  |                 |                         |        |                                                                                            |                                            |                |                                  |                   |                   |                        |                         |
|              |                    |                |               |       | 6                 | smart00219   | TyrKc, Tyrosine kinase, catalytic domain; Phosphotransferases. Tyrosine-specific kinase subfamilv.                                                                                                                                | 78.68        | 4.74E-16              | 26%             | 47%          | 17205            | 2196            | 6                       | 3 0    | CAA99015 CMK2                                                                              |                                            | 93.59          | 6.75E-18                         | 26%               | 49%               | 20265                  | 51-309                  |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                   |              |                       |                 |              |                  |                 | 7                       | 3 E    | ABA60892 calmodulin-like domain protein kina<br>EAA76933 hypothetical protein FG07121.1    |                                            | 93.59<br>93.20 | 6.75E-18<br>8.81E-18             | 26%<br>27%        | 46%<br>46%        | 15265<br>14269         | 75-333<br>285-556       |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                   |              |                       |                 |              |                  |                 | 9<br>10                 | ) (    | CAA40281 calmodulin-dependent protein kinas<br>P_667020 calmodulin-domain protein kinase 1 | se type II                                 | 92.82<br>92.82 | 1.15E-17<br>1.15E-17             | 26%<br>24%        | 49%<br>48%        | 20265<br>16270         | 51-309<br>54-315        |
| M546R 2      | 12942216115        | 1058           | 117,663       | 9.64  | 1                 | smart00433   | TOP2c, TopoisomeraseII; Eukaryotic DNA topoisomerase II, GyrB, ParE                                                                                                                                                               | 498.55       | 2.12E-142             | 33%             | 49%          | 50610            | 1589            | 1                       | I A    | AAU95770 topoisomerase II                                                                  |                                            | 1928.68        | 0.00E+00                         | 91%               | 93%               | 11058                  | 11058                   |
|              |                    |                |               |       | 2                 | smart00434   | TOP4c, DNA Topoisomerase IV; Bacterial DNA topoisomerase IV, GyrA, ParC .                                                                                                                                                         | 423.06       | 1.28E-119             | 34%             | 50%          | 6261051          | 1457            | 2                       |        | P_048939 PBCV-1 DNA topoisomerase II                                                       |                                            | 1360.90        | 0.00E+00                         | 63%               | 76%               | 31056                  | 2-1061                  |
|              |                    |                |               |       |                   |              | TOP4c, DNA Topoisomerase, subtype IIA; domain A'; bacterial<br>DNA topoisomerase IV (C subunit, ParC), bacterial DNA gyrases (A                                                                                                   |              |                       |                 |              |                  |                 |                         |        |                                                                                            |                                            |                |                                  |                   |                   |                        |                         |
|              |                    |                |               |       | 3                 | cd00187      | 7 subunit, GyrA), mammalian DNA toposiomerases II. DNA topoisomerases<br>are essential enzymes that regulate the conformational changes in DNA<br>topology by catalysing the concerted breakage and reigining of DNA              | 398.01       | 4.23E-112             | 33%             | 50%          | 6441056          | 1445            | 3                       | 3 C    | CAD25222 DNA TOPOISOMERASE II                                                              |                                            | 893.65         | 0.00E+00                         | 46%               | 62%               | 51053                  | 81066                   |
|              |                    |                |               |       | 4                 | COGO493      | strands during normal cellular growth GyrB, Type IIA topoisomerase (DNA gyrase/topo II, topoisomerase IV),                                                                                                                        | 360.60       | 8.10E-101             | 28%             | 44%          | 2626             | 8635            | 4                       |        | AW40881 DNA topoisomerase II, putative                                                     |                                            | 09.938         | 0.00E+00                         | 44%               | 59%               | 51057                  | 1121235                 |
|              |                    |                |               |       | 5                 | pfam00521    | B subunit IUNA replication recombination and repair!  1 DNA_topoisolV, DNA gyrase/topoisomerase IV, subunit A                                                                                                                     | 300.93       | 6.09E-83              | 31%             | 46%          | 6461053          | 1434            | 5                       | 5 F    | BAD86854 DNA topoisomerase II                                                              |                                            | 860.14         | 0.00E+00                         | 42%               | 59%               | 51057                  | 1031249                 |
|              |                    |                |               |       | 6                 | COG0188      | 3 GyrA, Type IIA topoisomerase (DNA gyrase/topo II, topoisomerase IV),<br>A subunit IDNA reolication. recombination. and reoairl.<br>DNA_gyraseB, DNA gyrase B. This family represents the second                                 | 267.90       | 5.37E-73              | 25%             |              | 6171051          |                 | 6                       | S NI   | P_189031 TOPII (TOPOISOMERASE II); A topoisomerase (ATP-hydrolyzing)                       | ATP binding / DNA binding / DNA            | 859.75         | 0.00E+00                         | 43%               | 60%               | 51056                  | 351166                  |
|              |                    |                |               |       | 7                 | pfam00204    | domain of DNA gyrase B. Which has a ribosomal S5 domain 2-like fold.  This family is structurally related to PF01119                                                                                                              | 72.97        | 3.02E-14              | 27%             | 45%          | 237382           | 14168           |                         |        | P_467311 putative DNA topoisomerase II                                                     |                                            |                | 0.00E+00                         | 43%               | 58%               | 51054                  |                         |
|              |                    |                |               |       |                   |              |                                                                                                                                                                                                                                   |              |                       |                 |              |                  |                 | 8                       | , Α    | BAE06274 topoisomerase II<br>AAN85208 DNA topoisomerase II                                 |                                            | 850.89         | 0.00E+00<br>0.00E+00             | 44%<br>43%        | 58%<br>59%        | 51054<br>51054         | 351166<br>271148        |
| o deaaw      | 16182216589        | 120            | 15.899        | 11.01 |                   | No Hit Found |                                                                                                                                                                                                                                   |              |                       |                 |              |                  |                 |                         |        | P 759648 hypothetical protein UM03501.1 P 048933 A577L                                     |                                            |                | 0.00E+00<br>4.19E-23             | 42%<br>65%        | 58%<br>69%        | 51057<br>26111         | 1001235<br>285          |
|              | 16631217152        |                | 19,111        | 9.91  |                   | No Hit Found |                                                                                                                                                                                                                                   |              |                       |                 |              |                  |                 |                         |        | P_048933 A577L<br>P_048732 A375R                                                           |                                            |                | 4.19E-23<br>5.18E-51             | 58%               | 70%               | 4-171                  | 6172                    |
| 2            |                    |                | ,             | 5.51  |                   |              |                                                                                                                                                                                                                                   |              |                       |                 |              |                  |                 |                         |        | /10/01                                                                                     |                                            |                |                                  | -570              | . 370             |                        |                         |

| Gene<br>Name | Genome<br>Position               | A.A.<br>length | Peptide<br>Mw | e pl   | CDD Hit<br>Number | COGs                    | COG Definition                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | Bit<br>Score | E-value I            | %<br>Identity I | %<br>Positive | Query<br>from-to | Hit o from-to            | BLASTp<br>Hit<br>Number |              | Hit BLASTp Definition                                                                                                                                                | Bit<br>Score              | E-value <sub>I</sub>             | %<br>dentity Po   |                   | Query<br>from-to        | Hit from-<br>to               |
|--------------|----------------------------------|----------------|---------------|--------|-------------------|-------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|----------------------|-----------------|---------------|------------------|--------------------------|-------------------------|--------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------|----------------------------------|-------------------|-------------------|-------------------------|-------------------------------|
| M557         | L 218777217155                   | 541            | 56,442        | 2 10.4 | 43                | 1 pfam05616             | Neisseria_TspB, Neisseria meningitidis TspB protein. This family consists of several Neisseria meningitidis TspB virulence factor proteins                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | 48.97        | 5.00E-07             | 33%             | 40%           | 1231             | 96 322-395               | 1                       |              | Po-048762 Pro-, Lys-rich, PAPK (30x); similar to wheat Pro-, Lys-rich protein, corresponds to GenBank Accession Number X52472                                        | 200.29                    | 1.39E-49                         | 37%               | 57%               | 266541                  | 214496                        |
|              |                                  |                |               |        |                   | 2 pfam05518             | Totivirus coat, Totivirus coat protein                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | 48.44        | 6.02E-07             | 24%             | 29%           | 161              | 45 630-751               | 2                       | 2 B/         | AE02830 surface protein                                                                                                                                              | 55.07                     | 7.22E-06                         | 44%               | 62%               | 201258                  | 572629                        |
| M560         | L 219451218903                   | 183            | 20,996        | 3.2    | 20                | 1 pfam04931             | DNA_pol_V, DNA polymerase V. This family includes the fifth essential<br>DNA polymerase in yeast EC:2.7.7.7. Pol5p is localised exclusively to the<br>nucleolus and binds near or at the enhancer region of rRNA-encoding<br>DNA repeating units                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 36.57        | 0.0024               | 24%             | 42%           | 551              | 54 644740                | 1                       | 1 NP         | P_048761 Glu-, Asn-rich                                                                                                                                              | 78.95                     | 7.56E-14                         | 63%               | 77%               | 157                     | 259                           |
| M563         | L 219757219494                   | 88             | 10,052        | 2 10.2 | 28                | No Hit Found            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |                      |                 |               |                  |                          | 1                       | 1 NP.        | P_048760 A403R                                                                                                                                                       | 95.13                     | 6.33E-19                         | 52%               | 69%               | 288                     | 693                           |
| M564         | L 220484219801                   | 228            | 24,869        | 9 6.8  | 88                | No Hit Found            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |                      |                 |               |                  |                          | 1                       | 1 NP         | P_048759 A402R                                                                                                                                                       | 231.88                    | 1.14E-59                         | 48%               | 72%               | 4228                    | 5227                          |
| M567         | L 221450220581                   | 290            | 32,586        | 5.4    | 48                | No Hit Found            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |                      |                 |               |                  |                          | 1 2                     | 2 B/         | P_048758 A401R<br>AC51116 bil5851                                                                                                                                    | 337.42<br>140.58          | 2.98E-91<br>5.35E-32             | 60%<br>34%        | 75%<br>53%        | 4287<br>34284           | 2274<br>7245                  |
|              |                                  |                |               |        |                   |                         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |                      |                 |               |                  |                          | 4                       | 4 YP         | 00863599 conserved hypothetical protein 7 190685 hypothetical protein GOX0246 00577151 conserved hypothetical protein                                                | 130.57<br>119.01<br>69.32 | 5.54E-29<br>1.67E-25<br>1.52E-10 | 33%<br>32%<br>24% | 53%<br>50%<br>41% | 34284<br>28287<br>23289 | 11-249<br>5-253<br>2-263      |
| M569         | L 221845221495                   | 117            | 13,086        | 6 10.0 | 03                | No Hit Found            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |                      |                 |               |                  |                          |                         | _            | 2_048757 similar to PBCV-1 ORF A214, corresponds to GenBank Accession Number U42580                                                                                  | 84.34                     | 1.12E-15                         | 37%               | 57%               | 5117                    | 4116                          |
| M570         | L 221876222109                   | 78             | 18,239        | 9.9    | 90                | 1 pfam00075             | RnaseH, RNase H. RNase H digests the RNA strand of an RNA/DNA hybrid. Important enzyme in retroviral replication cycle, and often found as a domain associated with reverse transcriptases. Structure is a mixed                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 73.40        | 1.85E-14             | 30%             | 47%           | 161              | 36 2124                  | 1                       | 1 NP.        | P_048756 A399R                                                                                                                                                       | 147.90                    | 9.70E-35                         | 51%               | 64%               | 4158                    | 17–180                        |
|              |                                  |                |               |        |                   | 2 COG0328               | aloha+beta fold with three a/b/a lavers<br>RnhA, Ribonuclease HI [DNA replication, recombination, and repair].                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | 57.29        | 1.52E-09             | 23%             | 37%           | 171              | 47 3153                  | 2                       |              | P_662792 hypothetical protein AN5188.2<br>AA41925 unnamed protein product                                                                                            |                           | 2.49E-06<br>4.25E-06             | 32%<br>32%        | 44%<br>51%        | 18136<br>15140          | 228-373<br>238-369            |
|              |                                  |                |               |        |                   |                         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |                      |                 |               |                  |                          | 4                       | 4            | AA41925 unnamed protein product<br>S37765 hypothetical protein - fruit fly (Drosophila miranda) transposon TRIM<br>-AL40632 ENSANGP00000008819                       |                           | 4.25E-06<br>4.25E-06<br>1.61E-05 | 32%<br>32%<br>28% | 51%<br>51%<br>46% | 15140<br>15140<br>12137 | 238-369<br>143-274<br>247-372 |
|              |                                  |                |               |        |                   |                         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |                      |                 |               |                  |                          | 6                       | 6            | 1RDH_B Chain B, Hiv-1 Reverse Transcriptase (Ribonuclease H Domain)                                                                                                  | 50.45                     | 2.11E-05                         | 27%               | 42%               | 9136                    | 1141                          |
|              |                                  |                |               |        |                   |                         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |                      |                 |               |                  |                          | 7                       | 8 A/         | AL40633 ENSANGP0000020605<br>AB65093 Lian-Aa1 retrotransposon protein                                                                                                | 49.68                     | 2.75E-05<br>3.59E-05             | 28%<br>31%        | 46%<br>52%        | 12137<br>15147          | 9101046                       |
|              |                                  |                |               |        |                   |                         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |                      |                 |               |                  |                          | 10                      | 9 E/         | EAL88427 ribonuclease h1<br>AE61068 unnamed protein product                                                                                                          | 49.29<br>48.52            | 4.70E-05<br>8.01E-05             | 31%<br>31%        | 46%<br>45%        | 18136<br>18136          | 239-383<br>236-380            |
| M571F        | R 222368221877                   | 164            | 9,065         | 5 8.8  | B1                | No Hit Found            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |                      |                 |               |                  |                          |                         | No H         | Hit Found No Hit Found                                                                                                                                               |                           |                                  |                   |                   |                         |                               |
|              | R 222471222824                   |                | 12,885        |        |                   | No Hit Found            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |                      |                 |               |                  |                          |                         |              | P 048755 A398L                                                                                                                                                       |                           | 9.08E-26                         | 49%               | 63%               | 1118                    |                               |
|              | L 223289222831<br>L 223684223316 |                | 17,499        |        |                   | No Hit Found            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |                      |                 |               |                  |                          |                         |              | AU06284 hypothetical protein  9_048752 A395R                                                                                                                         |                           | 2.80E-11<br>1.24E-22             | 32%<br>66%        | 49%<br>78%        | 7143<br>43122           | 23141                         |
|              |                                  |                |               |        |                   |                         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |                      |                 |               |                  |                          |                         |              | Lys- Am-rich: contains eukaryotic putative RNA-binding region RNP-1                                                                                                  |                           |                                  |                   |                   |                         |                               |
| M578F        | R 223670224293                   | 208            | 23,718        | 8.2    | 20                | No Hit Found            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |                      |                 |               |                  |                          |                         |              | P_04884 signature; similar to PBCV-1 ORF A267L, corresponds to GenBank<br>Accession Number U42580<br>P_048846 Lys., Glu-rich                                         |                           | 6.39E-40<br>5.98E-38             | 43%<br>42%        | 58%<br>59%        | 19207<br>17195          | 106294                        |
|              |                                  |                |               |        |                   |                         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |                      |                 |               |                  |                          | 3                       | 3 YP         | u-o-o-o-u tys., Giu-ncn<br>> 142777 unknown<br>>_048621 A267L                                                                                                        | 98.60                     | 1.25E-19<br>3.08E-18             | 34%<br>32%        | 49%<br>49%        | 19185<br>13207          |                               |
| M580         | L 224861224292                   | 190            | 21,463        | 3 3.8  | B1                | No Hit Found            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |                      |                 |               |                  |                          |                         |              |                                                                                                                                                                      |                           | 7.21E-18                         | 44%               | 63%               | 6-109                   |                               |
| M584F        | R 224914225219                   | 102            | 12,138        | 3 10.3 | 30                | No Hit Found            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |                      |                 |               |                  |                          |                         | No H         | Hit Found No Hit Found                                                                                                                                               |                           |                                  |                   |                   |                         |                               |
| M586         | L 226085225321                   | 255            | 29,033        | 3 9.2  | 23                | 1 smart00382            | AAA, ATPases associated with a variety of cellular activities; AAA - ATPases associated with a variety of cellular activities. This profile/alignment only detects a fraction of this vast family. The poorly conserved N-terminal helix is missing from the alignment.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | 40.89        | 0.000126             | 17%             | 37%           | 161              | 54 2146                  | 1                       | 1 NP         | P_048749 contains ATP/GTP-binding site motif A; similar to frog virus 3 ATPase, corresponds to GenBank Accession Number M80551                                       | 405.60                    | 7.17E-112                        | 78%               | 86%               | 1252                    | 1251                          |
|              |                                  |                |               |        |                   |                         | PvdE, ABC-type siderophore export system, fused ATPase and<br>permease components [Secondary metabolites biosynthesis, transport,<br>and catabolism / Inorganic ion transport and metabolism].                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | 37.62        | 0.001405             | 28%             | 40%           | 121              | 02 345448                | 2                       | 2 B/         | 8AB69884 AGB-1                                                                                                                                                       | 152.14                    | 1.43E-35                         | 35%               | 55%               | 1245                    | 1253                          |
|              |                                  |                |               |        |                   | 3 COG1123               | and calabolasm includes but inations and instabolation in termination and instabolation.  COG1123, ATPase components of various ABC-type transport systems, contain dunlicated ATPase (General function practiction only ReA-like, NTPases, ReA-like NTPases, This family includes the NTP binding domain of F1 and V1 H+ATPases. DnaB and related helicases                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | 36.70        | 0.002422             | 32%             | 52%           | 12               | 98 3197                  | 3                       | 3 AA         | AR26836 FirrV-1-A12                                                                                                                                                  | 147.90                    | 2.69E-34                         | 39%               | 57%               | 32230                   | 36-243                        |
|              |                                  |                |               |        |                   | 4 cd01120               | The state of the s | 35.09        | 0.007282             | 23%             | 40%           | 18-1             | 05 1-95                  | 4                       | 4 NP         | P_077511 EsV-1-26                                                                                                                                                    | 135.19                    | 1.81E-30                         | 35%               | 57%               | 32234                   | 47–258                        |
|              |                                  |                |               |        |                   |                         | secretion                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |              |                      |                 |               |                  |                          | 5                       | 6 A          | AAT09674 AAA-ATPase<br>AAL77796 putative ATPase                                                                                                                      | 122.87<br>122.87          | 9.28E-27<br>9.28E-27             | 33%<br>33%        | 56%<br>56%        | 32235<br>32235          | 43248<br>43248                |
|              |                                  |                |               |        |                   |                         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |                      |                 |               |                  |                          | 7                       | 7 AA<br>8 YP | AM82544 putative ATPase P 003858 ATPase                                                                                                                              | 121.71                    | 1.21E-26<br>2.07E-26             | 33%<br>33%        | 56%<br>56%        | 32235<br>32231          | 43-248<br>43-244              |
|              |                                  |                |               |        |                   |                         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |                      |                 |               |                  |                          | 10                      |              | AS18149 ATPase AV91100 ATPase-like protein                                                                                                                           | 120.55<br>120.55          | 4.61E-26<br>4.61E-26             | 31%<br>31%        | 56%<br>56%        | 32235<br>32235          | 46-251<br>46-251              |
|              |                                  |                |               |        |                   |                         | Trypan_PARP, Procyclic acidic repetitive protein (PARP). This family consists of several Trypanosoma brucei procyclic acidic repetitive protein                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |                      |                 |               |                  |                          |                         |              |                                                                                                                                                                      |                           |                                  |                   |                   |                         |                               |
| M591         | R 226168227925                   | 586            | 61,374        | 4 6.7  | 77                | 1 pfam05887             | (PARP) like sequences. The procyclic acidic repetitive protein (parp) genes of Trypanosoma brucei encode a small family of abundant surface proteins whose expression is restricted to the procyclic form of the parasite. They are found at two unlinked loci, parpA and parpB;                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 56.53        | 2.70E-09             | 31%             | 42%           | 4605             | 34 50124                 | 1                       | 1 NP         | 2_048741 Lys-, Pro-rich, PAPK (10x); similar to wheat Pro-, Lys-rich protein, corresponds to GenBank Accession Number X52472                                         | 177.56                    | 1.06E-42                         | 39%               | 50%               | 214464                  | 8263                          |
|              |                                  |                |               |        |                   | 2 pfam05616             | transcription of both loci is developmentally regulated Neisseria_TspB, Neisseria meningitidis TspB protein. This family                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 54.75        | 8.62E-09             | 39%             | 43%           | 4785             | 32 323-377               | 2                       | 2 NP         | P_048921 A565R                                                                                                                                                       | 115.93                    | 3.80E-24                         | 34%               | 41%               | 130375                  | 252466                        |
|              |                                  |                |               |        |                   |                         | TonB, Periplasmic protein TonB, links inner and outer membranes [Cell                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |              | 2.99E-07             | 32%             | 39%           |                  | 35 31126                 | 3                       |              | BAA11343 DNA binding protein                                                                                                                                         |                           | 3.22E-23                         | 34%               | 41%               |                         | 252458                        |
|              |                                  |                |               |        |                   |                         | enveloce biocenesis, outer membranel.  DEC-1_N, DEC-1 protein, N terminal region. The defective chorion-1 gene (dec-1) in Drosophila encodes follicle cell proteins necessary for                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |              |                      |                 |               |                  |                          |                         |              |                                                                                                                                                                      |                           |                                  |                   |                   |                         |                               |
|              |                                  |                |               |        |                   |                         | proper eggshell assembly. Multiple products of the dec-1 gene are formed by alternative RNA splicing and proteolytic processing. Cleavage products include S80 (80 kDa) which is incorporated into the eggshell, and further proteolysis of S80 pixes S80 (80 kDa).                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 46.03        | 4.01E-06             | 33%             | 34%           | 4685             | 32 78142                 | 4                       | 4 NP         | P_048688 a332L                                                                                                                                                       | 73.94                     | 1.66E-11                         | 61%               | 80%               | 544585                  | 243                           |
|              |                                  |                |               |        |                   |                         | and further proteolysis of SR0 gives SR0 (R0 kDa)  DedD, Uncharacterized protein conserved in bacteria [Function unknown].                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |              | 6.90E-06             | 35%             | 42%           |                  | 35 79150                 |                         |              | P_473058 hypothetical protein PFB0650w                                                                                                                               |                           | 1.40E-10                         | 57%               | 76%               | 959                     |                               |
|              |                                  |                |               |        |                   | 6 pfam05518 7 pfam06735 | Totivirus coat, Totivirus coat protein  DUF1210, Protein of unknown function (DUF1210). This family                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |              | 7.71E-06<br>7.32E-05 | 25%<br>42%      | 29%<br>44%    |                  | 35 680-753<br>32 136-202 | 6<br>7                  | 6 BA<br>7 NP | AC69973 putative serine/threonine protein kinase<br>- 048917 similar to Chlorella virus CVK2 DNA binding protein, corresponds to<br>GenRank Accession Number 1778305 |                           | 2.03E-09<br>2.24E-08             | 67%<br>39%        | 80%<br>51%        |                         | 557602<br>306403              |
|              |                                  |                |               |        |                   | B COG5373               | COG5373. Predicted membrane protein (Function unknown).                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |              | 0.000154             | 31%             |               |                  | 32 49111                 |                         | 8 BA         | GenBank Accession Number D78305<br>8AA11342 DNA binding protein                                                                                                      |                           | 2.48E-07                         | 38%               | 50%               |                         | 304401                        |
|              |                                  |                |               |        |                   | 9 COG3898               | COG3898, Uncharacterized membrane-bound protein [Function unknown]. MSP1_C, Merozoite surface protein 1 (MSP1) C-terminus. This family represents the C-terminal region of merozoite surface protein 1 (MSP1) which are found in a number of Plasmodium species. MSP-1 is a 200-tb2 protein expressed on the surface of the P. vivax merozoite. MSP-1                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 40.43        | 0.00016              | 24%             | 34%           | 4615             | 44 439-525               | g                       | 9            | P12348 Period circadían protein                                                                                                                                      | 59.69                     | 3.23E-07                         | 52%               | 58%               | 959                     | 744794                        |
|              |                                  |                |               |        | 1                 | 0 pfam07462             | of Plasmodium species is synthesised as a high-molecular-weight precursor and then processed into several fragments. At the time of red                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | 40.05        | 0.000259             | 23%             | 27%           | 4785             | 31 259-311               | 10                      | 0 XP         | P_500901 hypothetical protein                                                                                                                                        | 57.00                     | 2.10E-06                         | 37%               | 62%               | 1360                    | 327-374                       |
|              |                                  |                |               |        |                   |                         | cell invasion by the mercozite, only the 19-kDa C-terminal fragment (MSP-119), withor contains two epidemal growth factor-like domains, remains on the surface. Antibodies against MSP-119 inhibit mercozite entry into red cells, and immunisation with MSP-119 protects monkeys from challenging infections. Hence, MSP-119 is considered a promising vaccine candidate.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |              |                      |                 |               |                  |                          |                         | •            |                                                                                                                                                                      |                           |                                  |                   |                   |                         |                               |
| M599F        | R 228277229137                   | 287            | 30,575        | 5 6.5  | 51                | 1 COG3889               | COG3889, Predicted solute binding protein [General function prediction only].                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | 38.10        | 0.00079              | 25%             | 40%           | 1532             | 28 764841                |                         | No H         | Hit Found No Hit Found                                                                                                                                               |                           |                                  |                   |                   |                         |                               |

| Gene<br>Name | Genome<br>Position           | A.A.<br>length | Peptide<br>Mw | pl           | CDD Hit<br>Number | COGs                      | COG Definition                                                                                                                                                                                                                                                          | Bit<br>Score   | E-value  | %<br>Identity I | %<br>Positive | Query<br>from-to | Hit<br>from-to   | BLASTp<br>Hit<br>Number | Hit<br>Accession       | n BLASTp Definition                                                                                          | Bit<br>Score     | E-value               | %<br>Identity P | %<br>ositive | Query<br>from-to | Hit from-<br>to  |
|--------------|------------------------------|----------------|---------------|--------------|-------------------|---------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------|----------|-----------------|---------------|------------------|------------------|-------------------------|------------------------|--------------------------------------------------------------------------------------------------------------|------------------|-----------------------|-----------------|--------------|------------------|------------------|
| M601L        | 230203229127                 | 359            | 40,298        | 7.39         | 1                 | COG0076                   | GadB, Glutamate decarboxylase and related PLP-dependent proteins<br>[Amino acid transport and metaholism]                                                                                                                                                               | 147.45         | 1.02E-36 | 24%             | 46%           | 55310            | 102-380          | 1                       | NP_04895               | similar to tomato histidine decarboxylase, corresponds to Swiss-Prot<br>Accession Number P54772              | 464.54           | 2.27E-129             | 61%             | 78%          | 4-353            | 11-362           |
|              |                              |                |               |              | 2                 | pfam00282                 | Pyridoxal_deC, Pyridoxal-dependent decarboxylase conserved domain                                                                                                                                                                                                       | 118.87         | 3.90E-28 | 27%             | 45%           | 56289            | 81352            | 2                       | P2857                  | 8 Histidine decarboxylase (HDC)                                                                              | 216.08           | 1.40E-54              | 35%             | 55%          | 14347            | 26-369           |
|              |                              |                |               |              | 3                 |                           | CsdB, Selenocysteine lyase [Amino acid transport and metabolism].<br>NifS, Cysteine sulfinate desulfinase/cysteine desulfurase and related                                                                                                                              | 48.40          | 6.36E-07 | 25%             | 41%           |                  | 3 75339          | 3                       |                        | 1 histidine decarboxylase                                                                                    | 216.08           | 1.40E-54              | 35%             | 55%          | 14347            | 26-369           |
|              |                              |                |               |              | 4                 |                           | enzymes [Amino acid transport and metaholism]                                                                                                                                                                                                                           | 48.63          | 6.49E-07 | 24%             | 44%           |                  | 7 36298          | 4                       | ZP_0051052             | 9 Pyridoxal-dependent decarboxylase<br>6 COG0076: Glutamate decarboxylase and related PLP-dependent          |                  | 9.10E-54              | 35%             | 57%          | 14347            | 47-390           |
|              |                              |                |               |              | 5                 |                           | Aminotran_5, Aminotransferase class-V  GLY1, Threonine aldolase [Amino acid transport and metabolism].                                                                                                                                                                  | 36.83<br>34.84 | 0.002041 | 24%<br>19%      | 39%<br>35%    |                  | 55206<br>3 38206 | 5                       |                        | 6 proteins<br>5 histidine decarboxylase                                                                      | 207.22           | 6.52E-52<br>2.48E-51  | 36%<br>33%      | 55%<br>55%   | 15349<br>14347   | 25-365<br>26-369 |
|              |                              |                |               |              |                   |                           |                                                                                                                                                                                                                                                                         |                |          |                 |               |                  |                  | 7                       | AAR1253                | l3 histidine decarboxylase<br>l4 Histidine decarboxylase (HDC)                                               | 205.30<br>202.60 | 2.48E-51<br>1.61E-50  | 33%<br>32%      | 55%<br>53%   | 14347<br>14355   | 40-383<br>26-377 |
|              |                              |                |               |              |                   |                           |                                                                                                                                                                                                                                                                         |                |          |                 |               |                  |                  | 9                       |                        | putative pyridoxal 5' phosphate-dependent histidine decarboxylase                                            |                  | 3.03E-49              | 32%             | 54%          | 14347            | 26-369           |
|              |                              |                |               |              |                   |                           |                                                                                                                                                                                                                                                                         |                |          |                 |               |                  |                  | 10                      | BAA7833                | 11 serine decarboxylase                                                                                      | 198.36           | 3.03E-49              | 32%             | 54%          | 11347            | 112458           |
| M605L        | 230627230226                 | 134            | 15,194        | 4.86         |                   | No Hit Found              |                                                                                                                                                                                                                                                                         |                |          |                 |               |                  |                  | 1                       | NP_04895               | 8 A602L                                                                                                      | 54.68            | 9.47E-07              | 47%             | 76%          | 853              | 68-113           |
| M606R        | 230722231036                 | 105            | 11,996        | 6.00         |                   | No Hit Found              |                                                                                                                                                                                                                                                                         |                |          |                 |               |                  |                  | 1                       | NP 048959              | 9 A603R                                                                                                      | 96.67            | 2.20E-19              | 43%             | 63%          | 1103             | 1104             |
| M607L        | 231719231219                 | 167            | 18,477        | 11.04        |                   | No Hit Found              |                                                                                                                                                                                                                                                                         |                |          |                 |               |                  |                  |                         |                        | d No Hit Found                                                                                               |                  |                       |                 |              |                  |                  |
| M610L        | 232680232162                 | 173            | 3 20,120      | 6.04         |                   | No Hit Found              |                                                                                                                                                                                                                                                                         |                |          |                 |               |                  |                  | 1 2                     | NP_04884               |                                                                                                              |                  | 4.05E-24<br>5.31E-16  | 37%<br>39%      | 59%<br>57%   | 7173<br>28165    | 18189<br>29174   |
| M612R        | 232763233833                 | 357            | 42,324        | 7.99         |                   | No Hit Found              |                                                                                                                                                                                                                                                                         |                |          |                 |               |                  |                  | 1 2                     | NP_048850<br>YP 142783 | i0 A494R<br>i3 PBCV1-A494R-like protein                                                                      |                  | 4.70E-111<br>7.18E-35 | 55%<br>34%      | 71%<br>54%   | 7357<br>92357    | 4360<br>166431   |
|              |                              |                |               |              |                   |                           |                                                                                                                                                                                                                                                                         |                |          |                 |               |                  |                  | 3                       | NP_077586              | 6 EsV-1-101<br>5 FirrV-1-B20                                                                                 | 129.41<br>122.87 | 1.71E-28<br>1.60E-26  | 36%<br>32%      | 54%<br>51%   | 166356<br>148356 | 121318<br>86300  |
|              |                              |                |               |              |                   |                           |                                                                                                                                                                                                                                                                         |                |          |                 |               |                  |                  | 5<br>6                  | YP_29416<br>NP 14974   | 1 hypothetical protein EhV 403                                                                               | 112.85<br>70.86  | 1.66E-23<br>7.23E-11  | 33%<br>22%      | 52%<br>43%   | 152353<br>96354  | 156362<br>101380 |
|              |                              |                |               |              |                   |                           |                                                                                                                                                                                                                                                                         |                |          |                 |               |                  |                  | 7                       | YP_14278               | 4 unknown                                                                                                    | 57.38            | 8.27E-07              | 27%             | 46%          | 236351           | 131-247          |
| M615R        | 233914234339                 | 142            | 14,986        | 10.14        |                   | No Hit Found              |                                                                                                                                                                                                                                                                         |                |          |                 |               |                  |                  |                         | No Hit Found           | d No Hit Found                                                                                               |                  |                       |                 |              |                  |                  |
|              |                              |                |               |              |                   |                           | Trypan_PARP, Procyclic acidic repetitive protein (PARP). This family consists of several Trypanosoma brucei procyclic acidic repetitive protein                                                                                                                         |                |          |                 |               |                  |                  |                         |                        |                                                                                                              |                  |                       |                 |              |                  |                  |
| M616L        | 235330234347                 | 328            | 35,318        | 4.75         | 1                 | pfam05887                 | (PARP) like sequences. The procyclic acidic repetitive protein (parp) genes of Trypanosoma brucei encode a small family of abundant surface proteins whose expression is restricted to the procyclic form of the                                                        | 54.60          | 1.08E-08 | 40%             | 55%           | 129196           | 60127            | 1                       | NP_04885               | 77 A501L                                                                                                     | 63.54            | 1.02E-08              | 48%             | 67%          | 473              | 566              |
|              |                              |                |               |              | 2                 | pfam05616                 | parasite. They are found at two unlinked loci, parpA and parpB;<br>transcription of both loci is developmentally regulated<br>Neisseria_TspB, Neisseria meningitidis TspB protein. This family                                                                          | 51.29          | 9.36E-08 | 32%             | 38%           | 121193           | 3 323-397        | 2                       | NP 04885               | i6 Pro-, Ser-rich                                                                                            | 60.46            | 8.61E-08              | 50%             | 67%          | 255303           | 464              |
|              |                              |                |               |              |                   |                           | consists of several Neisseria meningitidis TspB virulence factor proteins                                                                                                                                                                                               |                |          |                 |               |                  |                  |                         |                        |                                                                                                              |                  |                       |                 |              |                  |                  |
|              | 235642235418<br>236651235692 | 75<br>320      |               | 4.28<br>5.24 |                   | No Hit Found No Hit Found |                                                                                                                                                                                                                                                                         |                |          |                 |               |                  |                  | 1                       | NP 04884               | 7 A491R<br>4 Pro-,Gin-rich                                                                                   | 63.16<br>358.61  | 2.68E-09<br>1.47E-97  | 38%<br>60%      | 56%<br>72%   | 475<br>1320      | 576<br>1317      |
| Mb2UL .      | 230051235092                 | 320            | 34,621        | 5.24         |                   | No Hit Found              |                                                                                                                                                                                                                                                                         |                |          |                 |               |                  |                  |                         | YP_14277               | 1 unknown                                                                                                    |                  | 3.14E-07              | 25%             | 39%          | 8-256            | 5-245            |
| M622L        | 237725236679                 | 349            | 40,425        | 9.69         | 1                 | pfam01844                 | HNH, HNH endonuclease                                                                                                                                                                                                                                                   | 38.39          | 0.000787 | 16%             | 24%           | 237286           | 1-51             | 1                       |                        | g similar to Bacteriophage SP01 gene 31 intron, corresponds to Swiss-Prot<br>Accession Number P34081         |                  | 5.69E-21              | 30%             | 50%          | 64316            | 59-313           |
|              |                              |                |               |              |                   |                           |                                                                                                                                                                                                                                                                         |                |          |                 |               |                  |                  | 3                       | YP 142599              | 2 putative endonuclease<br>9 HNH endonuclease                                                                | 86.27<br>83.96   | 1.60E-15<br>7.96E-15  | 37%<br>29%      | 50%<br>41%   | 193346<br>16283  | 15-164<br>10-267 |
|              |                              |                |               |              |                   |                           |                                                                                                                                                                                                                                                                         |                |          |                 |               |                  |                  | 5                       | YP 14260               | 5 putative HNH homing endonuclease<br>11 HNH endonuclease                                                    | 79.72<br>73.56   | 1.50E-13<br>1.08E-11  | 33%<br>36%      | 50%<br>49%   | 197346<br>190346 | 20179<br>18173   |
|              |                              |                |               |              |                   |                           |                                                                                                                                                                                                                                                                         |                |          |                 |               |                  |                  | 7                       | YP 164778              | 7 putative endodeoxyribonuclease<br>8 orf143                                                                 | 70.48<br>68.55   | 9.11E-11<br>3.46E-10  | 35%<br>33%      | 50%<br>46%   | 215346<br>183346 | 72–203<br>4–173  |
|              |                              |                |               |              |                   |                           |                                                                                                                                                                                                                                                                         |                |          |                 |               |                  |                  | 8<br>9                  | NP_047162              | 7 prophage LambdaBa03, HNH endonuclease family protein<br>2 putative HNH homing endonuclease                 | 67.01<br>66.63   | 1.01E-09<br>1.32E-09  | 33%<br>32%      | 51%<br>47%   | 180311<br>190348 | 9146<br>10165    |
|              |                              | _              |               |              |                   |                           |                                                                                                                                                                                                                                                                         |                |          |                 |               |                  |                  | 10                      |                        | 6 putative HNH endonuclease                                                                                  | 65.08            | 3.83E-09              | 30%             | 44%          | 185346           | 5184             |
| M625R        | 237774238019                 | 82             | 9,216         | 6.80         |                   | No Hit Found              |                                                                                                                                                                                                                                                                         |                |          |                 |               |                  |                  |                         | No Hit Founi           | d No Hit Found                                                                                               |                  |                       |                 |              |                  |                  |
| M627L        | 238682238251                 | 144            | 16,779        | 10.18        | 1                 | pfam03013                 | Pyr_excise, Pyrimidine dimer DNA glycosylase. Pyrimidine dimer DNA glycosylases excise pyrimidine dimers by hydrolysis of the glycosylic bond of the 5' pyrimidine, followed by the intra-pyrimidine phosphodiester bond. Pyrimidine dimers are the major UV-lesions of | 160.21         | 1.43E-40 | 47%             | 60%           | 1134             | 1-131            | 1                       | AAD3337                | 4 pyrimidine dimer-specific glycosylase                                                                      | 163.70           | 1.43E-39              | 59%             | 69%          | 1134             | 1133             |
|              |                              |                |               |              |                   |                           | DNA                                                                                                                                                                                                                                                                     |                |          |                 |               |                  |                  | 2                       |                        | 8 PBCV-1 pyrimidine dimer-specific glycosylase                                                               | 163.31           | 1.87E-39              | 59%             | 69%          | 1-134            | 1133             |
|              |                              |                |               |              |                   |                           |                                                                                                                                                                                                                                                                         |                |          |                 |               |                  |                  | 3                       |                        | 9 pyrimidine dimer-specific qlycosylase<br>9 pyrimidine dimer-specific qlycosylase                           | 162.93<br>162.16 | 2.45E-39<br>4.17E-39  | 58%<br>58%      | 70%<br>69%   | 1-134<br>1-134   | 1133<br>1133     |
|              |                              |                |               |              |                   |                           |                                                                                                                                                                                                                                                                         |                |          |                 |               |                  |                  | 5<br>6                  | AAD3339<br>AAD3338     | 11 pyrimidine dimer-specific glycosylase<br>11 pyrimidine dimer-specific glycosylase                         | 161.77<br>161.77 | 5.45E-39<br>5.45E-39  | 58%<br>58%      | 71%<br>71%   | 1-134<br>1-134   | 1133<br>1133     |
|              |                              |                |               |              |                   |                           |                                                                                                                                                                                                                                                                         |                |          |                 |               |                  |                  | 7                       | AAD3338                | 0 pyrimidine dimer-specific glycosylase<br>7 pyrimidine dimer-specific glycosylase                           | 161.77<br>161.38 | 5.45E-39<br>7.11E-39  | 58%<br>58%      | 71%<br>70%   | 1-134<br>1-134   | 1133<br>1133     |
|              |                              |                |               |              |                   |                           |                                                                                                                                                                                                                                                                         |                |          |                 |               |                  |                  | 9                       | AAD3335                | 2 pyrimidine dimer-specific glycosylase<br>5 pyrimidine dimer-specific glycosylase                           | 161.38<br>161.00 | 7.11E-39<br>9.29E-39  | 58%<br>58%      | 70%<br>71%   | 1-134<br>1-134   | 1133<br>1133     |
| M628L        | 239312239040                 | 91             | 9,598         | 10.67        |                   | No Hit Found              |                                                                                                                                                                                                                                                                         |                |          |                 |               |                  |                  | 1                       | NP_04883               |                                                                                                              |                  | 3.11E-18              | 59%             | 81%          | 368              | 469              |
| M629L        | 240038239343                 | 232            | 26,198        | 4.58         | 1                 | COG5219                   | COG5219, Uncharacterized conserved protein, contains RING Zn-finger [General function prediction only].                                                                                                                                                                 | 39.74          | 0.000253 | 28%             | 41%           | 149198           | 3 1711525        | 1                       | NP_04883               | 7 A481L                                                                                                      | 208.38           | 1.40E-52              | 44%             | 65%          | 1-228            | 1224             |
| M631L        | 241199240105                 | 365            | 41,836        | 10.81        |                   | No Hit Found              |                                                                                                                                                                                                                                                                         |                |          |                 |               |                  |                  |                         |                        | 1 similar to Chlamydia histone-like protein, corresponds to GenBank<br>Accession Number D71563               |                  | 9.68E-43              | 39%             | 59%          | 1229             | 46-296           |
|              |                              |                |               |              |                   |                           |                                                                                                                                                                                                                                                                         |                |          |                 |               |                  |                  | 2                       | NP_048439              | 9 a91L<br>2 similar to Chlorella virus PBCV-1 ORF A282L, corresponds to GenBank<br>Accession Number I I42580 | 120.17<br>60.46  | 1.08E-25<br>1.01E-07  | 49%<br>56%      | 64%<br>70%   | 241364<br>361    | 1126<br>260      |
|              |                              |                |               |              |                   |                           |                                                                                                                                                                                                                                                                         |                |          |                 |               |                  |                  |                         | NP 04863               | similar to PBCV-1 ORF A34R, corresponds to GenBank Accession                                                 |                  | 1.12E-06              | 72%             | 82%          | 341              | 530-569          |
|              |                              |                |               |              |                   |                           |                                                                                                                                                                                                                                                                         |                |          |                 |               |                  |                  |                         |                        | Number U17055<br>32 similar to bovine cylicin I, corresponds to Swiss-Prot Accession Number<br>P35662        |                  | 1.12E-06              | 72%             | 82%          |                  | 571610           |
| Wesve        | 241316241588                 | 9.             | 10,827        | 4.50         |                   | No Hit Found              |                                                                                                                                                                                                                                                                         |                |          |                 |               |                  |                  |                         | NP_048809              |                                                                                                              | 40 14            | 8.77E-05              | 37%             | 56%          | 11-91            | 379              |
|              | 241632242276                 |                | 24,419        |              |                   | No Hit Found              |                                                                                                                                                                                                                                                                         |                |          |                 |               |                  |                  | 1                       | NP 04883               | 8 A482R                                                                                                      |                  | 4.25E-50              | 44%             | 64%          | 3-214            | 1213             |
|              | 2-1210                       | -10            | ,0            | 5.20         |                   |                           |                                                                                                                                                                                                                                                                         |                |          |                 |               |                  |                  | 2                       | AAR2697                | 10 MacAr<br>10 FirrV-1-I5<br>16 hypothetical protein EhV 438                                                 | 64.70            | 2.14E-09<br>4.76E-09  | 32%<br>34%      | 44%<br>50%   | 33130<br>33130   | 31–146<br>73–168 |
|              |                              |                |               |              |                   |                           |                                                                                                                                                                                                                                                                         |                |          |                 |               |                  |                  | 4                       | AAR2688                | 9 FirrV-1-B14<br>11 EsV-1-96                                                                                 | 63.16            | 6.21E-09<br>1.69E-06  | 31%<br>31%      | 46%<br>46%   | 25130<br>37130   | 20-145<br>70-181 |
| M637L        | 242741242277                 | 155            | 18,317        | 6.80         |                   | No Hit Found              |                                                                                                                                                                                                                                                                         |                |          |                 |               |                  |                  | 1                       | NP_04884               |                                                                                                              |                  | 8.81E-45              | 54%             | 75%          | 1154             | 1154             |
|              |                              |                |               |              |                   |                           | PAF acetylesterase like. PAF acetylhydrolase (PAF-AH) like subfamily                                                                                                                                                                                                    |                |          |                 |               |                  |                  |                         |                        |                                                                                                              |                  |                       |                 |              |                  |                  |
|              |                              |                |               |              |                   |                           | of SGNH-hydrolases. Platelet-activating factor (PAF) and PAF-AH are                                                                                                                                                                                                     |                |          |                 |               |                  |                  |                         |                        |                                                                                                              |                  |                       |                 |              |                  |                  |
| M638R        | 242798244501                 | 568            | 61,762        | 10.39        | 1                 | cd01820                   | independent phospholipase A2 which exhibits strong substrate specificity towards PAF, hydrolyzing an acetyl ester at the sn-2 position. PAF-AH                                                                                                                          | 69.51          | 3.01E-13 | 26%             | 42%           | 254412           | 35210            | 1                       | NP_04848               | 8 PBCV-1 surface protein                                                                                     | 86.27            | 3.10E-15              | 26%             | 41%          | 321566           | 659890           |
|              |                              |                |               |              |                   |                           | also degrades a family of oxidized PAF-like phospholipids with short sn-2 residues. In addition, PAF and PAF-AH are associated with neural                                                                                                                              |                |          |                 |               |                  |                  |                         |                        |                                                                                                              |                  |                       |                 |              |                  |                  |
|              |                              |                |               |              |                   |                           | migration and mammalian reproduction                                                                                                                                                                                                                                    |                |          |                 |               |                  |                  |                         |                        |                                                                                                              |                  |                       |                 |              |                  |                  |

| Gene<br>Name | Genome<br>Position | A.A.<br>length | Peptide<br>Mw | pl    | CDD Hit<br>Number | COGs                 | COG Definition                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | Bit<br>Score | E-value              | %<br>Identity F | %<br>Positive | Query<br>from-to | Hit<br>from-to | BLASTp<br>Hit<br>Number | Hit<br>Access                        |                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | Bit<br>Score                         | E-value                                      | %<br>Identity Po                |                                        | Query<br>from-to                                   | Hit from-<br>to                               |
|--------------|--------------------|----------------|---------------|-------|-------------------|----------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|----------------------|-----------------|---------------|------------------|----------------|-------------------------|--------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------|----------------------------------------------|---------------------------------|----------------------------------------|----------------------------------------------------|-----------------------------------------------|
|              |                    |                |               |       | 2                 | cd01828              | sialate O-acetylosterase. Iše2, sialate O-acetylosterase. Išes subfamily of the SGNH-hydrolases, a diverse family of lipases and esterases. The tertiary fold of the enzyme is substantially different from that of the alpha/beta hydrolase family and unique among all known hydrolases; its active site closely resembles the Ser-His-Asp(Glu) triad found in other                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | 60.68        | 1.37E-10             | 24%             | 41%           | 254393           | 2155           |                         | 2 BAD2                               | 22850 surface protein                                                                                                                                                                                                                                                                                                                                                                                                                                                     | 83.96                                | 1.54E-14                                     | 33%                             | 48%                                    | 415566                                             | 753-890                                       |
|              |                    |                |               |       | 3                 | cd01833              | serion hurdralease.<br>XynB, Bike, SGNH- hydrolase subfamily, similar to Ruminococcus flavefaciens XynB. Most likely a secreted hydrolase with xylanase activity. SGNH hydrolases are all verses family of lipsoses and esterases. The teriary fold of the enzyme is substantially different from that of the alphabeth sylvolase family and unique among all known hydrolases; its active site dosely resembles the Ser-His-Asp(Glu) thad found in other serion hurdralease.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | 48.02        | 9.22E-07             | 20%             | 30%           | 254399           | 3190           | 3                       | 3 BAD1                               | 12236 surface protein                                                                                                                                                                                                                                                                                                                                                                                                                                                     | 81.65                                | 7.63E-14                                     | 30%                             | 44%                                    | 396566                                             | 734890                                        |
|              |                    |                |               |       | 4                 | cd00229              | SGNH-hydrolase, SGNH-hydrolase, or GDSL-hydrolase, is a diverse family of lipases and esterases. The tertiary fold of the enzyme is substantially different from that of the alpha/beta hydrolase family and unique among all known hydrolases; its active site closely resembles the typical Sex-His-Asp(Su) triad from other serine hydrolases, but may lack Nnac Jike, Nhac C(MP-NeuVAc synthetase). Jike subfamily of                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | 43.29        | 2.12E-05             | 18%             | 30%           | 253400           | 1192           | 4                       | \$ T1                                | 17636 proline-rich protein A145R - Chlorella virus PBCV-1                                                                                                                                                                                                                                                                                                                                                                                                                 | 74.71                                | 9.33E-12                                     | 37%                             | 50%                                    | 452566                                             | 299                                           |
|              |                    |                |               |       | 5                 | cd01841              | SGNH_hydrolases, a diverse family of lipases and esterases. The tetrialy did fine enzyme is substantially different from that of the alphabetal hydrolase family and unique among all known hydrolases: its cuties tile closely resembles two of the three components of hydrolases. Each discharge family and unique among all known hydrolases. Each size of the components of hydrolases. Each size of the components of hydrolases. Each size of the components of the com | 43.33        | 2.57E-05             | 25%             | 43%           | 253400           | 1160           | 5                       | 5 BAE0                               | 02830 surface protein                                                                                                                                                                                                                                                                                                                                                                                                                                                     | 70.86                                | 1.35E-10                                     | 30%                             | 48%                                    | 418564                                             | 504629                                        |
|              |                    |                |               |       | 6                 | cd01838              | Isoamyl, acetate, hydrolase, like, Isoamyl-acetate hydrolyzing esterase-like proteins. SGNH- hydrolase sublamyl similar to the Saccharomyces corevisiae IAH1. IAH1 may be the major esterase that hydrolyses isoamyl acetate in sake mash. The SGNH-family of hydrolases is a diverse family of Ipases and esterases. The tertiary loci of the enzyme is subclamially different from that of the shall be shall be eliquidise family and SSNH-family of hydrolases. STR establish that is subclamially different from that of the shall be sh | 40.24        | 0.000193             | 18%             | 34%           | 254408           | 2198           | 6                       | 6 AAH5                               | 59352 MGC69176 protein                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 54.30                                | 1.30E-05                                     | 27%                             | 44%                                    | 254414                                             | 42–227                                        |
| M647L 2      | 245060244572       | 163            | 18,896        | 3.57  | 1                 | COG5406              | COG5406, Nucleosome binding factor SPN, SPT16 subunit [Transcription / DNA replication, recombination, and repair / Chromatin structure and dvnamics].                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | 40.08        | 0.000241             | 29%             | 55%           | 96163            | 929995         |                         |                                      | Found No Hit Found                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                                      |                                              |                                 |                                        |                                                    |                                               |
| M649R :      | 245243246649       | 469            | 53,138        | 10.92 | 2                 | No Hit Found         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |                      |                 |               |                  |                |                         | ND OA                                | similar to Chlamydia histone-like protein, corresponds to G<br>Accession Number D71563<br>48439 a91L<br>48636 similar to PBCV-1 ORF A34R, corresponds to GenBank Ac                                                                                                                                                                                                                                                                                                       | 126 22                               | 2.12E-27                                     | 31%<br>52%                      | 49%<br>68%                             | 23331<br>343469                                    | 31–296<br>1–126                               |
|              |                    |                |               |       |                   |                      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |                      |                 |               |                  |                | 3                       |                                      | Number U17055<br>similar to bovine cylicin I, corresponds to Swiss-Prot Accession I<br>P35662                                                                                                                                                                                                                                                                                                                                                                             |                                      | 7.08E-07<br>2.97E-05                         | 36%<br>45%                      | 42%<br>52%                             | 7122<br>760                                        | 417-542<br>516-590                            |
| M651R :      | 246879247628       | 250            | 28,263        | 8.34  | •                 | No Hit Found         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |                      |                 |               |                  |                | 2                       | 2 NP_04                              | 48357 A9R 48626 similar to PBCV-1 ORF A79R, corresponds to GenBank Ac AUMDBET U17055 48607 similar to PBCV-1 ORF A275R, encoded by GenBank Accession i                                                                                                                                                                                                                                                                                                                    | 88 97                                | 1.29E-17                                     | 67%<br>28%<br>27%<br>28%        | 84%<br>48%<br>47%<br>47%               | 79249<br>3247<br>5226<br>5248                      | 3173<br>5248<br>4217<br>4246                  |
|              |                    |                |               |       |                   |                      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |                      |                 |               |                  |                | 5                       | 5 NP_04<br>5 NP_04                   | 1142/58(1)<br>49005<br>Accession Number U42580<br>48525<br>A177R                                                                                                                                                                                                                                                                                                                                                                                                          | enBank 84.34<br>83.96                | 3.52E-15<br>4.59E-15                         | 29%<br>27%                      | 46%<br>49%                             | 5-248<br>3-233                                     | 10-253<br>5-227                               |
|              |                    |                |               |       |                   |                      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |                      |                 |               |                  |                | 7<br>8<br>9<br>10       | B AAU0<br>B NP 04                    | 06304 hypothetical protein A275R<br>06301 hypothetical protein A275R<br>84723 A36E.<br>06302 hypothetical protein A275R                                                                                                                                                                                                                                                                                                                                                   | 62.77<br>61.62<br>55.07<br>50.83     | 2.44E-08<br>2.28E-06                         | 28%<br>27%<br>26%<br>28%        | 49%<br>49%<br>47%<br>50%               | 83-247<br>83-247<br>3-144<br>116-248               | 3167<br>3167<br>10155<br>1132                 |
| M653L :      | 248608247631       | 326            | 37,173        | 4.45  | 5 1               | cd01049              | RNRD2, Ribonudeside Reductase, RObeta subunit (RNRD2) is a member of a broad superfamily of ferritin-like disron-carboyale proteins, and the production of t | 295.61       | 3.04E-81             | 49%             | 64%           | 21293            | 4-284          | 1                       | 1 NP_04                              | 48832 contains ribonucleotide reductase (RR) signature; similar to toba<br>small subunit, corresponds to Swiss-Prot Accession Number P49                                                                                                                                                                                                                                                                                                                                  | cco RR 507.29<br>730                 | 2.63E-142                                    | 75%                             | 84%                                    | 5326                                               | 3324                                          |
|              |                    |                |               |       |                   |                      | raulica anti a dilinfuji fudiser. Inite sessinital yiusyi raulica is pilopuser<br>to generate a thiyl radical, located on a cysteline residue in the RT active<br>site that initiates ribonucleotide reduction. The beta subunit is composed<br>of 10-13 helices, the 8 longest helices form an alpha-helical bundle;<br>some have 2 addition beta strands. Yeast is unique in that it assembles<br>with bomodimers and helaroritimers of ISMIPS? The waset helaroritimer                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |              |                      |                 |               |                  |                |                         |                                      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |                                      |                                              |                                 |                                        |                                                    |                                               |
|              |                    |                |               |       | 3                 | pfam00268<br>COG0208 | Ribonuc red sm. Ribonucleotide reductase, small chain.<br>NrdF. Ribonucleotide reductase, beta subunit [Nucleotide transport and metabolism].<br>Ferritin Jike, Ferritin-like, diiron-carboxylate proteins participate in a<br>range of functions including iron regulation, mono-oxygenation, and<br>reactive ardical production. These proteins are characterized by the fact                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              | 9.98E-79<br>2.72E-63 | 44%<br>32%      | 63%<br>48%    | 7289<br>21326    | 1285<br>30348  | 3                       |                                      | 89342 fibonucleoside-diphosphate reductase<br>62422 ribonucleotide reductase small subunit                                                                                                                                                                                                                                                                                                                                                                                |                                      | 1.12E-100<br>7.26E-100                       | 54%<br>54%                      | 72%<br>71%                             | 6-326<br>6-326                                     | 6-332<br>6-333                                |
|              |                    |                |               |       | 4                 | cd00657              | that they catalyze dioxygen-dependent oxidation-hydroxylation reactions within circon centers; one exception is minageness catalase, which catalyzes percoxide-dependent oxidation-reduction within a diamagnesse catalase, and the catalyzes percoxide-dependent oxidation-reduction within a diamagnesse presence of duplicates the test (pands, guitamates and histidines (Exxt) and two additional glutamates within a four-helix bundle. Outside of these conserved residues there is little oxivous homology, Members include be phytroxylases (AAMH); obcouncidedire reductase RZ (PINRTZ), apr/AGP hydroxylases (AAMH); obcouncidedire reductase RZ (PINRTZ), apr/AGP demethoxybioquinone hydroxylases (ADMP); DNA protecting proteins (DPS), and bulgenio oxidases (AOM). Additional members include the FGPS) and bulgenio oxidases (AOM).                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |              | 4.31E-08             | 20%             | 30%           | 69–215           | 2140           | 4                       | 4 CAF2                               | 24073 probable ribonucleoside-diphosphate reductase small chain                                                                                                                                                                                                                                                                                                                                                                                                           | 363.61                               | 4.70E-99                                     | 56%                             | 73%                                    | 7–326                                              | 5323                                          |
|              |                    |                |               |       |                   |                      | containing exhunt of the semble curises existen (ACSE) the fertific like of                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |              |                      |                 |               |                  |                | 5<br>6<br>7<br>8<br>9   | XP_55<br>7 NP 91<br>8 CAA7<br>9 EAN8 | 48317 putative ribonucleotide reductase R2 50881 putative ribonucleotide reductase R2 50881 putative ribonucleotide reductase R2 5085 putative ribonucleotide-diphosphate reductase 71741 ribonucleotide reductase (Class I) 5029 ribonucleotide diphosphate reductase small chain 5091 Ribonucleotide diphosphate reductase small chain 61901 Ribonucleotide reductase small chain 61901 Ribonucleotide reductase R2 subuntil 61901 Ribonucleotide reductase R2 subuntil | 355.14<br>355.14<br>354.76<br>353.21 |                                              | 53%<br>53%<br>53%<br>54%<br>54% | 70%<br>71%<br>71%<br>70%<br>70%<br>70% | 2-326<br>6-326<br>6-326<br>6-326<br>6-326<br>6-326 | 10-339<br>20-345<br>2-327<br>16-337<br>16-337 |
| M657L :      | 249244248729       | 172            | 20,087        | 9.95  | 5                 | No Hit Found         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |                      |                 |               |                  |                | 3                       | 2 YP_14<br>3 AAR2                    | 48827 A471R<br>42861 unknown<br>26829 FinV-1-A5<br>77626 EsV-1-141                                                                                                                                                                                                                                                                                                                                                                                                        | 106.69<br>87.81                      | 6.96E-53<br>2.85E-22<br>1.37E-16<br>2.96E-11 | 52%<br>37%<br>31%<br>31%        | 77%<br>61%<br>60%<br>58%               | 2172<br>16155<br>2145<br>3119                      | 3-173<br>43-191<br>14-158<br>20-137           |
|              | 249878249285       |                | 22,211        | 4.64  |                   | No Hit Found         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |                      |                 |               |                  |                |                         | 2 BAB6                               | 48826 A470R<br>69883 UKCH-2                                                                                                                                                                                                                                                                                                                                                                                                                                               | 92.43                                | 1.11E-59<br>8.06E-18                         | 58%<br>39%                      | 71%<br>58%                             | 1197<br>6127                                       | 1202<br>49170                                 |
| M664L 2      | 251214249922       | 431            | 48,625        | 8.54  | ı                 | No Hit Found         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |                      |                 |               |                  |                | 1<br>2<br>3             | 2 AAR2                               | 48824 A468R<br>26870 FirrV-1-A46<br>77530 EsV-1-45                                                                                                                                                                                                                                                                                                                                                                                                                        | 66.63                                | 2.75E-119<br>1.78E-09<br>2.57E-08            | 55%<br>26%<br>26%               | 72%<br>48%<br>47%                      | 3-377<br>15-178<br>6-187                           | 4393<br>17180<br>1187                         |
| M667R :      | 251310252230       | 307            | 35,056        | 7.09  | )                 | No Hit Found         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |              |                      |                 |               |                  |                |                         |                                      | 48823 A467L<br>65022 Hypothetical protein                                                                                                                                                                                                                                                                                                                                                                                                                                 |                                      | 1.57E-109<br>1.20E-08                        | 62%<br>28%                      | 80%<br>46%                             | 4300<br>48221                                      | 6302<br>17218                                 |

| Gene<br>Name | Genome<br>Position               | A.A.<br>length | Peptide<br>Mw       | pl    | CDD Hit<br>Number | COGs         | COG Definition                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | Bit<br>Score     | E-value              | %<br>Identity | %<br>Positive | Query<br>from-to | Hit<br>from-to | BLASTp<br>Hit<br>Number | BLASTp Definition                                                                                                                                                                                                                                                                                                                                          |                                             | Bit<br>Score                      | E-value I                                                                                    | %<br>dentity P                                       |                                                      | Query<br>from-to        | Hit from-<br>to                                                                    |
|--------------|----------------------------------|----------------|---------------------|-------|-------------------|--------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------|----------------------|---------------|---------------|------------------|----------------|-------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------|-----------------------------------|----------------------------------------------------------------------------------------------|------------------------------------------------------|------------------------------------------------------|-------------------------|------------------------------------------------------------------------------------|
| M670F        | R 252261252608                   | 11             | 6 13,165            | 10.47 | ' 1               | pfam04777    | Evr1_AIr, Erv1 / AIr family. Biogenesis of Fe/S clusters involves a<br>number of essential mitochondrial proteins. Evr1p of Saccharomyces<br>cerevisiae mitochondria is required for the maturation of Fe/S proteins in<br>the cytosol. The ALR (augmenter of liver regeneration) represents a                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | 78.03            | 8.91E-16             | 40%           | 52%           | 18110            | ) 2–90         |                         | 1 NP_048821 PBCV-1 thiol oxidoreductase                                                                                                                                                                                                                                                                                                                    |                                             | 177.95                            | 7.45E-44                                                                                     | 67%                                                  | 80%                                                  | 2-116                   | 4118                                                                               |
|              |                                  |                |                     |       |                   |              | mammalian orthologue of yeast Erv1p. Both Erv1p and full-length ALR<br>are located in the mitochondrial intermembrane and it thought to operate<br>downstream of the mitochondrial ABC transporter.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                  |                      |               |               |                  |                |                         |                                                                                                                                                                                                                                                                                                                                                            |                                             |                                   |                                                                                              |                                                      |                                                      |                         |                                                                                    |
|              |                                  |                |                     |       | 2                 | COG5054      | ERV1, Mitochondrial sulfhydryl oxidase involved in the biogenesis of<br>cytosolic Fe/S proteins [Posttranslational modification, protein turnover,<br>chaperones].                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 55.83            | 4.41E-09             | 30%           | 45%           | 1110             | 70175          |                         | 2 YP_142722 putative thiol oxidoreductase                                                                                                                                                                                                                                                                                                                  |                                             | 67.40                             | 1.42E-10                                                                                     | 36%                                                  | 51%                                                  | 10-97                   | 390                                                                                |
|              |                                  |                |                     |       |                   |              | cilabetoriesi.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |                  |                      |               |               |                  |                |                         | 3 XP_503294 hypothetical protein<br>4 NP_078699 Thiol oxidoreductase                                                                                                                                                                                                                                                                                       |                                             | 62.00<br>62.00                    | 5.96E-09<br>5.96E-09                                                                         | 33%<br>33%                                           | 50%<br>55%                                           | 9111<br>299             | 88186<br>21115                                                                     |
|              |                                  |                |                     |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                  |                      |               |               |                  |                |                         | 5 CAH02199 unnamed protein product<br>6 CAA48192 ERV1<br>7 CAA97017 FRV1                                                                                                                                                                                                                                                                                   |                                             | 58.54                             | 2.26E-08<br>6.59E-08<br>6.59E-08                                                             | 31%<br>31%<br>31%                                    | 50%<br>48%<br>48%                                    | 14111<br>14110<br>14110 | 85-178<br>19-111<br>91-183                                                         |
|              |                                  |                |                     |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                  |                      |               |               |                  |                |                         | 8 NP_612265 ORF043L<br>9 AAX82354 thiol oxidoreductase                                                                                                                                                                                                                                                                                                     |                                             | 56.61<br>56.23                    | 2.50E-07<br>3.27E-07                                                                         | 34%<br>34%                                           | 52%<br>52%                                           | 799<br>799              | 596<br>596                                                                         |
|              |                                  |                |                     |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                  |                      |               |               |                  |                |                         | 0 NP_149810 347L                                                                                                                                                                                                                                                                                                                                           | o Suice Brot Accession                      | 56.23                             | 3.27E-07                                                                                     | 30%                                                  | 50%                                                  | 798                     | 398                                                                                |
| M672         | L 253412252615                   | 26             | 6 29,615            | 6.63  | 1 2               |              | Rnc, dsRNA-specific ribonuclease [Transcription]. RIBOc, Ribonuclease III family:                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | 172.76<br>130.75 | 2.47E-44<br>1.10E-31 | 35%<br>42%    | 51%<br>58%    | 38263<br>58187   | 4234<br>2130   |                         | 1 NP_048820 similar to Bacillus ribonuclease III, corresponds t<br>Number P51833<br>2 YP_445467 ribonuclease III                                                                                                                                                                                                                                           | D GWISS-T TOT ACCESSION                     | 348.59<br>133.65                  | 1.12E-94<br>5.64E-30                                                                         | 66%<br>35%                                           | 81%<br>53%                                           | 16265<br>33258          | 25–274<br>9–241                                                                    |
|              |                                  |                |                     |       |                   |              | RIBOc. RIBDo. Ribonuclease III.C terminal domain. This group consists of eukaryotic. Scherial and archeri Prounclease III. (RNAse III) proteins. RNAse III is a double stranded RNA-specific endonuclease. Prokaryotic RNAse III is important in post-transcriptional control of mRNA stability and translational efficiency. It is involved in the processing of ribosomal RNA precursors. Prokaryotic RNAse III also plays are lot in the maturation                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                  |                      |               |               |                  |                |                         |                                                                                                                                                                                                                                                                                                                                                            |                                             |                                   |                                                                                              |                                                      |                                                      |                         |                                                                                    |
|              |                                  |                |                     |       | 3                 | cd00593      | of IRNA precursors and in the processing of phage and plasmid<br>transcripts. Eukaryotic Rhase III Blagons, participate (through direct<br>cleavage) in rRNA processing, in processing of small nucleolar RNAs<br>(snoRNAs) and snrNAsAppos, components of the spiceosome). In<br>eukaryotes RNase III or RNaseIII like enzymes such as Dicer are<br>involved in RNAs (RNA interference) and miRNA (micro-RNA) gene                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 127.32           | 1.16E-30             | 48%           | 62%           | 70181            | 16127          |                         | 3 ZP_00590199 Ribonuclease III                                                                                                                                                                                                                                                                                                                             |                                             | 128.26                            | 2.37E-28                                                                                     | 36%                                                  | 54%                                                  | 45255                   | 47–265                                                                             |
|              |                                  |                |                     |       | 4                 | pfam00636    | eilencinn Ribonuclease 3, RNase3 domain DSRM, Double-stranded RNA binding motif. Binding is not sequence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 123.30           | 1.78E-29             | 56%           | 70%           | 76166            | 1-91           |                         | 4 ZP 00532592 Ribonuclease III                                                                                                                                                                                                                                                                                                                             |                                             | 127.49                            | 4.04E-28                                                                                     | 33%                                                  | 53%                                                  | 16255                   | 21-259                                                                             |
|              |                                  |                |                     |       | 5                 | cd00048      | specific but is highly specific for double stranded RNA. Found in a variety of proteins including dsRNA dependent protein kinase PKR, RNA helicases, Drosophila staufen protein, E. coli RNase III, RNases H1, and                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 58.84            | 5.27E-10             | 31%           | 52%           | 192258           | 1-68           |                         | 5 ABB23018 Ribonuclease III                                                                                                                                                                                                                                                                                                                                |                                             | 122.87                            | 9.95E-27                                                                                     | 35%                                                  | 55%                                                  | 58255                   | 47-252                                                                             |
|              |                                  |                |                     |       | 6                 | COG1939      | dsRNA denendent adenosine deaminases COG1939, Uncharacterized protein conserved in bacteria [Function                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 34.87            | 0.007537             | 29%           | 46%           | 78163            | 18112          |                         | 6 ZP_00591208 Ribonuclease III                                                                                                                                                                                                                                                                                                                             |                                             | 122.48                            | 1.30E-26                                                                                     | 34%                                                  | 52%                                                  | 45256                   | 44-266                                                                             |
|              |                                  |                |                     |       | 7                 | pfam05948    | DUF880, Protein of unknown function (DUF880). This family consists of<br>a number of hypothetical bacterial and plant proteins. The family also<br>contains the C terminal region of a Cysteinyl-tRNA synthetase from                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 35.17            | 0.007546             | 26%           | 43%           | 78167            | 11111          |                         | 7 ZP_00528534 Ribonuclease III                                                                                                                                                                                                                                                                                                                             |                                             | 122.48                            | 1.30E-26                                                                                     | 33%                                                  | 54%                                                  | 41255                   | 58281                                                                              |
|              |                                  |                |                     |       |                   |              | Staphylococcus epidermidis. The function of this family is unknown                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |                  |                      |               |               |                  |                |                         | 8 ZP 00511103 Ribonuclease III<br>9 ZP_00661649 Ribonuclease III<br>0 AAM73335 ribonuclease III                                                                                                                                                                                                                                                            |                                             | 121.32                            | 2.90E-26<br>2.90E-26<br>8.43E-26                                                             | 34%<br>32%<br>32%                                    | 54%<br>50%<br>52%                                    | 45257<br>35257<br>30262 | 42-264<br>19-254<br>15-260                                                         |
| M674F        | R 253807255765                   | 65             | 3 75,012            | 5.23  | 1                 | COG3378      | COG3378, Predicted ATPase [General function prediction only]. Pox D5. Poxvirus D5 protein-like. This family includes D5 from                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | 70.01            | 2.38E-13             | 24%           | 37%           | 281593           | 146445         |                         | 1 NP_048813 contains ATP/GTP-binding site motif A                                                                                                                                                                                                                                                                                                          |                                             | 870.54                            | 0.00E+00                                                                                     | 65%                                                  | 81%                                                  | 20648                   | 22-650                                                                             |
|              |                                  |                |                     |       | 2                 | pfam03288    | Poxviruses which is necessary for viral DNA replication, and is a nucleic acid independent nucleoside triphosphatase. Members of this family are also found outside of poxviruses                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | 63.73            | 1.55E-11             | 22%           | 39%           | 290614           | 3313           |                         | 2 NP_077594 EsV-1-109                                                                                                                                                                                                                                                                                                                                      |                                             | 278.10                            | 6.60E-73                                                                                     | 34%                                                  | 54%                                                  | 190607                  | 141-565                                                                            |
|              |                                  |                |                     |       |                   |              | asu tionia dustre o trouvituses                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                  |                      |               |               |                  |                |                         | 3 AAR26902 FirrV-1.827 4 YP_29427 jouther nucleic acid independent nucleoside tripl 5 ZP 0012428 COG3378. Predicted ATPase 6 BAE69402 jouther Duk Dimase-phase associated 8 AR27348 phase phi-R73 primase-like protein 8 ZP_00503786 Phase(pilsmid primase P4. C-terminal 9 ZP_00593388 Phase(pilsmid primase P4. C-terminal 9 ZP_00503788 Podicted ATPase | nosphatase                                  | 166.78<br>65.86<br>61.62<br>61.23 | 8.92E-70<br>2.14E-39<br>5.15E-09<br>9.71E-08<br>1.27E-07<br>2.16E-07<br>8.22E-07<br>1.07E-06 | 29%<br>25%<br>28%<br>24%<br>24%<br>22%<br>23%<br>30% | 51%<br>44%<br>48%<br>42%<br>43%<br>42%<br>42%<br>53% | 292566<br>265523        | 6-547<br>168-672<br>310-495<br>107-366<br>213-458<br>241-499<br>307-592<br>439-581 |
| M680F        | R 255802256674                   | 29             | 1 31,259            | 4.95  | i                 | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                  |                      |               |               |                  |                |                         | 1 NP 048811 A454L                                                                                                                                                                                                                                                                                                                                          |                                             | 354.37                            | 2.37E-96                                                                                     | 64%                                                  | 77%                                                  | 20291                   | 20-289                                                                             |
| M683         | L 257390256689                   | 23             | 4 26,931            | 10.80 | ) 1               | cd00283      | GIY-YIG Clerm, GIYX(10-11)YIG family of class I homing<br>endonucleases C-termius (GIY-YIC, Clerm), Homing endonucleases<br>promote the mobility of intron or inteni by recognizing and cleaving a<br>homologous allele that lacks the sequence. They catalyze a double-<br>strand treak in the DAN ears the insertion star of that element to facilitate<br>strand treak in the DAN ears the insertion star of that element to facilitate<br>families based on the presence of these mofflis in their respective N-<br>termini: LAGLIDADC, His-Cyg box, HNH, and GIY-YIG. This CD contains<br>several but not all members of the GIY-YIG family. The Cetermius of                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 36.90            | 0.002104             | 38%           | 55%           | 120199           | 1–77           |                         | 1 NP_048851 elimilar to PBCV-1 ORF A315L, corresponds to Number M744440                                                                                                                                                                                                                                                                                    | o GenBank Accession                         | 150.98                            | 2.70E-35                                                                                     | 41%                                                  | 57%                                                  | 12234                   | 2-219                                                                              |
|              |                                  |                |                     |       |                   |              | GIY-YIG is a DNA-binding domain which is separated from the N-<br>terminus by a long, flexible linker. The DNA-binding domain consists of a<br>minor-grove binding alpha-helix, and a helix-furn-helix. Some also<br>contain a zinc finger (ie. 1-Tevl) which is not required for DNA binding or<br>catalysis, but is a component of the linker and directs the catalytic<br>domain to cleave the homing site at a fixed distance from the intron<br>insertion: site                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |                  |                      |               |               |                  |                |                         | ···········                                                                                                                                                                                                                                                                                                                                                |                                             |                                   |                                                                                              |                                                      |                                                      |                         |                                                                                    |
|              |                                  |                |                     |       | 2                 | smart00465   | GIYc, GIY-YIG type nucleases (URI domain); .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | 36.21            | 0.003097             | 29%           | 49%           | 1099             | 1-83           |                         | <ol> <li>NP_049007 similar to Chlorella virus PBCV-1 ORF A315L, or<br/>Accession Number U42580</li> <li>NP 048671 A315L</li> </ol>                                                                                                                                                                                                                         | orresponds to GenBank                       |                                   | 9.33E-20<br>6.05E-19                                                                         | 34%<br>32%                                           | 51%<br>50%                                           | 12205<br>12211          | 2195<br>2216                                                                       |
|              |                                  |                |                     |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                  |                      |               |               |                  |                |                         | 4 NP 048641 PBCV-1 33kd peptide<br>5 NP_899393 SegD                                                                                                                                                                                                                                                                                                        |                                             |                                   | 4.20E-12<br>3.32E-09                                                                         | 31%<br>31%                                           | 51%<br>50%                                           | 15194<br>21182          | 6-203<br>12-154                                                                    |
|              |                                  |                |                     |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                  |                      |               |               |                  |                |                         | 6 AAT53588 group I intron GIY-YIG endonuclease<br>7 ZP_00391680 COG0532: Translation initiation factor 2 (IF-2; G1<br>8 AAU16837 GIY-YIG catalytic domain containing protein; p                                                                                                                                                                            | Pase)<br>ossible intron encoded             | 62.00<br>56.23<br>55.84           | 1.65E-08<br>9.05E-07<br>1.18E-06                                                             | 29%<br>28%<br>23%                                    | 47%<br>47%<br>41%                                    | 11186<br>35186<br>10197 | 2190<br>2165<br>2182                                                               |
|              |                                  |                |                     |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                  |                      |               |               |                  |                |                         | endonuclease 9 ABA03239 putative homing endonuclease 0 CAC51107 putative GIY-YIG endonuclease                                                                                                                                                                                                                                                              |                                             | 54.68                             | 2.63E-06<br>1.71E-05                                                                         | 42%<br>24%                                           | 64%<br>42%                                           | 118185                  | 109-171<br>7-206                                                                   |
| M686I        | L 258678257446                   | 41             | 1 44,516            | 4.13  | 1                 | pfam07415    | Herpes_LMP2, Gammaherpesvirus latent membrane protein (LMP2) protein. This family consists of several Gammaherpesvirus latent membrane protein (LMP2) proteins. Epstein-Bar virus is a human Gammaherpesvirus that infects and establishes latency in B lymphocytes in vivo. The latent membrane protein 2 (LMP2) gene the expressed in latently infected 5 cells and encodes two protein isoforms, LMP2A and conditionally infected 5 cells and encodes two protein isoforms, LMP2A and conditionally consistent with the control in the LMP2A logorithm. LMP2A is conditionally consistent with its present in the LMP2A logorithm. LMP2A is set to the conditional consistent with the consistent consistent consistent with the consistent of the LMP2A logorithm. LMP2A is set to the consistent consistent with the consistent consisten | 34.95            | 0.007536             | 19%           | 32%           | 266331           | 3194           |                         | contains Pro-rich Px molif EPSPEPxP (5X), and 1 NP_048415 to trypanosome procyclin precursor, correr Accession Number P04869                                                                                                                                                                                                                               | PEST sequence; similar sponds to Swiss-Prot |                                   | 3.24E-05                                                                                     | 24%                                                  | 48%                                                  | 16149                   |                                                                                    |
|              |                                  |                |                     |       |                   |              | thought to play a key role in either the establishment or the maintenance of latency and/or the reactivation of productive infection from the latent state. The significance of LMP2B and its role in pathogenesis remain unclear                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                  |                      |               |               |                  |                |                         |                                                                                                                                                                                                                                                                                                                                                            |                                             |                                   |                                                                                              |                                                      |                                                      |                         |                                                                                    |
|              | R 258822259511                   |                | 0 25,409            |       |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                  |                      |               |               |                  |                |                         | No Hit Found No Hit Found                                                                                                                                                                                                                                                                                                                                  |                                             |                                   |                                                                                              |                                                      |                                                      |                         |                                                                                    |
|              | R 259590259793<br>L 260573260178 |                | 8 7,882<br>2 15,046 |       |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                  |                      |               |               |                  |                |                         | No Hit Found No Hit Found  1 NP 048926 A570L                                                                                                                                                                                                                                                                                                               |                                             | 177 18                            | 1.27E-43                                                                                     | 65%                                                  | 76%                                                  | 1118                    | 1121                                                                               |
|              | R 260638261159                   |                | 4 19,819            |       |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                  |                      |               |               |                  |                |                         | 1 NP_048928 A572R                                                                                                                                                                                                                                                                                                                                          |                                             |                                   | 1.55E-63                                                                                     | 65%                                                  | 85%                                                  | 4-163                   | 11-170                                                                             |
| M697         | L 261989261162                   | 27             | 6 30,772            | 4.22  | ! 1               | pfam02747    | PCNA_C, Proliferating cell nuclear antigen, C-terminal domain. N-terminal and C-terminal domains of PCNA are topologically identical.<br>Three PCNA nelocules are tightly associated to form a closed ring encirclinn dunler DNA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 64.96            | 7.73E-12             | 28%           | 49%           | 152267           | 3128           |                         | 1 NP_048930 similar to Periwinkle PCNA, corresponds to GenE<br>X55052                                                                                                                                                                                                                                                                                      | lank Accession Number                       | 230.34                            | 4.75E-59                                                                                     | 44%                                                  | 67%                                                  | 21276                   | 4264                                                                               |

| Gene<br>Name | Genome<br>Position | A.A.<br>length | Peptide<br>Mw | pl    | CDD Hit<br>Number | COGs         | COG Definition                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | Bit<br>Score   | E-value              | %<br>Identity P | %<br>ositive | Query<br>from-to | Hit<br>from-to | BLASTp<br>Hit<br>Number              | BLASTo Definition                                                                                                                                                                                                                                                                                                                                                                               | Bit<br>Score                                                         | E-value                                                              | %                                                           |                                                             | Query<br>from-to                                                                       | Hit from-<br>to                                                                    |
|--------------|--------------------|----------------|---------------|-------|-------------------|--------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------|----------------------|-----------------|--------------|------------------|----------------|--------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------|----------------------------------------------------------------------|-------------------------------------------------------------|-------------------------------------------------------------|----------------------------------------------------------------------------------------|------------------------------------------------------------------------------------|
|              |                    | •              |               |       | 2                 | COG0592      | DnaN, DNA polymerase sliding clamp subunit (PCNA homolog) [DNA reolication. recombination. and reoair].                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 59.16          | 3.81E-10             | 17%             | 37%          |                  | 47323          | Number<br>2                          |                                                                                                                                                                                                                                                                                                                                                                                                 | 123.64                                                               | 6.26E-27                                                             | 29%                                                         | 50%                                                         | 32274                                                                                  | 2260                                                                               |
|              |                    |                |               |       | 3                 | pfam00705    | PCNA_N, Proliferating cell nuclear antigen, N-terminal domain. N-terminal and C-terminal domains of PCNA are topologically identical.  Three PCNA molecules are tightly associated to form a closed ring                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | 49.88          | 2.40E-07             | 25%             | 55%          | 42129            | 12100          | 3                                    | 3 XP_502661 hypothetical protein                                                                                                                                                                                                                                                                                                                                                                | 115.93                                                               | 1.30E-24                                                             | 25%                                                         | 52%                                                         | 42275                                                                                  | 12-260                                                                             |
|              |                    |                |               |       |                   |              | endirdina dualex DNA.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                |                      |                 |              |                  |                | 4<br>5<br>6<br>7<br>8<br>9           | CAA38836 proliferating cell nuclear antigen     AA42827 Prona (proliferating cell nuclear antigen) homolog protein 1     CAA38830 proliferating cell nuclear antigen (PCNA)     CAMAY38 proliferating cell nuclear antigen (PCNA)     AA624980 proliferating cell nuclear antigen     XP_1X255 proliferating cell nuclear antigen     AA627491 PCNA     AA627491 PCNA                           | 115.16<br>115.16<br>113.24<br>113.24<br>112.85<br>112.85             | 2.22E-24<br>8.45E-24<br>8.45E-24<br>1.10E-23<br>1.10E-23             | 29%<br>29%<br>30%<br>28%<br>30%<br>29%<br>28%               | 49%<br>51%<br>49%<br>50%<br>49%<br>51%<br>50%               | 37-274<br>65-273<br>42-275<br>42-274<br>42-274<br>38-268<br>42-275                     | 7-260<br>1-226<br>12-262<br>12-261<br>12-261<br>7-261<br>12-262                    |
| M699L        | 262481261993       | 163            | 18,190        | 10.90 |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |                |                      |                 |              |                  |                |                                      | No Hit Found No Hit Found                                                                                                                                                                                                                                                                                                                                                                       |                                                                      |                                                                      |                                                             |                                                             |                                                                                        |                                                                                    |
| M700L        | 263013262513       | 167            | 19,586        | 10.14 |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |                |                      |                 |              |                  |                | 1                                    | 1 NP_048931 A575L                                                                                                                                                                                                                                                                                                                                                                               | 112.46                                                               | 4.82E-24                                                             | 35%                                                         | 66%                                                         | 31166                                                                                  | 30167                                                                              |
| M701R        | 263097264128       | 344            | 37,917        | 8.48  | 1                 |              | Trypan_PARP. Procyclic acidir espetitive protein (PARP). This family consists of several Trypansoma bruce procyclic acidir espetitive protein (PARP) like sequences. The procyclic acidir repetitive protein (parp) genes of Trypansoma brucei encode a small family of abundant surface proteins whose expression is restricted to the procyclic form of the parasite. They are found at two unlinked loci, parpA and parpB; transcription of both lois developmentality remutated.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 53.06          | 3.06E-08             | 33%             | 52%          | 222276           | 68122          | 1                                    | 1 NP_048889 PLPRNLLL (4X), SPPPSKP (3X)                                                                                                                                                                                                                                                                                                                                                         | 345.13                                                               | 1.87E-93                                                             | 70%                                                         | 80%                                                         | 7220                                                                                   | 1213                                                                               |
|              |                    |                |               |       | 2                 | pfam05616    | Neisseria_TspB, Neisseria meningitidis TspB protein. This family consists of several Neisseria meningitidis TspB virulence factor proteins                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 50.13          | 1.91E-07             | 38%             | 54%          | 224274           | 323-373        | 2                                    | 2 NP_048688 a332L                                                                                                                                                                                                                                                                                                                                                                               | 88.20                                                                | 4.12E-16                                                             | 74%                                                         | 90%                                                         | 284326                                                                                 | 143                                                                                |
|              |                    |                |               |       | 3                 | COG0810      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 46.29          | 2.85E-06             | 51%             | 53%          | 222278           | 68125          | 3                                    | 3 EAA01148 ENSANGP0000018413                                                                                                                                                                                                                                                                                                                                                                    | 71.25                                                                | 5.22E-11                                                             | 27%                                                         | 42%                                                         | 25217                                                                                  | 4202                                                                               |
|              |                    |                |               |       | 4                 |              | DedD, Uncharacterized protein conserved in bacteria [Function unknown]. DEC-1 protein, N terminal region. The defective chorion-1 page (dec-1) in Drosophila encodes follicle cell proteins pages any for                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | 44.63          | 9.90E-06             | 28%             | 40%          | 222275           | 97150          | 4                                    | 4 EAA06469 ENSANGP0000012390                                                                                                                                                                                                                                                                                                                                                                    | 69.71                                                                | 1.52E-10                                                             | 27%                                                         | 42%                                                         | 33218                                                                                  | 1196                                                                               |
|              |                    |                |               |       | 5                 | pfam04625    | proper eggshell assembly. Multiple products of the dec-1 gene are formed by alternative RNA splicing and proteolytic processing. Cleavage products include S80 (80 kDa) which is incorporated into the eggshell,                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | 39.48          | 0.0003               | 24%             | 45%          | 226275           | 94143          | 5                                    | 5 AAS15667 LP24064p                                                                                                                                                                                                                                                                                                                                                                             | 66.24                                                                | 1.68E-09                                                             | 28%                                                         | 43%                                                         | 18216                                                                                  | 8210                                                                               |
|              |                    |                |               |       | 6                 |              | and further proteolysis of S80 gives S60 (60 kDa) ChtBD2, Chitin-binding domain type 2; . COG3397, Uncharacterized protein conserved in bacteria [Function                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 39.34          | 0.0004               | 40%             | 50%          | 282325           |                | 6                                    | 6 AAM50982 RE24790p                                                                                                                                                                                                                                                                                                                                                                             |                                                                      | 2.19E-09                                                             | 29%                                                         | 42%                                                         | 20218                                                                                  | 38-249                                                                             |
|              |                    |                |               |       | 7                 | COG3397      | Nucleo_P87, Nucleopolyhedrovirus capsid protein P87. This family                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | 39.27          | 0.000431             | 21%             | 31%          | 11235            | 8226           | 7                                    | 7 AAF46012 CG15786-PA                                                                                                                                                                                                                                                                                                                                                                           | 65.86                                                                | 2.19E-09                                                             | 29%                                                         | 42%                                                         | 20218                                                                                  | 19230                                                                              |
|              |                    |                |               |       | 8                 | pfam07267    | P87 is expressed late in infection and concentrated in infected cell nuclei  CBM 14, Chitin binding Peritrophin-A domain. This domain is called the                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | 37.48          | 0.001331             | 19%             | 32%          | 215275           | 319-378        | 8                                    | 8 EAL29006 GA18133-PA                                                                                                                                                                                                                                                                                                                                                                           | 64.31                                                                | 6.38E-09                                                             | 27%                                                         | 42%                                                         | 18216                                                                                  | 6210                                                                               |
|              |                    |                |               |       | 9                 | pfam01607    | peritorybic matrix proteins of insects and animal chrinases. Copies of the<br>domain are also found in some baculoviruses. Relevant references that<br>describe proteins with this domain include. It is an extracellular domain<br>that contains six conserved cysteries that probably from three disulphide<br>bridges. Chilin binding has been demonstrated for a protein containing<br>only two not these domains.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | 37.34          | 0.001689             | 41%             | 62%          | 294325           | 1847           | 9                                    | 9 EAL32472 GA13958-PA                                                                                                                                                                                                                                                                                                                                                                           | 64.31                                                                | 6.38E-09                                                             | 27%                                                         | 41%                                                         | 30218                                                                                  | 11-209                                                                             |
|              |                    |                |               |       | 10                | pfam03067    | Chitin_bind_3, Chitin binding domain. This domain is found associated with a wide variety of cellulose binding domain. This domain however is a chitin binding domain. This domain is found in isolation in baculoviral spheroidins and spindolins, protein of unknown function                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 37.28          | 0.001746             | 23%             | 30%          | 18216            | 1211           | 10                                   | 0 EAA00829 ENSANGP00000011567                                                                                                                                                                                                                                                                                                                                                                   | 60.08                                                                | 1.20E-07                                                             | 28%                                                         | 39%                                                         | 17216                                                                                  | 7215                                                                               |
| M704R        | 264221264748       | 176            | 19,645        | 11.69 |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |                |                      |                 |              |                  |                | 1                                    | 1 NP_048877 contains Gln-rich, neutral zinc metallopeptidase, zinc binding signature                                                                                                                                                                                                                                                                                                            | region 200.68                                                        | 1.56E-50                                                             | 58%                                                         | 71%                                                         | 1-175                                                                                  | 212391                                                                             |
|              | 266241264841       | 467            | 51,912        | 8.35  |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |                |                      |                 |              |                  |                | 1                                    | 1 NP_048863 a507R                                                                                                                                                                                                                                                                                                                                                                               | 83.96                                                                | 1.20E-14                                                             | 34%                                                         | 52%                                                         | 143313                                                                                 | 6182                                                                               |
|              | 266984266688       | 99             | 11,584        | 8.38  |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |                |                      |                 |              |                  |                |                                      | No Hit Found No Hit Found                                                                                                                                                                                                                                                                                                                                                                       |                                                                      |                                                                      |                                                             |                                                             |                                                                                        |                                                                                    |
| M715R        | 266983267303       | 107            | 12,723        | 6.80  |                   | No Hit Found | ANK, ankyrin repeats; ankyrin repeats mediate protein-protein interactions in very diverse families of proteins. The number of ANK                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                |                      |                 |              |                  |                | 1                                    | 1 NP_048858 A502L                                                                                                                                                                                                                                                                                                                                                                               | 103.99                                                               | 1.37E-21                                                             | 54%                                                         | 65%                                                         | 21107                                                                                  | 395                                                                                |
| M717R        | 267359268519       | 387            | 43,765        | 6.67  | 1                 | cd00204      | repeats in a protein can range from 2 to over 20 (ankyrins, for example).<br>ANK repeats may occur in combinations with other types of domains. The<br>structural repeat unit contains two antiparallel helices and a beta-hairpin,<br>repeats are stacked in a superhelical arrangement; this alignment                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | 60.10          | 1.95E-10             | 30%             | 58%          | 65-170           | 8112           | 1                                    | 1 NP_048963 contains 3 ankyrin repeat-like elements; similar to Drosophila al corresponds to GenBank Accession Number L35601                                                                                                                                                                                                                                                                    | nkyrin, 196.05                                                       | 1.67E-48                                                             | 45%                                                         | 63%                                                         | 4220                                                                                   | 3217                                                                               |
|              |                    |                |               |       | 2                 | COG0666      | contains 4 consecutive remark  Arp, FOG: Ankyrin repeat [General function prediction only].                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | 40.26          | 0.000187             | 17%             | 36%          | 63219            | 72-234         | 2<br>3<br>4<br>5<br>6<br>7<br>8<br>9 | NP_048964 A609R     NP_04897 PREDICTED similar to CQ8599-PA     AV025891 ankyrin repeat protein E4 2     AV22517 PREDICTED similar to CQ8599-PA     AV2525 PREDICTED similar to ankyrin repeat protein     XP_17255 PREDICTED similar to ankyrin repeat domain 28     XP_272539 PREDICTED similar to ankyrin repeat domain 28     EAL45264 ankyrin repeat protein, putative     NP_048786 A429L | 96.67<br>70.86<br>59.31<br>58.54<br>57.77<br>56.61<br>56.61<br>55.84 | 8.11E-11<br>2.44E-07<br>4.17E-07<br>7.11E-07<br>1.58E-06<br>1.58E-06 | 33%<br>27%<br>31%<br>26%<br>28%<br>28%<br>24%<br>27%<br>22% | 54%<br>44%<br>51%<br>44%<br>49%<br>47%<br>41%<br>44%<br>40% | 239-384<br>1-258<br>46-170<br>28-231<br>42-170<br>94-286<br>33-265<br>42-300<br>26-308 | 1149<br>134368<br>29152<br>144328<br>202332<br>121323<br>188415<br>192430<br>57357 |
| M719L        | 269688268522       | 389            | 43,206        | 5.23  | 1                 |              | Ugd, Predicted UDP-glucose 6-dehydrogenase [Cell envelope<br>hionenasis. outer membrane]<br>WecC, UDP-N-acetyl-D-mannosaminuronate dehydrogenase [Cell                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | 354.11         |                      | 36%             | 56%          | 4375             |                |                                      | 1 NP_048965 PBCV-1 UDP-glucose dehydrogenase                                                                                                                                                                                                                                                                                                                                                    |                                                                      | 8.73E-130                                                            | 59%                                                         | 77%                                                         | 5386                                                                                   | 4386                                                                               |
|              |                    |                |               |       | 2                 |              | envelope biogenesis, outer membrane1.  UDPG MGDP dh N. UDP-glucose/GDP-mannose dehydrogenase                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | 124.96         | 5.80E-30             | 28%             | 44%          | 1364             | 8393           | 2                                    | 2 AAK02860 unknown                                                                                                                                                                                                                                                                                                                                                                              | 422.94                                                               | 8.48E-117                                                            | 54%                                                         | 74%                                                         | 3389                                                                                   | 2389                                                                               |
|              |                    |                |               |       | 3                 | pfam03721    | family, NAD binding domain. The UDP-glucose/GDP-mannose<br>dehydrogenaseses are a small group of enzymes which possesses the<br>ability to catalyse the NAD-dependent 2-fold oxidation of an alcohol to an<br>acid without the release of an alidehyde intermediate<br>UDPG MGDP dh, UDP-glucose/GDP-mannose dehydrogenase family,                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 122.65         | 3.33E-29             | 32%             | 49%          | 3176             | 1186           | 3                                    | 3 AAC67251 UDP-glucose dehydrogenase                                                                                                                                                                                                                                                                                                                                                            | 421.78                                                               | 1.89E-116                                                            | 53%                                                         | 74%                                                         | 3389                                                                                   | 2389                                                                               |
|              |                    |                |               |       | 4                 |              | central domain. The UDP-glucose/GDP-mannose dehydrogenaseses are<br>a small group of enzymes which possesses the ability to catalyse the<br>NAD-dependent 2-fold oxidation of an alcohol to an acid without the<br>release of an aldehyde intermediate.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |                | 9.51E-20             | 39%             | 55%          | 193285           |                | 4                                    | 4 AAK17922 UDP-glucose dehydrogenase FcbC                                                                                                                                                                                                                                                                                                                                                       |                                                                      | 3.02E-114                                                            | 53%                                                         | 73%                                                         | 3-389                                                                                  | 2389                                                                               |
|              |                    |                |               |       | 5                 | COG1893      | ApbA, Ketopantoate reductase [Coenzyme metabolism].<br>GpsA, Glycerol-3-phosphate dehydrogenase [Energy production and                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | 46.85<br>45.92 | 1.95E-06<br>3.46E-06 | 16%<br>32%      | 39%<br>57%   | 4249<br>387      |                | 5<br>6                               | 5 AAK17904 UDP-qlucose dehydrogenase DcbC<br>6 CAG21035 putative UDP-glucose dehydrogenase                                                                                                                                                                                                                                                                                                      |                                                                      | 1.95E-113<br>3.12E-111                                               | 53%<br>51%                                                  | 72%<br>73%                                                  | 3389<br>5389                                                                           | 2389<br>3388                                                                       |
|              |                    |                |               |       | 7                 |              | conversion).  MmsR 3-hydroxyischutyrate dehydrogenase and related heta-                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 43.29          | 2.72E-05             | 29%             | 45%          | 3232             | 1204           | 7                                    | 7 ZP_00851874 UDP-glucose 6-dehydrogenase                                                                                                                                                                                                                                                                                                                                                       |                                                                      | 5.88E-110                                                            | 51%                                                         | 71%                                                         | 5389                                                                                   | 3388                                                                               |
|              |                    |                |               |       | 8                 |              | hydroxyacid dehydrocenases [Lioid metabolism]. FadB, 3-hydroxyacyl-CoA dehydroqenase [Lioid metabolism]. UDPG_MGDP_dh_C, UDP-glucose/GDP-mannose dehydrogenase family, UDP binding domain. The UDP-glucose/GDP-mannose                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | 41.40          | 8.91E-05             | 25%             | 47%          | 188              |                |                                      | 8 ZP 00880860 UDP-glucose 6-dehydrogenase                                                                                                                                                                                                                                                                                                                                                       |                                                                      | 8.50E-109                                                            | 50%                                                         | 71%                                                         | 5389                                                                                   | 3388                                                                               |
|              |                    |                |               |       | 9                 |              | dehydrogenaseses are a small group of enzymes which possesses the<br>ability to catalyse the NAD-dependent 2-fold oxidation of an alcohol to an<br>acid without the release of an alciholwde intermediate<br>UbiH, 2-polyprenyl-6-methoxyphenol hydroxylase and related FAD-                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | 37.96          | 0.00092              | 17%             | 38%          | 302383           |                |                                      | 9 ZP_00579286 UDP-glucose 6-dehydrogenase                                                                                                                                                                                                                                                                                                                                                       |                                                                      | 7.19E-108                                                            | 48%                                                         | 72%                                                         | 5-389                                                                                  | 11–396                                                                             |
| M7041        | 271560269749       | 00.            | 69.547        | 7.89  | 10                |              | dependent oxidoreductases [Coenzyme metabolism / Energy production and conversion].  COG1215, Glycosytransferases, probably involved in cell wall become in the conversion of |                | 0.001585<br>2.39E-25 | 39%             | 68%          | 131              | 1-32           |                                      | AAM38424 UDP-glucose dehydrogenase      NP_048569 similar to cellulose synthase catalytic subunit (UDP-forming)                                                                                                                                                                                                                                                                                 |                                                                      | 9.39E-108<br>0.00E+00                                                | 51%                                                         | 70%                                                         | 5-386                                                                                  | 3-385<br>72-677                                                                    |
| WI721L       | 2/1000-209/49      | 604            | 00,347        | 1.09  | '                 | JUG 1215     | hindenesis [Cell envelone hindenesis outer membrane]                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 109.04         | 4.50E-23             | 1870            | 3/70         | JU-44/           | 14-402         | 1                                    |                                                                                                                                                                                                                                                                                                                                                                                                 | 090.00                                                               | 0.00E*00                                                             | UO70                                                        | 0170                                                        | 1-004                                                                                  | 12-011                                                                             |

| Gene Genome<br>Name Position | A.A.<br>length | Peptide<br>Mw | pl    | CDD Hit<br>Number | COGs         | COG Definition                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | Bit<br>Score    | E-value              | %          | %<br>Positive | Query<br>from-to | Hit<br>from-to      | BLASTp<br>Hit | Hit<br>Accession                                                                                            | BLASTp Definition                                                                                                                                                                                                                                                      | Bit<br>Score            | E-value                                                                          | %<br>Identity P                               |                                               | Query<br>from-to                                            | Hit from-                               |
|------------------------------|----------------|---------------|-------|-------------------|--------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------|----------------------|------------|---------------|------------------|---------------------|---------------|-------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------|----------------------------------------------------------------------------------|-----------------------------------------------|-----------------------------------------------|-------------------------------------------------------------|-----------------------------------------|
|                              |                |               |       |                   |              | Cellulose_synt, Cellulose synthase. Cellulose, an aggregate of<br>unbranched polymers of beta-1,4-linked glucose residues, is the major                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |                 |                      |            |               |                  |                     | Number        |                                                                                                             |                                                                                                                                                                                                                                                                        |                         |                                                                                  |                                               |                                               |                                                             |                                         |
|                              |                |               |       | 2                 | pfam03552    | component of wood and thus paper, and is synthesised by plants, most algae, some bacteria and fungi, and even some animast. The genes that synthesise cellulose in higher plants differ greatly from the well-characterized genes found in Acothobacter and Agrobacterium sp. More correctly designated as Aspos.cellulose synthase cetaltytic subunitsAspos., plant cellulose synthase (cestal) proteins are integral                                                                                                                                                                                                                                                | 40.66           | 0.00015              | 24%        | 43%           | 26246            | 7 499698            | 2             | AAL44127 cellulo                                                                                            | sse synthase                                                                                                                                                                                                                                                           | 305.06                  | 4.60E-81                                                                         | 41%                                           | 57%                                           | 82474                                                       | 91–478                                  |
|                              |                |               |       |                   |              | membrane proteins, approximately 1,000 amino acids in length. There are a number of highly conserved residues, including several motifs shown to be necessary for processive allocosyltransferase activity Glycos transf 2. Glycosyl transferase. Diverse family, transferring sugar                                                                                                                                                                                                                                                                                                                                                                                  |                 |                      |            |               |                  |                     |               |                                                                                                             |                                                                                                                                                                                                                                                                        |                         |                                                                                  |                                               |                                               |                                                             |                                         |
|                              |                |               |       | 3                 | pfam0053     | from UDP-glucose, UDP-N-acetyl- galactosamine, GDP-mannose or CDP-abequose, to a range of substrates including cellulose, dolichol phosphate and teichoic acids                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 38.66           | 0.000678             | 15%        | 32%           | 10627            | 3 10168             | 3             | BAB54246 mlr787                                                                                             | 73                                                                                                                                                                                                                                                                     | 288.12                  | 5.81E-76                                                                         | 38%                                           | 53%                                           | 93531                                                       | 112546                                  |
|                              |                |               |       |                   |              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |                 |                      |            |               |                  |                     | 6             | G CAC48842 putativ<br>G AAQ87082 Cellulo<br>ZP_00913218 similar<br>G YP 418867 glycos<br>ABA79331 cellulo   | se synthase-like protein ce cellulose synthase protein se synthase protein se synthase catalytic subunit to cellulose synthase to cellulose synthase to cellulose synthase to recomp 2 family protein ses synthase-like protein yl transferase, group 2 family protein | 278.10<br>276.17        | 6.43E-75<br>1.10E-74<br>7.11E-74<br>1.21E-73<br>6.02E-73<br>2.29E-72<br>2.29E-72 | 36%<br>38%<br>38%<br>40%<br>37%<br>35%<br>37% | 53%<br>53%<br>53%<br>57%<br>53%<br>52%<br>53% | 93557<br>90526<br>93535<br>93471<br>90523<br>93559<br>90523 | 102-539<br>166-538<br>93-516<br>110-568 |
| M727L 2719922716             | 36 11          | 9 13,075      | 10.35 | 5 1               | smart00317   | y SET, SET (Su(var)3-9, Enhancer-of-zeste, Trithorax) domain; Putative<br>methyl transferase. based on outlier olant homologues.<br>SET, SET domain. SET domains are protein lysine methyltransferase<br>enzymes. SET domains appear to be protein-protein interaction domains.                                                                                                                                                                                                                                                                                                                                                                                       | 58.87           | 5.39E-10             | 29%        | 38%           | 611              | 3 2125              | 1             | NP_048968 PBCV                                                                                              | -1 histone H3-Lys 27 methyltransferase (vSET)                                                                                                                                                                                                                          | 141.35                  | 7.64E-33                                                                         | 57%                                           | 73%                                           | 2114                                                        | 1113                                    |
|                              |                |               |       | 2                 | pfam00856    | enzymes. Set i comains appeal to be protein-protein interaction contains<br>that has been demonstrated that SET domains mediate interactions with a<br>family of proteins that display similarity with dual-specificity phosphatases<br>(dsPTPases). A subset of SET domains have been called PR AGmains.<br>These domains are divergent in sequence from other SET domains, but<br>also appear to mediate protein-protein interaction.                                                                                                                                                                                                                               | 48.14           | 8.37E-07             | 24%        | 39%           | 610              | 8 8128              | 2             | . ZP_00588496 Nuclea                                                                                        | ar protein SET                                                                                                                                                                                                                                                         | 65.86                   | 4.08E-10                                                                         | 34%                                           | 52%                                           | 7114                                                        | 39150                                   |
|                              |                |               |       | 3                 | COG294       | COG2940, Proteins containing SET domain [General function prediction only].                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | 42.79           | 3.73E-05             | 22%        | 35%           | 110              | 8 328450            |               | ZP_00661322 Nuclea                                                                                          |                                                                                                                                                                                                                                                                        |                         | 1.71E-08                                                                         | 35%                                           | 51%                                           | 6117                                                        |                                         |
|                              |                |               |       |                   |              | S TKc. Serine/Threonine protein kinases, catalviic domain.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                 |                      |            |               |                  |                     |               | i ZP_00511449 Nuclea<br>i EAM63832 Nuclea<br>i ZP_00543998 Nuclea<br>i ABB23988 Nuclea<br>i AAM32541 hypoth | ar protein SET<br>ar protein SET                                                                                                                                                                                                                                       | 58.54<br>57.77<br>56.61 | 2.24E-08<br>3.82E-08<br>6.51E-08<br>1.11E-07<br>2.47E-07<br>5.51E-07<br>9.40E-07 | 31%<br>34%<br>31%<br>31%<br>34%<br>32%<br>32% | 45%<br>50%<br>49%<br>52%<br>51%<br>55%<br>58% | 3-111<br>7-114<br>7-116<br>7-106<br>6-114<br>16-106         | 3109<br>38150                           |
| M729L 2737762720             | 28 58          | 3 65,559      | 11.57 | 7 1               | cd0018i      | Phosphotransferases of the serine or threonine-specific kinase subfamily.  The enzymatic activity of these protein kinases is controlled by phosphorylation of specific residues in the activation segment of the catalytic domain, sometimes combined with reversible conformational                                                                                                                                                                                                                                                                                                                                                                                 | 43.66           | 2.15E-05             | 30%        | 46%           | 53-23            | 8 6151              | ,             | NP_048970 RPQT                                                                                              | -like (9x)                                                                                                                                                                                                                                                             | 478.02                  | 3.78E-133                                                                        | 44%                                           | 62%                                           | 2583                                                        | 6577                                    |
|                              |                |               |       | 2                 | COG051       | channes in the C-terminal authoreulatory tail SPS1, Serine/threonine protein kinase [General function prediction only / 5 Signal transduction mechanisms / Transcription / DNA replication,                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | 41.30           | 9.80E-05             | 13%        | 28%           | 8447             | 4 11383             | 2             | NP 048632 similar                                                                                           | to bovine cylicin I, corresponds to Swiss-Prot Accession Number                                                                                                                                                                                                        | 300.06                  | 1.41E-79                                                                         | 35%                                           | 55%                                           | 1-524                                                       | 9527                                    |
|                              |                |               |       | 3                 |              | recombination. and reoairl.  § T.KC, Serine/Threonine protein kinases, catalytic domain; Phosohotransferases. Serine or threonine-specific kinase subfamilv  APH, Phosphotransferase enzyme family. This family consists of bacterial antibiotic resistance proteins, which confer resistance to various                                                                                                                                                                                                                                                                                                                                                              |                 | 0.000152             | 25%        | 48%           |                  | 8 5150              |               | NP 048636 Similar                                                                                           | IZ: to PBCV-1 ORF A34R, corresponds to GenBank Accession er U17055                                                                                                                                                                                                     |                         | 1.73E-77                                                                         | 36%                                           | 54%                                           | 1-490                                                       |                                         |
|                              |                |               |       | 4                 | pfam01636    | aminoglycosides they include: aminoglycoside 3':<br>phosphotransferase or knampyinch knaser / neonypicn-haramycin<br>phosphotransferase and steptomycin 3', ', -kinase or<br>steptomycin 3', ', -kinase or<br>steptomycin 3', ', -kinase or<br>steptomycin 3', ', -kinase or<br>steptomycin 3', -kinase or<br>phosphorylation. This family also includes homoserine kinase. This family<br>is related in furchosamine kinase andm3381                                                                                                                                                                                                                                 | 40.52           | 0.000167             | 37%        | 70%           | 20723            | 3 171-198           | 4             | CAB10894 extens                                                                                             | sin                                                                                                                                                                                                                                                                    | 82.80                   | 3.54E-14                                                                         | 33%                                           | 43%                                           | 371502                                                      | 104-235                                 |
|                              |                |               |       | 5                 | pfam00069    | Pkinase, Protein kinase domain     COG2334, Putative homoserine kinase type II (protein kinase fold)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | 39.50<br>38.02  | 0.000343             | 21%<br>36% | 38%<br>55%    |                  | 1 5257<br>0 200233  |               |                                                                                                             | filament triplet H1-like protein<br>ICTED: similar to p87, partial                                                                                                                                                                                                     |                         | 7.39E-12<br>1.65E-11                                                             | 36%<br>24%                                    | 51%<br>46%                                    | 306506<br>366515                                            |                                         |
|                              |                |               |       | 7                 | COG542       | ROM1, RhoGEF, Guanine nucleotide exchange factor for Rho/Rac/Cdc42-like GTPases (Signal transduction mechanisms).                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |                 | 0.001145             | 18%        | 34%           |                  | 8 19197             |               |                                                                                                             | ICTED: similar to neurofilament, heavy polypeptide 200kDa isoform                                                                                                                                                                                                      |                         | 2.15E-11                                                                         | 32%                                           | 51%                                           | 327499                                                      |                                         |
|                              |                |               |       | 8                 |              | TEL1, Phosphatidylinositol kinase and protein kinases of the PI-3 kinase<br>jamily [Signal transduction mechanisms / Cell division and chromosome<br>partitioning / Chromatin structure and dynamics / DNA replication,<br>recombination, and repair / Intracellular trafficking and secretion].                                                                                                                                                                                                                                                                                                                                                                      | 36.21           | 0.002982             | 44%        | 50%           | 20924            | 3 }471981           |               | 2                                                                                                           | ICTED: similar to neurofilament, heavy polypeptide 200kDa isoform                                                                                                                                                                                                      | 73.17                   | 2.81E-11                                                                         | 32%                                           | 51%                                           | 327499                                                      | 634820                                  |
|                              |                |               |       | 9                 |              | AarF, Predicted unusual protein kinase [General function prediction only].                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 36.08           | 0.003979             | 29%        | 57%           |                  | 4 284-333           |               |                                                                                                             | ICTED: similar to neurofilament, heavy polypeptide 200kDa isoform                                                                                                                                                                                                      |                         | 2.81E-11                                                                         | 32%                                           | 51%                                           | 327499                                                      |                                         |
| M734L 2741052738             | 66 8           | 0 9.277       | 9.60  | 10                | No Hit Found | Pl3 Pl4 kinase, Phosphatidylinositol 3- and 4-kinase                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | 35.72           | 0.004574             | 32%        | 40%           | 20925            | 3 137184            | 10            | No Hit Found No Hit                                                                                         | netical protein, conserved                                                                                                                                                                                                                                             | 73.17                   | 2.81E-11                                                                         | 28%                                           | 46%                                           | 306499                                                      | 141351                                  |
| M735L 2752682748             |                |               | 3.97  |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |                 |                      |            |               |                  |                     |               | NP_048974 A618L                                                                                             |                                                                                                                                                                                                                                                                        | 80.49                   | 1.58E-14                                                                         | 55%                                           | 80%                                           | 82151                                                       | 63-125                                  |
| M738L 2759022752             | 31 22          | 4 25,232      | 4.23  | 3                 | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |                 |                      |            |               |                  |                     |               | No Hit Found No Hit                                                                                         | Found                                                                                                                                                                                                                                                                  |                         |                                                                                  |                                               |                                               |                                                             |                                         |
| M740L 2762472759             | 57 9           | 7 11,140      | 8.64  |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |                 |                      |            |               |                  |                     | 1             | NP_048976 similar<br>Numbi<br>NP_048991 A635F                                                               | to Synechocystis orf 90, corresponds to GenBank Accession er D90902                                                                                                                                                                                                    |                         | 1.85E-10<br>8.06E-06                                                             | 34%<br>30%                                    | 50%<br>51%                                    | 195<br>191                                                  | 181<br>182                              |
| M741L 2766822763             | 35 116         | 12,705        | 8.65  | 5                 | No Hit Found | I                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |                 |                      |            |               |                  |                     | 1             | NP 048977 A621L                                                                                             |                                                                                                                                                                                                                                                                        | 130.18                  | 1.78E-29                                                                         | 52%                                           | 71%                                           | 4116                                                        | 5117                                    |
| M742R 2767532794             | 55 90          | 1 99,481      | 6.19  | 1                 | COG048       | Uup, ATPase components of ABC transporters with duplicated ATPase domains [General function prediction only].     ABC, tran, ABC transporter. ABC transporters for a large family of proteins responsible for translocation of a variety of compounds across proteins responsible for translocation of a variety of compounds across.                                                                                                                                                                                                                                                                                                                                 | 277.14          | 9.58E-76             | 27%        | 45%           | 30084            | 8 2530              | 1             | NP_049022 Chlore<br>GenBa                                                                                   | ella virus CVK2 translation elongation factor-3 homolog, refer to ank Accession Number D16505                                                                                                                                                                          | 1105.51                 | 0.00E+00                                                                         | 65%                                           | 77%                                           | 37901                                                       | 55–918                                  |
|                              |                |               |       | 2                 | pfam0000     | biological membranes. ABC transporters are the largest family of proteins<br>5 in many completely sequenced bacteria. ABC transporters are composed<br>of two copies of this domain and two copies of a transmembrane domain<br>pfam00664. These four domains may belong to a single polypeptide, or<br>belanar in different colvenetitie chains.                                                                                                                                                                                                                                                                                                                     | 117.70          | 1.07E-27             | 31%        | 48%           | 32749            | 1 1-182             | 2             | A48779 transla                                                                                              | ation elongation factor EF-3 homolog - Chlorella virus CVK2                                                                                                                                                                                                            | 1010.36                 | 0.00E+00                                                                         | 61%                                           | 74%                                           | 37885                                                       | 2721120                                 |
|                              |                |               |       | 3                 |              | ABC_ATPase, ABC_(ATP-binding cassette) transporter nucleotide-<br>binding domain; ABC transporters are a large family of proteins involved<br>in the transport of a wide variety of different compounds. Ilse sugars,<br>and the summary of the summary of the summary of the summary of the<br>binding domain shows the highest smillarly between all members of the<br>family. ABC transporters are a subset of nucleotide hydrolases that<br>contain a signature motif. C-loop, and H-loopswitch region in addition to<br>the Walter A mottip-loop and/Walker B motif commonly found in a<br>number of ATP—and GTP-binding and hydrolyzing protein in<br>addition. | 105.67          | 4.18E-24             | 30%        | 51%           | 30647            | 6 5194              | 3             | CAG58023 unnam                                                                                              | ned protein product                                                                                                                                                                                                                                                    | 761.91                  | 0.00E+00                                                                         | 45%                                           | 63%                                           | 3900                                                        | 1131043                                 |
|                              |                |               |       | 4                 |              | ZnuC, ABC-type Mn/Zn transport systems, ATPase component     Illnomanic ion transport and metabolism!     CcmA, ABC-type multidrug transport system, ATPase component                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |                 | 1.07E-22             | 31%        | 51%           |                  | 3 9216              |               |                                                                                                             | ation elongation factor3                                                                                                                                                                                                                                               |                         | 0.00E+00                                                                         | 45%                                           | 62%                                           |                                                             | 1131043                                 |
|                              |                |               |       | 5                 | COG113       | IDefense mechanisms1.<br>  SunT, ABC-type bacteriocin/lantibiotic exporters, contain an N-terminal                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 100.43<br>94.52 | 1.58E-22<br>1.02E-20 | 28%<br>28% | 47%<br>46%    |                  | 7 16208<br>4 476666 |               |                                                                                                             | ation elongation factor 3                                                                                                                                                                                                                                              |                         | 0.00E+00<br>0.00E+00                                                             | 46%<br>45%                                    | 62%<br>62%                                    |                                                             | 1361050<br>1361050                      |
|                              |                |               |       | 7                 |              | double-olvcine peolidase domain (Defense mechanisms).  DppF, ABC-type dipeptide/oligopeptide/nickel transport system, ATPase 4 component [Amino acid transport and metabolism / Inorganic ion transport and metabolism).                                                                                                                                                                                                                                                                                                                                                                                                                                              |                 | 1.44E-20             | 29%        | 50%           |                  | 7 8213              |               |                                                                                                             | ation elongation factor 3                                                                                                                                                                                                                                              |                         | 0.00E+00                                                                         | 46%                                           | 62%                                           |                                                             | 1361050                                 |
|                              |                |               |       | 8                 | COG498       | CydD, ABC-type transport system involved in cytochrome bd<br>biosynthesis, ATPase and permesse components [Energy production<br>and conversion / Posttranslational modification, protein turnover,<br>chaperonesi.                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 91.04           | 1.01E-19             | 29%        | 44%           | 26247            | 5 279–514           | 8             | CAA22654 SPCC                                                                                               | 417.08                                                                                                                                                                                                                                                                 | 752.67                  | 0.00E+00                                                                         | 45%                                           | 63%                                           | 35898                                                       | 1421046                                 |

| Gene<br>Name | Genome<br>Position | A.A.<br>length | Peptide<br>Mw | pl    | CDD Hit<br>Number | COGs                 | COG Definition                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | Bit<br>Score     | E-value        | %<br>Identity | %<br>Positive | Query<br>from-to | Hit<br>from-to | BLASTp<br>Hit                   | Hit<br>Accession B                                                                                                                                                                                                                                                                                                                                                                                    | LASTp Definition                                                                            | Bit<br>Score                                                       | E-value                                                                                                           | %<br>Identity P                                             |                                                      | Query<br>from-to                                                    | Hit from-<br>to                                                                    |
|--------------|--------------------|----------------|---------------|-------|-------------------|----------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------|----------------|---------------|---------------|------------------|----------------|---------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------|--------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------|------------------------------------------------------|---------------------------------------------------------------------|------------------------------------------------------------------------------------|
| Name         | FOSILION           | leligtii       | INIV          |       | Number<br>9       | COG1122              | CbiO, ABC-type cobalt transport system, ATPase component [Inorganic                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 86.87            | 1.88E-18       | 29%           | 47%           | 305492           |                | Number<br>9                     |                                                                                                                                                                                                                                                                                                                                                                                                       |                                                                                             |                                                                    | 0.00E+00                                                                                                          | 45%                                                         | 62%                                                  |                                                                     | 1361049                                                                            |
|              |                    |                |               |       | 10                |                      | ion transport and metabolismi. Ttg2A, ABC-type transport system involved in resistance to organic solvents, ATPase component [Secondary metabolites biosynthesis,                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | 86.76            | 2.09E-18       | 26%           | 46%           | 306487           | 13217          | 10                              | CAG98340 unnamed protein product                                                                                                                                                                                                                                                                                                                                                                      |                                                                                             | 751.51                                                             | 0.00E+00                                                                                                          | 46%                                                         | 62%                                                  | 35901                                                               | 1371044                                                                            |
| M748L        | 281025279463       | 521            | 57,929        | 5.26  | i 1               | pfam04451            | transcort. and catabolism.  Capsid_Iridovir, Iridovirus and capsid protein. This family includes the major capsid protein of iridoviruses, chlorella virus and Spodoptera ascovirus, which are all dsDNA viruses with no RNA stage. This is the most abundant structural protein and can account for up to 45% of virion                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 365.79           | 1.79E-102      | 41%           | 53%           | 172517           | 75443          | 1                               | AAC27493 putative capsid protein                                                                                                                                                                                                                                                                                                                                                                      |                                                                                             | 1039.25                                                            | 0.00E+00                                                                                                          | 97%                                                         | 97%                                                  | 1-521                                                               | 1521                                                                               |
|              |                    |                |               |       |                   |                      | protein. In Chlorella virus MT325 the major capsid protein is a alvooprotein                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |                  |                |               |               |                  |                |                                 | ainsilas ta Cimulium iridaasa                                                                                                                                                                                                                                                                                                                                                                         | ant views associal associais associates de Corina                                           |                                                                    |                                                                                                                   |                                                             |                                                      |                                                                     |                                                                                    |
|              |                    |                |               |       |                   |                      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                  |                |               |               |                  |                | 2                               | NP_048978 Prot Accession Number P2 BAE06835 hypothetical major capsid p                                                                                                                                                                                                                                                                                                                               | ent virus capsid protein, corresponds to Swiss-<br>12166<br>protein                         |                                                                    | 0.00E+00<br>2.10E-63                                                                                              | 70%<br>35%                                                  | 80%<br>52%                                           | 1521<br>126521                                                      | 1520<br>34440                                                                      |
|              |                    |                |               |       |                   |                      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                  |                |               |               |                  |                | 4<br>5                          | BAA76600 major capsid protein<br>AAC27492 major capsid protein Vp49                                                                                                                                                                                                                                                                                                                                   |                                                                                             | 197.98<br>195.67                                                   | 6.56E-49<br>3.26E-48                                                                                              | 34%<br>32%                                                  | 51%<br>50%                                           | 196521<br>196521                                                    | 92-436<br>89-432                                                                   |
|              |                    |                |               |       |                   |                      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                  |                |               |               |                  |                | 6                               | BAA76601 major capsid protein MCP1<br>NP 048787 PBCV-1 major capsid protein                                                                                                                                                                                                                                                                                                                           | in Vp54, corresponds to GenBank Accession                                                   | 194.13<br>194.13                                                   |                                                                                                                   | 34%<br>34%                                                  | 51%<br>51%                                           | 196521<br>196521                                                    | 92-437<br>92-437                                                                   |
|              |                    |                |               |       |                   |                      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                  |                |               |               |                  |                | 8                               |                                                                                                                                                                                                                                                                                                                                                                                                       | Of Major Capsid Protein Of A Large, Lipid                                                   | 194.13                                                             |                                                                                                                   | 34%                                                         | 51%                                                  | 196521                                                              | 68-413                                                                             |
|              |                    |                |               |       |                   |                      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                  |                |               |               |                  |                | 9<br>10                         | 1M4X_C Chain C, Pbcv-1 Virus Cap<br>BAA22198 major capsid protein Vp54                                                                                                                                                                                                                                                                                                                                | sid, Quasi-Atomic Model                                                                     | 193.74<br>190.66                                                   |                                                                                                                   | 34%<br>34%                                                  | 51%<br>51%                                           | 196521<br>196521                                                    | 68413<br>92437                                                                     |
| M752R        | 281297281653       | 119            | 13,504        | 9.75  | 1                 | COG4852              | COG4852, Predicted membrane protein [Function unknown].                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | 37.20            | 0.001517       | 24%           | 36%           | 8117             | 10124          | 1 2                             | NP_048980 A624R<br>ZP_00843145 conserved hypothetical pro                                                                                                                                                                                                                                                                                                                                             | tein                                                                                        | 130.95<br>53.91                                                    | 1.03E-29<br>1.60E-06                                                                                              | 52%<br>30%                                                  | 72%<br>49%                                           | 5117<br>6108                                                        | 8120<br>9115                                                                       |
|              |                    |                |               |       |                   |                      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                  |                |               |               |                  |                | 3                               | ZP 00653117 conserved hypothetical pro<br>ZP 00234461 conserved hypothetical pro                                                                                                                                                                                                                                                                                                                      | tein<br>tein                                                                                | 50.45<br>50.45                                                     | 1.77E-05                                                                                                          | 31%<br>24%                                                  | 46%<br>43%                                           | 6108<br>6119                                                        | 8117<br>9127                                                                       |
|              |                    |                |               |       |                   |                      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                  |                |               |               |                  |                | 5<br>6                          | AAZ18596 conserved hypothetical pro<br>CAC98668 Imo0589                                                                                                                                                                                                                                                                                                                                               | tein                                                                                        | 49.29<br>48.52                                                     | 3.95E-05<br>6.74E-05                                                                                              | 27%<br>24%                                                  | 44%<br>42%                                           | 6108<br>6119                                                        | 8117<br>9127                                                                       |
| M753L        | 282564281662       | 301            | 33,934        | 10.83 |                   | No Hit Found         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                  |                |               |               |                  |                | 1                               | NP_048439 a91L                                                                                                                                                                                                                                                                                                                                                                                        |                                                                                             | 64.70                                                              | 3.99E-09                                                                                                          | 34%                                                         | 53%                                                  | 177291                                                              | 2112                                                                               |
| M756R        | 282608283156       | 183            | 20,183        | 3.48  |                   | No Hit Found         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                  |                |               |               |                  |                | 1                               | NP 049015 A659L                                                                                                                                                                                                                                                                                                                                                                                       |                                                                                             | 80.49                                                              | 2.60E-14                                                                                                          | 38%                                                         | 49%                                                  | 1139                                                                | 1125                                                                               |
| M757R        | 283302283733       | 144            | 16,320        | 6.80  |                   | No Hit Found         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                  |                |               |               |                  |                | 1                               | NP_049012 A656L                                                                                                                                                                                                                                                                                                                                                                                       |                                                                                             | 65.08                                                              | 6.94E-10                                                                                                          | 38%                                                         | 56%                                                  | 992                                                                 | 8100                                                                               |
| M758R        | 283779284369       | 197            | 21,905        | 5.09  | 1                 | pfam00583            | Acetyltransf_1, Acetyltransferase (GNAT) family. This family contains proteins with N-acetyltransferase functions                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | 45.24            | 6.81E-06       | 31%           | 52%           | 113163           | 3082           | 1                               | NP_049010 A654L                                                                                                                                                                                                                                                                                                                                                                                       |                                                                                             | 209.15                                                             | 5.83E-53                                                                                                          | 51%                                                         | 68%                                                  | 1197                                                                | 1197                                                                               |
|              |                    |                |               |       | 2                 | COG3153              | COG3153, Predicted acetyltransferase [General function prediction only].                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 37.17            | 0.001578       | 40%           | 62%           | 113162           | 81126          | 2                               | ZP_00663133 GCN5-related N-acetyltran                                                                                                                                                                                                                                                                                                                                                                 | sferase                                                                                     | 65.08                                                              | 1.36E-09                                                                                                          | 27%                                                         | 47%                                                  | 2184                                                                | 11-205                                                                             |
|              |                    |                |               |       | 3                 | COG1247              | COG1247, Sortase and related acyltransferases [Cell envelope biogenesis. outer membranel. Riml, Acetyltransferases [General function prediction only].                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | 36.40<br>35.37   |                | 35%<br>25%    | 49%<br>41%    |                  | 89138<br>97168 |                                 | ZP_00675288 hypothetical protein TeryDI<br>XP 392876 PREDICTED: similar to any                                                                                                                                                                                                                                                                                                                        |                                                                                             | 62.39<br>56.23                                                     |                                                                                                                   | 24%<br>27%                                                  | 44%<br>44%                                           | 10186<br>2173                                                       | 15-204<br>51-237                                                                   |
|              |                    |                |               |       | *                 | C000450              | Rinii, Acetylitansierases (General function prediction only).                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | 35.37            | 0.005900       | 2376          | 4170          | 113102           | 97100          | 5                               |                                                                                                                                                                                                                                                                                                                                                                                                       |                                                                                             | 52.76                                                              |                                                                                                                   | 21%                                                         | 41%                                                  | 7-173                                                               | 41-225                                                                             |
|              | 285067284549       | 173            | 19,326        | 9.68  |                   | No Hit Found         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                  |                |               |               |                  |                | 1                               | NP_049000 A644R                                                                                                                                                                                                                                                                                                                                                                                       |                                                                                             | 117.09                                                             | 2.15E-25                                                                                                          | 72%                                                         | 79%                                                  | 89172                                                               | 284                                                                                |
| M763L        | 286537285098       | 480            | 54,551        | 11.33 |                   |                      | COG4942, Membrane-bound metallopeptidase [Cell division and chromosome partitioning] ToIA, Membrane protein involved in colicin uptake [Cell envelope                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 42.26            |                | 16%           | 32%           |                  | 175334         |                                 | NP_048999 Gln-rich; KQQ (6X)                                                                                                                                                                                                                                                                                                                                                                          |                                                                                             | 224.94                                                             |                                                                                                                   | 47%                                                         | 58%                                                  | 196480                                                              | 1269                                                                               |
|              |                    |                |               |       | 2                 | COG3064              | biogenesis. outer membranel.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | 38.23            | 0.00091        | 19%           | 40%           | 233380           | 52196          | 2                               | NP_048998 A642R                                                                                                                                                                                                                                                                                                                                                                                       |                                                                                             | 60.46                                                              | 1.47E-07                                                                                                          | 68%                                                         | 80%                                                  | 19-53                                                               | 2155                                                                               |
| M766L        | 287657286569       | 363            | 41,125        | 5.49  | 1                 | pfam04371            | PAD_porth, Porphyromonas-type peptidy/-arginine deiminase. Peptidy-<br>arginine deiminase (PAD) enzymes catalyse the deimination of the<br>quanding group from carboxy-terminal arginine residues of various<br>peptides to produce ammonia. PAD from Porphyromonas gringvisias<br>(paramonas) per period produce ammonia pada period produce (paramonas period produce)<br>(plann03083), which is a metallorinzyme. PPAD is thought to belong to<br>the same superfamily as aminiotranteries and arginine deiminase, and<br>to form an alpha/beta propeller structure. This family has previously been<br>named PPADI (Porphyromonas peptidy)-arginine deiminase, and<br>homologues). The predicted catalytic residues in PPAD are Asp 130,<br>which the exception of Aso 167 which is absent in No family members.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | 435.03           | 2.64E-123      | 54%           | 73%           | 14-361           | 1-329          | 1                               | NP_048994 PBCV-1 Agmatine iminohy                                                                                                                                                                                                                                                                                                                                                                     | drolase                                                                                     | 493.43                                                             | 4.63E-138                                                                                                         | 62%                                                         | 78%                                                  | 7–361                                                               | 1358                                                                               |
|              |                    |                |               |       |                   |                      | PPAD is also able to catalyse the deimination of free L-arginine, but has<br>primarily peptidyl-arginine specificity. It may have a FMN cofactor                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |                  |                |               |               |                  |                |                                 |                                                                                                                                                                                                                                                                                                                                                                                                       |                                                                                             |                                                                    |                                                                                                                   |                                                             |                                                      |                                                                     |                                                                                    |
|              |                    |                |               |       | 2                 | COG2957              | COG2957, Peptidylarginine deiminase and related enzymes [Amino acid transcort and metabolism].                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | 388.49           | 2.64E-109      | 46%           | 65%           | 9361             | 10344          | 3<br>4<br>5<br>6<br>7<br>8<br>9 | CAC9823 Imo0038 ZP_0022985 peptidyl-arginine deiminass ZP_00228718 peptidyl-arginine deiminass ZP_00222858 COG2957: Peptidylarginin ZP_00402290 Perpitymonass-type pepti ZP_00229957 conserved hypothetical pro ArT02237 conserved hypothetical pro ArT02237 conserved hypothetical pro ArT02337 conserved hypothetical pro ArT02337 conserved hypothetical pro ZP_0049916 conserved hypothetical pro | s-like protein<br>e deiminase and related enzymes<br>dyl-arginine deiminase<br>tein<br>tein | 382.10<br>381.72<br>378.25<br>377.48<br>377.48<br>376.71<br>374.02 | 3.96E-105<br>1.50E-104<br>1.97E-104<br>2.17E-103<br>3.71E-103<br>3.71E-103<br>6.32E-103<br>4.10E-102<br>6.99E-102 | 52%<br>51%<br>51%<br>52%<br>51%<br>51%<br>51%<br>51%<br>51% | 69%<br>68%<br>68%<br>68%<br>67%<br>66%<br>66%<br>67% | 11361<br>11361<br>9361<br>9-362<br>9-361<br>9-361<br>9-362<br>9-362 | 10-362<br>10-362<br>10-362<br>8-362<br>15-368<br>8-362<br>8-362<br>8-364<br>15-368 |
| M770R        | 287716288117       | 134            | 15,340        | 9.58  |                   | No Hit Found         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                  |                |               |               |                  |                | 1                               | NP_048990 A634L                                                                                                                                                                                                                                                                                                                                                                                       |                                                                                             | 166.78                                                             | 1.71E-40                                                                                                          | 58%                                                         | 77%                                                  | 2129                                                                | 4131                                                                               |
| M773L        | 288496288116       | 127            | 14,776        | 8.63  |                   | No Hit Found         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                  |                |               |               |                  |                |                                 | No Hit Found No Hit Found                                                                                                                                                                                                                                                                                                                                                                             |                                                                                             |                                                                    |                                                                                                                   |                                                             |                                                      |                                                                     |                                                                                    |
|              |                    |                |               |       |                   |                      | RNR_1, RNR. class I. Ribonucleotide reductase (RNR) catalyzes the<br>reductive synthesis of deoxyrbonucleotides from their corresponding<br>ribonucleotides. Il provides the precursors necessary for DNA synthesis.<br>RNRs are separated into three classes based on their metallocofactor<br>usage. Class I RNRs, found in eukaryotes, bacteria, and many viruses,<br>use a diron-typor diracla. (Class I RNRs, found in bacteria, and<br>bacteriophages, use coenzyme B12 (adenosylcobalamin, AdoCDI). Class<br>III RNRs, found in bacteriophages, and archeae, use                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                  |                |               |               |                  |                |                                 | similar to Schivosarchar                                                                                                                                                                                                                                                                                                                                                                              | umyces ribonucleotide reductase M1 chain,                                                   |                                                                    |                                                                                                                   |                                                             |                                                      |                                                                     |                                                                                    |
| M777L        | 291250288956       | 765            | 85,431        | 7.89  | 1                 | cd01679              | as FaS cluster and S-adenosylmethionine to generate a glycyl radical.<br>Many organisms have more than one class of RNR present in their<br>genomes. All three RNRs have a ten-stranded alpha-beta barrel domain<br>that is structurally similar to the domain of PEL (pryunate formate) tyase).<br>Class I RNR is oxygen-dependent and can be subdivided into classes is<br>(eukaryotes, prokaryotes viruses and phages) and Ib (which is found in<br>prokaryotes only). It is a tetrametic enzyme of two alpha and two beta<br>subvinite: this model rowers the amain and if the solither of stress subvinite or<br>submitted.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 733.93           | 0              | 51%           | 68%           | 176746           | 1578           | 1                               | NP_048985 corresponds to Swiss-Prot                                                                                                                                                                                                                                                                                                                                                                   | Accession Number P36602                                                                     | 1215.29                                                            | 0.00E+00                                                                                                          | 78%                                                         | 88%                                                  | 9-765                                                               | 14-771                                                                             |
|              |                    |                |               |       | 2                 | pfam02867<br>COG0209 | Ribonuc red InC. Ribonucleotide reductase, barrel domain. NrdA, Ribonucleotide reductase, alpha subunit (Nucleotide transport and matahalismi) RNR_1 like, RNR, class Like family. Ribonucleotide reductase (RNR) catalyzes the reductive synthesis of deoxyribonucleotides from their corresponding ribonucleotides. It provides the precursors necessary for DNA synthesis. RNNS are separated into three classes based on their                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 634.64<br>533.82 | 0<br>5.62E-153 | 47%<br>37%    | 66%<br>53%    |                  | 1532<br>7696   |                                 | IP_001026008 ribonucleoside-diphosphate<br>AAH46846 RRM1 protein                                                                                                                                                                                                                                                                                                                                      | e reductase M1 chain                                                                        |                                                                    | 0.00E+00<br>0.00E+00                                                                                              | 50%<br>50%                                                  | 69%<br>69%                                           | 9759<br>9759                                                        | 1754<br>1754                                                                       |
|              |                    |                |               |       | 4                 | cd02888              | metalicoclactor usage. Class I RNRs, found in eukaryotes, bacteria, and many viruses, use a dirion-typosy radical, Class II RNRs, found in bacteria, and bacteria, and bacteriophages, use coenzyme B12 (adenosylcobalamin, AdoCb). Class II RNRs, found in anemorbic bacteria, bacteriophages, and acrbaea, use an FeS duster and S-adenosylmethionine to generate a glycyl redical. Many organisms have more than one class of RNR. Alternative of the common that is serious and common that it is seriou | 303.28           | 1.50E-83       | 34%           | 51%           | 196743           | 2-521          | 4                               | AAH74185 RRM1 protein                                                                                                                                                                                                                                                                                                                                                                                 |                                                                                             | 756.90                                                             | 0.00E+00                                                                                                          | 50%                                                         | 69%                                                  | 9759                                                                | 1754                                                                               |
|              |                    |                |               |       | 5                 | pfam00317            | sequence similarity and the predicted active site  Ribonuc red IqN, Ribonucleotide reductase, all-alpha domain                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | 77.59            | 1.24E-15       | 40%           | 56%           | 150222           | 1-78           | 5                               | CAB98233 ribonucleoside-diphosphate                                                                                                                                                                                                                                                                                                                                                                   | e reductase large chain (un-24)                                                             | 751.13                                                             | 0.00E+00                                                                                                          | 50%                                                         | 67%                                                  | 9759                                                                | 1755                                                                               |

| Gene<br>Name | Genome<br>Position           | A.A.<br>length | Peptide<br>Mw | pl         | CDD Hit<br>Number | COGs         | COG Definition                                                                                                                                                                                                                                                                                                                                               | Bit<br>Score            | E-value                          | %<br>Identity     | %<br>Positive     | Query<br>from-to | Hit<br>from-to       | BLASTp<br>Hit | Hit<br>Accession              | BLASTp Definition                                                                                                                      | Bit<br>Score              | E-value Ic                       | %<br>dentity Po   |                   | Query<br>from-to          | Hit from-                  |
|--------------|------------------------------|----------------|---------------|------------|-------------------|--------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------|----------------------------------|-------------------|-------------------|------------------|----------------------|---------------|-------------------------------|----------------------------------------------------------------------------------------------------------------------------------------|---------------------------|----------------------------------|-------------------|-------------------|---------------------------|----------------------------|
|              |                              |                |               |            |                   |              | RNR_PFL, RNR_PFL. Ribonucleotide reductase (RNR) and pyruvate formate lyase (PFL) have a structurally similar ten-stranded alpha-beta                                                                                                                                                                                                                        |                         |                                  |                   |                   |                  |                      | Number        |                               |                                                                                                                                        |                           |                                  |                   |                   |                           |                            |
|              |                              |                |               |            | 6                 | cd00576      | barrel active site domain and are believed to have diverged from a<br>common ancestor. RNRs are found in all organisms and provide the only<br>mechanism by which nucleotides are converted to deoxynucleotides,<br>while PFL, an essential enzyme in anaerobic bacteria, catalyzes the<br>conversion of pryruvate and CoA to acteVCOA and formate. Both RNR | 55.88                   | 3.78E-09                         | 22%               | 35%               | 249678           | 70461                | 6             | AAD4974                       | 3 ribonucleotide reductase large subunit                                                                                               | 751.13                    | 0.00E+00                         | 50%               | 67%               | 9759                      | 1755                       |
|              |                              |                |               |            | 7                 | pfam03477    | and PEI are glycyl radical enzymes ATP-cone, ATP cone domain                                                                                                                                                                                                                                                                                                 | 49.59                   | 2.95E-07                         | 24%               | 40%               | 9110             | 1-89                 | 7             |                               | Ribonucleotide reductase M1                                                                                                            |                           | 0.00E+00                         | 49%               | 68%               | 9759<br>9759              | 1754                       |
|              |                              |                |               |            |                   |              |                                                                                                                                                                                                                                                                                                                                                              |                         |                                  |                   |                   |                  |                      | 9<br>10       | NP_00102<br>AAA4006           | Ribonucleotide reductase M1     fibonucleoside-diphosphate reductase M1 chain     ribonucleotide reductase subunit M1                  | 743.81                    | 0.00E+00<br>0.00E+00<br>0.00E+00 | 49%<br>49%<br>49% | 68%<br>68%<br>68% | 9759<br>9759<br>9759      | 1754<br>1754<br>1754       |
| M785L        | 292819291521                 | 433            | 48,690        | 10.        | 98                | No Hit Found |                                                                                                                                                                                                                                                                                                                                                              |                         |                                  |                   |                   |                  |                      | 1 2           | NP_04857                      | 3 similar to Chlorella virus PBCV-1 ORF A231L, corresponds to GenBank<br>Accession Number U42590<br>9 contains ATP/GTP-binding motif A |                           | 1.41E-131<br>9.62E-80            | 54%<br>44%        | 73%<br>62%        | 25433<br>94423            | 31436<br>4337              |
| M789L        | 293169292840                 | 110            | 12,094        | 9.         | 36 1              |              | COG1293, Predicted RNA-binding protein homologous to eukaryotic snRNP ITranscrintinal DUF814, Domain of unknown function (DUF814). This domain occurs in proteins that have been annotated as Fibronectin/fibrinogen binding                                                                                                                                 | 78.49                   | 6.81E-16                         | 37%               | 57%               | 4110             | 445554               | 1             | NP_04852                      | 3 similar to Streptococcus pyogenes fibronectin protein, corresponds to<br>GenRank Accession Number I 28919                            | 117.86                    | 9.04E-26                         | 52%               | 71%               | 1106                      | 1105                       |
|              |                              |                |               |            | 2                 | pfam05670    | protein by similarity. This annotation comes from a sequence, where the N-terminal region is involved in this activity. Hence the activity of this C-terminal domain is unknown. This domain contains a conserved motif D/E X-W/Y-X-H that may be functionally important.                                                                                    | 72.55                   | 4.05E-14                         | 39%               | 60%               | 487              | 3-90                 |               |                               | Protein of unknown function DUF814:Fibronectin-binding A, N-terminal                                                                   |                           | 2.02E-09                         | 34%               | 50%               | 5109                      | 469-576                    |
|              |                              |                |               |            |                   |              |                                                                                                                                                                                                                                                                                                                                                              |                         |                                  |                   |                   |                  |                      | 3             | AAD3545<br>CAA0886<br>CAB1343 | Difibronectin-binding protein, putative 3 putative fibronectin-binding protein                                                         | 60.46<br>59.69<br>58.54   | 1.71E-08<br>2.92E-08<br>6.51E-08 | 32%<br>32%<br>30% | 53%<br>50%<br>50% | 6106<br>5106<br>5110      | 440542<br>439542<br>457563 |
|              |                              |                |               |            |                   |              |                                                                                                                                                                                                                                                                                                                                                              |                         |                                  |                   |                   |                  |                      | 6 7           | YP_00421                      | o yloA<br>3 fibronectin/fibrinogen-binding protein<br>1 probable RNA-biniding protein                                                  | 57.77<br>57.77            | 1.11E-07<br>1.11E-07             | 29%<br>29%        | 51%<br>51%        | 8106<br>8106              | 405-505<br>405-505         |
|              |                              |                |               |            |                   |              |                                                                                                                                                                                                                                                                                                                                                              |                         |                                  |                   |                   |                  |                      | 8             | NP_78221<br>AAK8006           | 3 fibronectin/fibrinogen-binding protein<br>9 Fibronectin-binding protein                                                              | 57.38<br>57.38            | 1.45E-07<br>1.45E-07             | 34%<br>32%        | 48%<br>50%        | 5106<br>5104              | 463-566<br>460-561         |
| 147040       |                              | 500            | 50.004        | -          |                   |              |                                                                                                                                                                                                                                                                                                                                                              | 00.00                   | 4.055.40                         | 0001              | 400/              | 40.00            | 4.05                 | 10            |                               | RNA-binding protein homologous to eukaryotic snRNP                                                                                     | 57.38                     | 1.45E-07                         | 33%               | 49%               | 5106                      | 468572                     |
| M791R        | 293278294837                 | 520            | 56,961        | 7.:        | 33 1<br>2<br>3    | pfam00704    | CBD II, CBD II domain Glyco hydro 18, Glycosyl hydrolases family 18 COG3469, Chitinase [Carbohydrate transport and metabolism].                                                                                                                                                                                                                              | 66.89<br>38.19<br>34.63 | 1.85E-12<br>0.000916<br>0.009514 | 30%<br>18%<br>33% | 43%<br>31%<br>47% | 1293<br>117301   | 485<br>8187<br>90169 |               |                               | 4 vChti-1<br>9 PBCV-1 chitinase<br>5 Cellulose-bindina, bacterial type                                                                 | 729.17                    | 0.00E+00<br>0.00E+00<br>2.93E-57 | 69%<br>68%<br>40% | 76%<br>75%<br>57% | 1-513<br>1-513<br>111-412 | 1-533<br>1-537<br>225-538  |
|              |                              |                |               |            |                   | 0003408      | COCONOS, Cilitilase [Calibonydiate transport and metabolism].                                                                                                                                                                                                                                                                                                | 34.03                   | 0.000014                         | 3370              | 47.70             | 178238           | 30103                | 4             | BAC7296                       | 4 putative sugar hydrolase<br>3 putative secreted sugar hydrolase                                                                      |                           | 2.57E-53<br>2.57E-53             | 32%<br>31%        | 46%<br>46%        | 14412                     | 46-487<br>45-483           |
|              |                              |                |               |            |                   |              |                                                                                                                                                                                                                                                                                                                                                              |                         |                                  |                   |                   |                  |                      |               | ZP_0068741                    | 4 Chitinase<br>7 Carbohydrate-binding domain, family V/XII                                                                             | 183.73                    | 1.28E-44<br>2.04E-42             | 36%<br>36%        | 50%<br>49%        | 118422<br>118422          | 26-350<br>26-350           |
|              |                              |                |               |            |                   |              |                                                                                                                                                                                                                                                                                                                                                              |                         |                                  |                   |                   |                  |                      | 8             | ABB0841                       | 3 Chitinase<br>2 putative exported chitinase                                                                                           | 175.25<br>174.87          | 4.54E-42<br>5.93E-42             | 35%<br>37%        | 49%<br>51%        | 118422<br>125409          | 26-350<br>37-339           |
|              |                              |                |               |            |                   |              |                                                                                                                                                                                                                                                                                                                                                              |                         |                                  |                   |                   |                  |                      | 10            | ZP_0047957                    | 9 COG3979: Uncharacterized protein contain chitin-binding domain type 3                                                                | 174.10                    | 1.01E-41                         | 37%               | 51%               | 125409                    | 33-335                     |
|              |                              |                |               |            |                   |              | S_TKc, Serine/Threonine protein kinases, catalytic domain.  Phosphotransferases of the serine or threonine-specific kinase subfamily.                                                                                                                                                                                                                        |                         |                                  |                   |                   |                  |                      |               |                               |                                                                                                                                        |                           |                                  |                   |                   |                           |                            |
| M794R        | 294905295714                 | 270            | 30,668        | 6.         | 52 1              | cd00180      | The enzymatic activity of these protein kinases is controlled by<br>phosphorylation of specific residues in the activation segment of the<br>catalytic domain, sometimes combined with reversible conformational<br>changes in the C-terminal autoregulatory tail                                                                                            | 190.03                  | 1.42E-49                         | 29%               | 53%               | 10-268           | 1256                 | 1             | AAU0628                       | D protein kinase A248R                                                                                                                 | 240.74                    | 3.37E-62                         | 48%               | 66%               | 4267                      | 41–304                     |
|              |                              |                |               |            | 2                 |              | Phosphotransferases Serine or threonine-specific kinase subfamily                                                                                                                                                                                                                                                                                            | 187.34                  | 9.57E-49                         | 31%               | 49%               | 11268            |                      | 2             |                               | 5 protein kinase A248R                                                                                                                 |                           | 3.37E-62                         | 48%               | 66%               | 4267                      | 20-283                     |
|              |                              |                |               |            | 3                 |              | Pkinase, Protein kinase domain<br>SPS1, Serine/threonine protein kinase [General function prediction only /<br>Signal transduction mechanisms / Transcription / DNA replication,                                                                                                                                                                             | 183.56                  | 1.48E-47<br>2.00E-27             | 32%<br>27%        | 53%<br>47%        | 11-268           |                      | 3             |                               | 2 protein kinase A248R<br>7 PBCV-1 protein kinase                                                                                      |                           | 4.87E-61<br>6.58E-58             | 46%<br>43%        | 65%<br>64%        | 4267<br>4267              | 41–304<br>38–303           |
|              |                              |                |               |            |                   |              | recombination, and repairl.                                                                                                                                                                                                                                                                                                                                  |                         | 9.29E-26                         |                   | 47%               |                  |                      |               |                               |                                                                                                                                        |                           | 9.50E-57                         | 43%               | 64%               |                           | 34-299                     |
|              |                              |                |               |            | 5                 |              | Tyrosine-specific kinase subfamily.  TyrKc, Tyrosine kinase, catalytic domain. Phosphotransferases; tyrosine-specific kinase subfamily. Enzymes with TyrKc domains belong to an extensive family of proteins which share a conserved catalytic core.                                                                                                         | 111.04                  | 9.29E-26                         | 26%               | 44%               | 12264            | 2-253                | 5             | AAA8706                       | 5 serine/threonine protein kinase                                                                                                      | 222.63                    | 9.50E-57                         | 43%               | 64%               | 4267                      | 34-299                     |
|              |                              |                |               |            | 6                 |              | common to both serine/threonine and tyrosine protein kinases. Enzymatic activity of tyrosine protein kinases is controlled by phosphorylation of specific tyrosine residues in the activation segment of the catalytic domain or a C-terminal tyrosine (tail) residue with reversible conformational channes                                                 | 103.35                  | 1.82E-23                         | 25%               | 43%               | 9263             | 7263                 | 6             | AAU0627                       | 4 protein kinase A248R                                                                                                                 | 201.83                    | 1.73E-50                         | 49%               | 66%               | 52267                     | 22-237                     |
|              |                              |                |               |            | 7                 | COG3642      | COG3642, Mn2+-dependent serine/threonine protein kinase [Signal transduction mechanisms].                                                                                                                                                                                                                                                                    | 35.98                   | 0.003518                         | 28%               | 44%               | 86152            | 67139                | 7             |                               | 5 protein kinase A248R                                                                                                                 | 192.97                    | 8.06E-48                         | 47%               | 64%               | 54267                     | 2216                       |
|              |                              |                |               |            |                   |              |                                                                                                                                                                                                                                                                                                                                                              |                         |                                  |                   |                   |                  |                      | 8<br>9        | ΔΔI I0627                     | 6 protein kinase A248R<br>D protein kinase A248R                                                                                       | 154.84<br>139.43          | 2.43E-36<br>1.06E-31             | 48%<br>45%        | 70%<br>67%        | 119267<br>124267          | 1151<br>1146               |
|              |                              |                |               |            |                   |              |                                                                                                                                                                                                                                                                                                                                                              |                         |                                  |                   |                   |                  |                      | 10            | _                             | 3 similar to PBCV-1 serine/threonine protein kinase, corresponds to GenBank Accession Number U14660                                    | 132.88                    | 9.90E-30                         | 31%               | 53%               | 9267                      | 16-278                     |
|              | 295768296901<br>297080297832 | 378<br>251     |               | 8.t<br>7.t |                   | No Hit Found |                                                                                                                                                                                                                                                                                                                                                              |                         |                                  |                   |                   |                  |                      | 1 2           | CAA6497                       | 4 QI74 protein                                                                                                                         | 133.65<br>60.08<br>239.20 | 9.83E-30<br>1.38E-07<br>8.61E-62 | 55%<br>22%<br>63% | 69%<br>33%<br>81% | 178281<br>130378<br>80248 | 69173<br>344580<br>3171    |
| moo n        | 207000 207002                | 201            | 20,000        |            | ~                 | NO THE FOUND |                                                                                                                                                                                                                                                                                                                                                              |                         |                                  |                   |                   |                  |                      | 2             | NP_04900                      | similar to Chlorella virus PBCV-1 ORF A450R, corresponds to GenBank                                                                    | 120.55                    | 4.46E-26                         | 31%               | 48%               | 1-250                     | 4254                       |
|              |                              |                |               |            |                   |              |                                                                                                                                                                                                                                                                                                                                                              |                         |                                  |                   |                   |                  |                      |               |                               | similar to PBCV-1 ORF A275R, encoded by GenBank Accession Number<br>1142580                                                            | 115.55                    | 1.44E-24                         | 29%               | 48%               | 4250                      | 1247                       |
|              |                              |                |               |            |                   |              |                                                                                                                                                                                                                                                                                                                                                              |                         |                                  |                   |                   |                  |                      | 4<br>5        | NP_04852<br>NP_04862          | similar to PBCV-1 ORF A79R, corresponds to GenBank Accession                                                                           |                           | 3.91E-22<br>8.71E-22             | 29%<br>27%        | 48%<br>47%        | 1234<br>1248              | 1227<br>1248               |
|              |                              |                |               |            |                   |              |                                                                                                                                                                                                                                                                                                                                                              |                         |                                  |                   |                   |                  |                      | 6             | NP 04842                      |                                                                                                                                        | 98.60                     | 1.82E-19                         | 30%               | 45%               | 4227                      | 1217                       |
|              |                              |                |               |            |                   |              |                                                                                                                                                                                                                                                                                                                                                              |                         |                                  |                   |                   |                  |                      | 7             | AAU0630                       | 4 hypothetical protein A275R                                                                                                           | 62.77<br>62.39            | 1.10E-08<br>1.44E-08             | 47%<br>25%        | 65%<br>46%        | 464<br>82248              | 158<br>1167                |
|              |                              |                |               |            |                   |              |                                                                                                                                                                                                                                                                                                                                                              |                         |                                  |                   |                   |                  |                      | 9<br>10       | AAU0630                       | 1 hypothetical protein A275R<br>2 hypothetical protein A275R                                                                           | 61.23<br>51.22            | 3.21E-08<br>3.33E-05             | 24%<br>24%        | 46%<br>48%        | 82248<br>117250           | 1167<br>1133               |
|              |                              |                |               |            |                   |              | Trypan_PARP, Procyclic acidic repetitive protein (PARP). This family consists of several Trypanosoma brucei procyclic acidic repetitive protein                                                                                                                                                                                                              |                         |                                  |                   |                   |                  |                      |               |                               |                                                                                                                                        |                           |                                  |                   |                   |                           |                            |
| M803L        | 299338297839                 | 500            | 53,361        | 10.        | 75 1              | pfam05887    | (PARP) like sequences. The procyclic acidic repetitive protein (parp)<br>genes of Trypanosoma brucei encode a small family of abundant surface<br>proteins whose expression is restricted to the procyclic form of the<br>parasite. They are found at two unlinked loci, para and paraB:                                                                     | 52.67                   | 3.85E-08                         | 34%               | 48%               | 123185           | 60122                | 1             | NP_04838                      | o contains Pro-rich Px motif, PAPK (8X); similar to Thermoproteus virus protein TPX, corresponds to Swiss-Prot Accession Number P19275 | 427.56                    | 4.85E-118                        | 65%               | 80%               | 186499                    | 99-410                     |
|              |                              |                |               |            | 2                 | COG0810      | transcrintion of hoth loci is developmentally regulated.  TonB, Periplasmic protein TonB, links inner and outer membranes [Cell envelope biogenesis, outer membrane].                                                                                                                                                                                        | 48.22                   | 7.13E-07                         | 27%               | 34%               | 98196            | 39138                | 2             | NP_04851                      | g similar to PBCV-1 ORF A41R, corresponds to GenBank Accession Number U17055                                                           | 410.22                    | 8.01E-113                        | 63%               | 79%               | 189499                    | 77-385                     |
|              |                              |                |               |            | 3                 | pfam05616    | Materials Teach Materials and Santa Teach and Santa Santa                                                                                                                                                                                                                                                                                                    | 47.43                   | 1.30E-06                         | 31%               | 35%               | 117191           | 329403               | 3             | BAD8696                       | B hypothetical protein                                                                                                                 | 67.01                     | 1.66E-09                         | 23%               | 41%               | 194476                    | 95-375                     |
|              |                              |                |               |            | 4                 | COG3147      | DedD, Uncharacterized protein conserved in bacteria [Function                                                                                                                                                                                                                                                                                                | 43.09                   | 2.99E-05                         | 33%               | 39%               | 99187            | 63151                | 4             |                               | 5 P0481E12.18                                                                                                                          | 67.01                     | 1.66E-09                         | 23%               | 41%               | 194476                    | 72-352                     |
|              |                              |                |               |            |                   |              | unknownl.  DEC-1_N, DEC-1 protein, N terminal region. The defective chorion-1 gene (dec-1) in Drosophila encodes follicle cell proteins necessary for                                                                                                                                                                                                        |                         |                                  |                   |                   |                  |                      |               |                               |                                                                                                                                        |                           |                                  |                   |                   |                           |                            |
|              |                              |                |               |            | 5                 |              | proper eggshell assembly. Multiple products of the dec-1 gene are formed by alternative RNA splicing and proteolytic processing. Cleavage products include S80 (80 kDa) which is incorporated into the eggshell, and further ornteolysis of S80 gives S80 (60 kDa)                                                                                           | 42.95                   | 3.28E-05                         | 33%               | 39%               | 108175           | 75142                | 5             | BAE0283                       | 9 surface protein                                                                                                                      | 64.70                     | 8.26E-09                         | 50%               | 68%               | 56113                     | 572-629                    |
|              |                              |                |               |            | 6                 | pfam06735    | DUF1210, Protein of unknown function (DUF1210). This family recresents a conserved region within plant proline-rich proteins MCPVI, Minor capsid protein VI. This minor capsid protein may act as a                                                                                                                                                          | 42.41                   | 4.83E-05                         | 33%               | 37%               | 75183            | 100202               | 6             | AAM6381                       | 7 unknown                                                                                                                              | 59.69                     | 2.66E-07                         | 21%               | 40%               | 195477                    | 63-341                     |
|              |                              |                |               |            | 7                 | pfam02993    | link between the external capsid and the internal DNA-protein core. The<br>C-terminal 11 residues may function as a protease cofactor leading to                                                                                                                                                                                                             | 41.17                   | 9.40E-05                         | 32%               | 38%               | 114185           | 122196               | 7             | NP_19143                      | 9 unknown protein                                                                                                                      | 59.69                     | 2.66E-07                         | 21%               | 40%               | 195477                    | 63-341                     |
|              |                              |                |               |            | 8                 | pfam05518    | enzyme activation Totivirus_coat, Totivirus coat protein                                                                                                                                                                                                                                                                                                     | 40.74                   | 0.000133                         | 18%               | 23%               | 103183           | 670750               | 8             | XP_46478                      | 3 unknown protein                                                                                                                      | 51.22                     | 9.45E-05                         | 22%               | 41%               | 214477                    | 78-346                     |

| Part                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | Ger<br>Nan |                    | A.A.<br>length | Peptide<br>Mw | pl    | CDD Hit | COGs         | COG Definition                                                                                                                               | Bit<br>Score | E-value  | %<br>dentity P | %    | Query<br>from-to | Hit<br>from-to | BLASTp<br>Hit | Hit<br>Accession | BLASTp Definition                                                                                                 | Bit<br>Score | E-value <sub>I/</sub> | %<br>dentity Po | %    | Query H  | Hit from-<br>to |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|--------------------|----------------|---------------|-------|---------|--------------|----------------------------------------------------------------------------------------------------------------------------------------------|--------------|----------|----------------|------|------------------|----------------|---------------|------------------|-------------------------------------------------------------------------------------------------------------------|--------------|-----------------------|-----------------|------|----------|-----------------|
| 1                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |            |                    | -              |               | 4 64  | 1       | COG5201      | SKP1, SCF ubiquitin ligase, SKP1 component [Posttranslational                                                                                |              |          |                |      |                  |                | Number        |                  | contains ATP/GTP-binding motif A; similar to Dictyostelium FP21                                                   |              |                       |                 |      |          |                 |
| Part                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |            |                    |                | ,.            |       |         |              | modification, protein turnover, chaperonesi.<br>Skp1, Found in Skp1 protein family; Family of Skp1 (kinetochore protein                      |              |          |                |      |                  |                |               |                  | divcobrotein. corresponds to Swiss-Prot Accession Number P52285                                                   |              |                       |                 |      |          |                 |
| Part                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |            |                    |                |               |       |         |              | polymerase II transcription factor SIII) homologues                                                                                          |              |          |                |      |                  |                |               |                  | , , , , , , , , , , , , , , , , , , , ,                                                                           |              |                       |                 |      |          |                 |
| Part                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |            |                    |                |               |       | _       | p            |                                                                                                                                              |              |          |                |      |                  |                |               |                  | Chain B. Structure Of A Beta-Trcp1-Skp1-Beta-Catenin Complex:                                                     |              |                       |                 |      |          |                 |
| Part      |            |                    |                |               |       | *       | piamosesi    | SKPT_POZ, SKPT family, tetramensation domain                                                                                                 | 39.33        | 3.30E-10 | 4076           | 0076 | 161              | 100            |               |                  | Ubiquitin Ligase                                                                                                  |              |                       |                 |      |          |                 |
| Part      |            |                    |                |               |       |         |              |                                                                                                                                              |              |          |                |      |                  |                | 6             | BAB85607         | kinetochore protein                                                                                               | 104.76       | 8.02E-22              | 36%             | 58%  | 3-144    | 6159            |
| Part      |            |                    |                |               |       |         |              |                                                                                                                                              |              |          |                |      |                  |                | 8             | AAO85510         | SKP1                                                                                                              | 103.99       | 1.37E-21              | 37%             | 59%  | 3144     | 4152            |
| March   Marc   |            |                    |                |               |       |         |              |                                                                                                                                              |              |          |                |      |                  |                | 10            | AAI02436         | Hypothetical protein LOC615427                                                                                    |              |                       |                 |      |          |                 |
| March   Marc   | M81        | OR 299894300142    | 83             | 9,480         | 10.61 |         | No Hit Found |                                                                                                                                              |              |          |                |      |                  |                |               | No Hit Found     | No Hit Found                                                                                                      |              |                       |                 |      |          |                 |
| Mile                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | M8         | 11L 300904300149   | 252            | 29,824        | 5.27  |         | No Hit Found |                                                                                                                                              |              |          |                |      |                  |                |               | No Hit Found     | No Hit Found                                                                                                      |              |                       |                 |      |          |                 |
| Ministry    | M81        | 3R 301197301880    | 228            | 25,821        | 4.93  |         | No Hit Found |                                                                                                                                              |              |          |                |      |                  |                |               | No Hit Found     | No Hit Found                                                                                                      |              |                       |                 |      |          |                 |
| Miles   Mile   | M8         | 14L 302784302212   | 191            | 21,820        | 10.62 |         | No Hit Found |                                                                                                                                              |              |          |                |      |                  |                |               | No Hit Found     | No Hit Found                                                                                                      |              |                       |                 |      |          |                 |
| Math                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | M81        | 7R 302866303330    | 155            | 18,715        | 6.79  |         | No Hit Found |                                                                                                                                              |              |          |                |      |                  |                |               | No Hit Found     | No Hit Found                                                                                                      |              |                       |                 |      |          |                 |
| Part      | M81        | 8R 303386303766    | 127            | 14,321        | 4.20  |         | No Hit Found |                                                                                                                                              |              |          |                |      |                  |                |               | No Hit Found     | No Hit Found                                                                                                      |              |                       |                 |      |          |                 |
| Part      | M81        | 9R 303786304223    | 146            | 16,768        | 10.18 |         | No Hit Found |                                                                                                                                              |              |          |                |      |                  |                |               | No Hit Found     | No Hit Found                                                                                                      |              |                       |                 |      |          |                 |
| Part      | M82        | OR 304499304954    | 152            | 18,365        | 6.28  |         | No Hit Found |                                                                                                                                              |              |          |                |      |                  |                |               | No Hit Found     | No Hit Found                                                                                                      |              |                       |                 |      |          |                 |
| Part      |            |                    |                |               |       |         |              |                                                                                                                                              |              |          |                |      |                  |                |               |                  |                                                                                                                   |              |                       |                 |      |          |                 |
| Part      | M82        | 2R 305057306853    | 599            | 65,968        | 10.89 | 1       | cd01828      | tertiary fold of the enzyme is substantially different from that of the                                                                      | 79.17        | 3.73E-16 | 25%            | 38%  | 120275           | 1167           | 1             | NP 048674        | A318R                                                                                                             | 243.43       | 1.62E-62              | 56%             | 71%  | 403599   | 6211            |
| Part      |            |                    |                |               |       |         |              | active site closely resembles the Ser-His-Asp(Glu) triad found in other                                                                      |              |          |                |      |                  |                |               |                  |                                                                                                                   |              |                       |                 |      |          |                 |
| Region   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970   1970    |            |                    |                |               |       | 2       | COG0810      | TonB, Periplasmic protein TonB, links inner and outer membranes [Cell                                                                        | 60.54        | 1.47E-10 | 35%            | 42%  | 17117            | 28127          | 2             | NP_048672        | PAPK (17X); similar to PBCV-1 ORF A41R, corresponds to Genbank                                                    | 80.11        | 2.38E-13              | 60%             | 69%  | 299360   | 130194          |
| Ministry   Control property control pr   | M83        | RR 307542308210    | 223            | 25 745        | 7 60  |         | No Hit Found | envelope bioderiesis, oder memoraner.                                                                                                        |              |          |                |      |                  |                |               |                  | Accession Number 6 17033                                                                                          | 73.56        | 5.00F-12              | 28%             | 42%  | 1205     | 1247            |
| No.   1                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | WIOZ       | .dic 30/342-300210 | 223            | 23,743        | 7.00  |         | NO THE FOUND | ANK ankyrin reneats: ankyrin reneats mediate protein-protein                                                                                 |              |          |                |      |                  |                |               | 147_040430       | noon                                                                                                              | 73.30        | J.00L-12              | 2070            | 4270 | 1-200    | 1-247           |
| No.   1                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |            |                    |                |               |       |         |              | interactions in very diverse families of proteins. The number of ANK                                                                         |              |          |                |      |                  |                |               |                  |                                                                                                                   |              |                       |                 |      |          |                 |
| Part      | M8:        | 29L 309223308348   | 292            | 32,821        | 6.70  | 1       | cd00204      | ANK repeats may occur in combinations with other types of domains. The                                                                       | 98.23        | 6.93E-22 | 41%            | 68%  | 29150            | 2125           | 1             | XP_681288        | hypothetical protein AN8019.2                                                                                     | 150.98       | 4.01E-35              | 40%             | 57%  | 5-233    | 8671100         |
| Part      |            |                    |                |               |       |         |              | contains 4 consecutive repeats                                                                                                               |              |          |                |      |                  |                |               |                  |                                                                                                                   |              |                       |                 |      |          |                 |
| Registration   Part     |            |                    |                |               |       | 2       | COG0666      | Arp, FOG: Ankyrin repeat [General function prediction only].  Ank, Ankyrin repeat. There's no clear separation between noise                 | 62.60        | 3.44E-11 | 30%            | 42%  | 5148             | 77224          | 2             | CAE64680         | Hypothetical protein CBG09456                                                                                     | 136.35       | 1.02E-30              | 36%             | 54%  | 5-230    | 463693          |
| Michael   Mich   |            |                    |                |               |       | 3       | pfam00023    | beta, alpha, alpha, beta order of secondary structures. The repeats                                                                          | 45.05        | 6.98E-06 | 59%            | 78%  | 193225           | 1-33           | 3             | XP_392578        | PREDICTED: similar to CG7462-PB, isoform B                                                                        | 135.19       | 2.28E-30              | 36%             | 54%  | 5-230    | 441671          |
| Part      |            |                    |                |               |       |         |              | ANK ankyrin reneate: Ankyrin reneate are about 33 amino acide long                                                                           |              |          |                |      |                  |                |               |                  |                                                                                                                   |              |                       |                 |      |          |                 |
| Part      |            |                    |                |               |       | 4       | smart00248   | protein interactions. The core of the repeat seems to be an helix-loop-                                                                      | 38.02        | 0.000936 | 59%            | 78%  | 193220           | 128            | 4             | AAF61702         | ankyrin 1                                                                                                         | 132.88       | 1.13E-29              | 39%             | 54%  | 5-230    | 506-736         |
| Part      |            |                    |                |               |       |         |              | neix suddure                                                                                                                                 |              |          |                |      |                  |                | 5             | XP_581734        | PREDICTED: similar to Ankyrin-1 (Erythrocyte ankyrin), partial                                                    |              |                       |                 |      |          |                 |
| All                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |            |                    |                |               |       |         |              |                                                                                                                                              |              |          |                |      |                  |                | 7             | AAN12046         | CG7462-PC, isoform C                                                                                              | 131.34       | 3.29E-29              | 36%             | 54%  | 5230     | 374604          |
| All                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |            |                    |                |               |       |         |              |                                                                                                                                              |              |          |                |      |                  |                | 9             | EAL87814         | NACHT domain protein putative                                                                                     | 130.95       | 4.29E-29              | 36%             | 57%  | 20240 13 | 2701489         |
| R93. 310081—300226                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |            |                    |                |               |       |         |              | ANK ankyrin reneats: ankyrin reneats mediate protein-protein                                                                                 |              |          |                |      |                  |                |               | 701001011        | oncoordinated protein 44, isolomi q                                                                               | 100.00       | 4.202.20              | 0070            | 0470 | U LL     | 402 000         |
| ## 1808-10000-1000000000000000000000000000                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |            |                    |                |               |       |         |              | interactions in very diverse families of proteins. The number of ANK repeats in a protein can range from 2 to over 20 (ankyrins for example) |              |          |                |      |                  |                |               |                  |                                                                                                                   |              |                       |                 |      |          |                 |
| Part      | M8:        | 33L 310061309258   | 268            | 28,814        | 5.83  | 1       | cd00204      | ANK repeats may occur in combinations with other types of domains. The                                                                       | 126.74       | 1.66E-30 | 46%            | 59%  | 38156            | 8126           | 1             | NP_049038        | contains 10 ankyrin-like repeats; similar to human ankyrin, corresponds to Swiss-Prot Accession Number P16157     | 167.93       | 2.74E-40              | 37%             | 54%  | 8-268    | 109368          |
| And A registration denotes in decided plants and the secondary adjunction control registration denotes in decided plants and succession of the registration denotes in color and spiration. The core of the repeat search to the secondary adjunction. The core of the repeat search to the secondary adjunction. The core of the repeat search to the secondary adjunction. The core of the repeat search to the secondary adjunction. The core of the repeat search to the secondary adjunction. The core of the repeat search to the secondary adjunction. The core of the repeat search to the secondary adjunction. The core of the repeat search to the secondary adjunction. The core of the repeat search to the secondary adjunction. The core of the repeat search to the secondary adjunction. The core of the repeat search to the secondary adjunction. The core of the repeat search to the secondary adjunction. The core of the repeat search to the secondary adjunction. The core of the repeat search to the secondary adjunction. The core of the repeat search to the secondary adjunction. The core of the repeat search to the secondary adjunction. The core of the repeat search to the secondary adjunction. The core of the repeat search to the secondary adjunction. The core of the repeat search to the registration. The core of the repeat search to the registration. The core of the repeat search to the registration. The core of the repeat search to the registration. The core of the repeat search to the registration. The core of the repeat search to the registration. The core of the repeat search to the registration. The core of the repeat search to the registration. The core of the repeat search to the registration. The core of the registration. The core of the repeat search to the registration. The core of the repeat search to the registration. The core of the repeat search to the registration. The core of the repeat search to the registration. The core of the repeat search to the registration. The core of the repeat search to the registration. T |            |                    |                |               |       |         |              | repeats are stacked in a superhelical arrangement; this alignment                                                                            |              |          |                |      |                  |                |               |                  |                                                                                                                   |              |                       |                 |      |          |                 |
| And A registration denotes in decided plants and the secondary adjunction control registration denotes in decided plants and succession of the registration denotes in color and spiration. The core of the repeat search to the secondary adjunction. The core of the repeat search to the secondary adjunction. The core of the repeat search to the secondary adjunction. The core of the repeat search to the secondary adjunction. The core of the repeat search to the secondary adjunction. The core of the repeat search to the secondary adjunction. The core of the repeat search to the secondary adjunction. The core of the repeat search to the secondary adjunction. The core of the repeat search to the secondary adjunction. The core of the repeat search to the secondary adjunction. The core of the repeat search to the secondary adjunction. The core of the repeat search to the secondary adjunction. The core of the repeat search to the secondary adjunction. The core of the repeat search to the secondary adjunction. The core of the repeat search to the secondary adjunction. The core of the repeat search to the secondary adjunction. The core of the repeat search to the secondary adjunction. The core of the repeat search to the secondary adjunction. The core of the repeat search to the registration. The core of the repeat search to the registration. The core of the repeat search to the registration. The core of the repeat search to the registration. The core of the repeat search to the registration. The core of the repeat search to the registration. The core of the repeat search to the registration. The core of the repeat search to the registration. The core of the repeat search to the registration. The core of the registration. The core of the repeat search to the registration. The core of the repeat search to the registration. The core of the repeat search to the registration. The core of the repeat search to the registration. The core of the repeat search to the registration. The core of the repeat search to the registration. T |            |                    |                |               |       | 2       | COG0666      | Arp, FOG: Ankyrin repeat [General function prediction only].                                                                                 | 93.41        | 1.76E-20 | 38%            | 53%  | 21159            | 57203          | 2             | NP_048353        | contains 4 ankyrin repeats; similar to reticulocyte ankyrin, corresponds to<br>Swiss-Prot Accession Number P16157 | 164.85       | 2.32E-39              | 37%             | 52%  | 1252     | 1251            |
| Settle 4. Sur 1. |            |                    |                |               |       | 3       | pfam00023    | and signal on the HMM search Ankyrin repeats generally consist of a                                                                          | 44.28        | 1.17E-05 | 57%            | 67%  | 104134           | 3-33           |               |                  |                                                                                                                   | 141.35       | 2.74E-32              | 44%             | 55%  | 8189     | 234415          |
| ### ### ### ### ### ### ### ### ### ##                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |            |                    |                |               |       |         |              | associate to form a higher order structure                                                                                                   |              |          |                |      |                  |                |               | _                |                                                                                                                   |              |                       |                 |      |          |                 |
| Facility    |            |                    |                |               |       | 4       | smart00248   | and occur in at least four consecutive copies. They are involved in protein                                                                  | 39.17        | 0.000447 | 56%            | 70%  | 104131           | 330            | 4             | XP_796302        | PREDICTED: similar to Ankyrin repeat domain protein 28, partial                                                   | 140.97       | 3.58E-32              | 41%             | 53%  | 8220     | 744954          |
| Register    |            |                    |                |               |       |         |              | protein interactions. The core of the repeat seems to be an neitx-roop-<br>helix structure.                                                  |              |          |                |      |                  |                | 5             | EAL84954         | ankvrin repeat protein                                                                                            | 134.42       | 3.35E-30              | 40%             | 56%  | 8215     | 430634          |
| S                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |            |                    |                |               |       |         |              |                                                                                                                                              |              |          |                |      |                  |                | 6             | XP 782299        | PREDICTED: similar to Ankyrin-3 (ANK-3) (Ankyrin G)                                                               |              |                       |                 |      |          |                 |
| M838L 310886-310315 124 13,324 8.64 No Hit Found                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |            |                    |                |               |       |         |              |                                                                                                                                              |              |          |                |      |                  |                | 8             | XP 782887        | PREDICTED: similar to Ankvrin-1 (Erythrocyte ankvrin) (Ankvrin R)                                                 | 129.80       | 8.26E-29              | 36%             | 57%  | 3189     | 374560          |
| M838L 31686-310315 124 33.24 8.64 No Hit Found 1 2 P00575949 RepA / Rep+ protein KID 2 No_70565 hypothetical marian ariangem ariangem an ariange   |            |                    |                |               |       |         |              |                                                                                                                                              |              |          |                |      |                  |                | 10            | XP_788092        | PREDICTED: similar to Ankyrin-2 (Brain ankyrin) (Ankyrin B) (Ankyrin, poperythroid) partial                       |              |                       | 41%             | 59%  |          |                 |
| M842R 31500—312168 22 3 £.253 6.70 No Hit Found 1 ZP_0057594P RepA/Rep+ protein KID 55.45 1.41E-06 25% 59% 14—113 61—100 2 NP_705165 hypothetical arrian antique 1 51.00 2 NP_601850 hypothetical protein BGP110 45.68 77.65-05 27% 54% 27—137 91—3129—3129 NP_601850 hypothetical protein BGP110 45.68 77.65-05 27% 54% 27—137 91—3129 NP_601850 hypothetical protein BGP110 45.68 77.65-05 27% 54% 27—137 91—3129 NP_601850 hypothetical protein BGP110 45.68 77.65-05 27% 54% 27—137 91—3129 NP_601850 hypothetical protein BGP110 45.68 NP_60850 hypothetical protein BGP110 45.68 NP | M8:        | 38L 310686310315   | 124            | 13,324        | 8.64  |         | No Hit Found |                                                                                                                                              |              |          |                |      |                  |                |               |                  |                                                                                                                   |              |                       |                 |      |          |                 |
| 2 NP_76165 Npcombeted malarial arriticen                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |            |                    |                |               |       |         |              |                                                                                                                                              |              |          |                |      |                  |                | 1             |                  |                                                                                                                   | 55.45        | 1.41E-06              | 25%             | 59%  | 14113    | 61160           |
| 1 NP_048728 A368L 153.0 0 No Hit Found 1 NP_048728 A368L 153.0 0 A368L 1 |            |                    | -              | .,            |       |         |              |                                                                                                                                              |              |          |                |      |                  |                | 2             | NP 705165        | hypothetical malaria antigen                                                                                      |              |                       |                 |      |          |                 |
| 2 NP_049005 similar to Chloreby AdsOR, corresponds to GenBank 59.31 1.14E-07 28% 40% 2-167 8-180 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | M84        | 4R 312868313596    | 243            | 27,062        | 8.50  |         | No Hit Found |                                                                                                                                              |              |          |                |      |                  |                | 1             | NP 048723        | A366L                                                                                                             |              |                       |                 |      |          |                 |
| 3 NP_049003 similar to Chlorolada virus PECV-1 ORF A450R, corresponds to GenBank 50,07 2.16E-06 41% 60% 1-60 1-59 4 NP_049003 similar to PECV-1 ORF A275R, encoded by GenBank Accession Number 53,18 18,19E-06 48% 63% 1-47 1-47 5 NP_048007 similar by PECV-1 ORF A275R, encoded by GenBank Accession Number 51,18 18,19E-06 48% 63% 1-47 1-47 1-45 18,19E-06 48% 63% 1-47 1-47 1-45 18,19E-06 48% 63% 1-47 1-45 18,19E-06 48% 63% 1-47 1-45 18,19E-06 18,19E |            |                    |                | ,             | 2.20  |         |              |                                                                                                                                              |              |          |                |      |                  |                | 2             | NP_049005        | similar to Chlorella virus PBCV-1 ORF A450R, corresponds to GenBank                                               |              |                       |                 |      |          |                 |
| 5 NP_048807 similar to PEGV-1 ORF A275R, encoded by GenBank Accession Number 53.14 8.19E-06 48% 63% 1-47 1-47 14590 6 NP_048525 A177R 51.60 2.38E-05 40% 63% 1-47 4-50                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |            |                    |                |               |       |         |              |                                                                                                                                              |              |          |                |      |                  |                | 3             | NP_049003        | similar to Chlorella virus PBCV-1 ORF A450R, corresponds to GenBank<br>Accession Number LI42580                   |              |                       |                 |      |          |                 |
| 6 NP_048525 A177R 51.60 2.38E-05 40% 63% 1-47 4-50                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |            |                    |                |               |       |         |              |                                                                                                                                              |              |          |                |      |                  |                | 4             | NP_048427        | A79R<br>similar to PBCV-1 ORF A275R, encoded by GenBank Accession Number                                          |              |                       |                 |      |          |                 |
| M845L 313927-313712 72 8,515 6.34 No Hit Found No Hit Found No Hit Found                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |            |                    |                |               |       |         |              |                                                                                                                                              |              |          |                |      |                  |                | 6             | NP_048525        | U42580<br>A177R                                                                                                   |              |                       |                 |      |          |                 |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | M8         | 15L 313927313712   | 72             | 8,515         | 6.34  |         | No Hit Found |                                                                                                                                              |              |          |                |      |                  |                |               | No Hit Found     | No Hit Found                                                                                                      |              |                       |                 |      |          |                 |