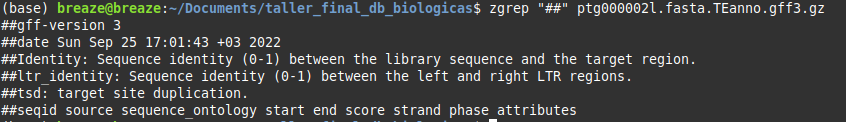
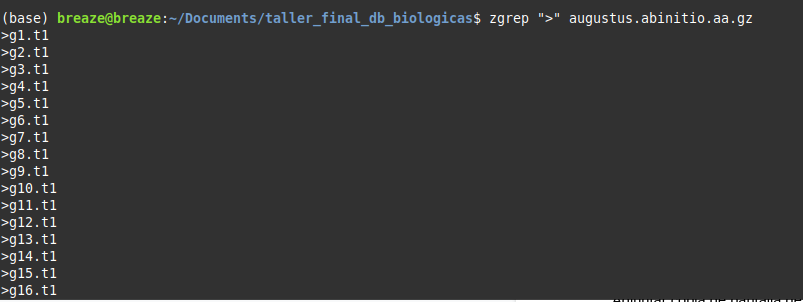
Repository: <https://github.com/breaze/biological_databases>

* This activity is located in the activity2 folder
* You can create a copy of the entire environment using the file environment.yml
* If you want to restore the database just execute a mongorestore using the folder “taller\_final”

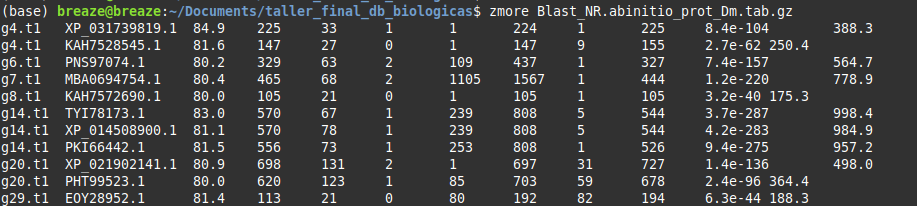
1. Headers of each download file:
2. ptg000002l.fasta.TEanno.gff3.gz



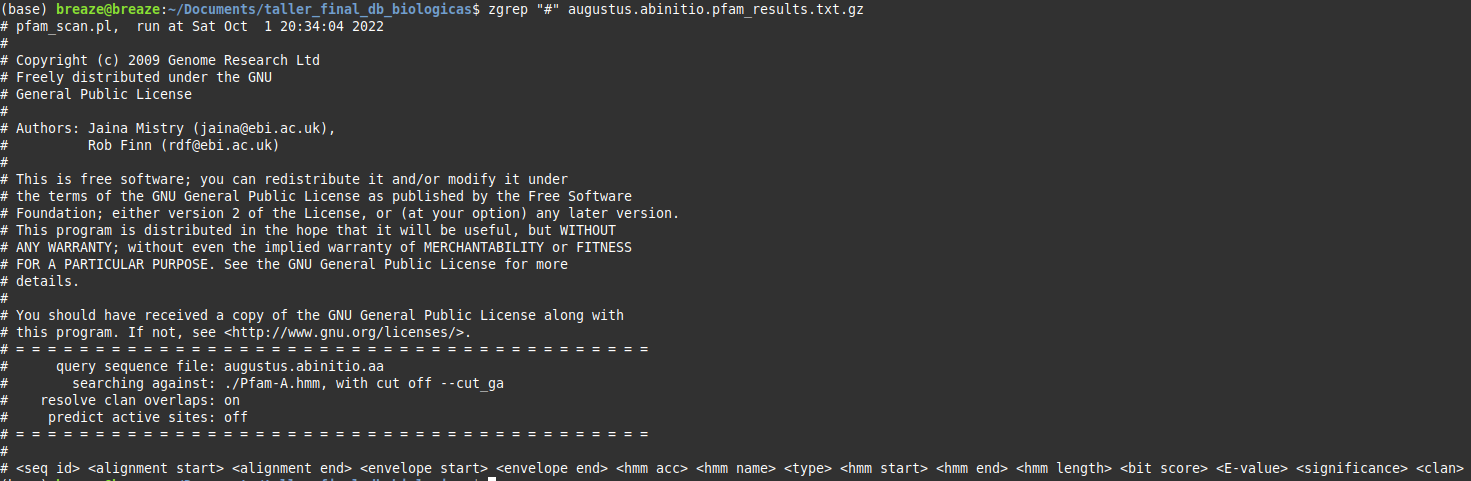
1. augustus.abinitio.aa.gz



1. Blast\_NR.abinitio\_prot\_Dm.tab.gz



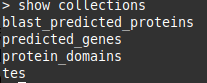
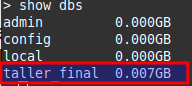
1. augustus.abinitio.pfam\_results.txt.gz



1. ptg000002l.fasta.gz



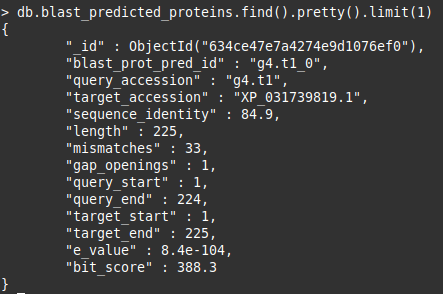
1. Database structure:



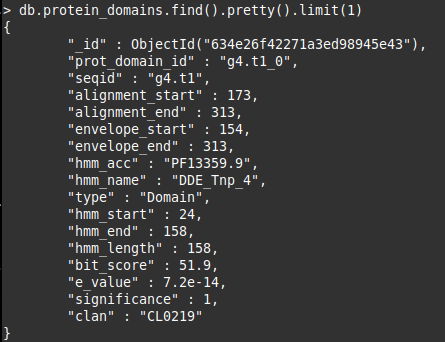
1. TEs



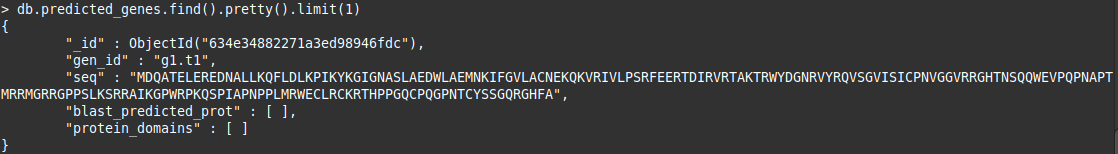
1. Blast predicted proteins



1. Protein domains

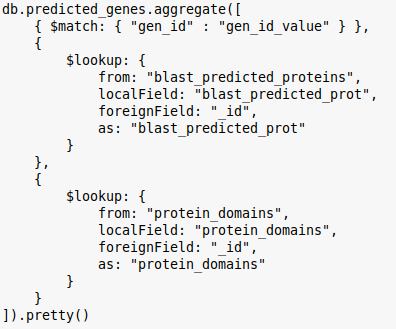


1. Predicted genes



1. Genes info

Query structure (Just replace “gen\_id\_value” for the id of each gene of interest):



1. g1026.t1

{

"\_id" : ObjectId("634e34882271a3ed989473dd"),

"gen\_id" : "g1026.t1",

"seq" : "MPVVNGGFPAFRILNPGVTPTPPKTRGLKRRRIGHWFNKVAGELLTPTAPMRVSHLTAIKGSGIKKRWGKEKFDHAAATPPGTVQQRGGSAGTGKNDVACSPDKRNGPEMSAARVDCSFCYLSGEMEALTELCDIIAQNPRQFSEKLAWICGRCPQPEWLLSRSSKVSWSHLNAVLAVARILSRSPDSTDNRAKSVVLEFLRAIPASFRRSFWPQSFVTDAISSFYADFLGYVSTAADLSSEFATEIAAFTGEVVIAAIEKREEGGESEVDAAISRAFLLAFSRNFPPILPSDGDKLITLLLDQCSGGPVSSSPREKLQQQLSINSETSSSQSSPVSVKPSANNEGSSPGNEAASHVSSASSKSSMSVVVNGASIVWKSGVDQFSFGFSDGSGGSGAVFRQHVASFEDESIETLEKQEIAFKLITHALEKVRIDPKLLEQVRVTAKKQLQSMSAFLKIRKRDWNEQGTLLKARVNAKLSAYQAAARMKIRSLVSLDTDGKTSKRLVLETLALLLDAAEACLLSVWRKMRVCEELFGSLLSGIAQIAVSRGGQPLRVLLIRLKPLVLAVCAQADTWGSSQGALFESILKTSCEIIESGWNKDRAPVDTFIMGLATSIRERNDYEEQVDKEMQTVPAMQLNVIRLLADLNFVVKKPEVVEMILPLFVENLEEGDASTPSLLRLRLLDAVSRIASLGFEKSYRETVVLMTRSYLSKLSSVGSAESKTSAPEAATECIETLPAGFLMIASGLADTKLRSDYRHRLLSLCSDVGLAAESKSGRSGADFLGPLLPAVAEICSDFDPTVHVEPSLLKLFRNLWFYVALFGLAPPIVRNQPPAKSISNVTNSLVSVGTISLQAVSGPYMWNAQWSSAVQRIAQGTPPLVVSTVKWLEDELELNALHNPGSRRGSGNEKAAVTQRAALSAALGGHVDVAAMGTISGVKATYLLAVTFLEIIRFSSNGGILNGGSTATASRSAFSCVFEYLKTPNLMPAVFQCLTALVHRAFETAVAWLEDQTLATGKEAEIRESTLFAHACFLIKRMSQREEHIRDISVNLLTQLRDRFPQVLWNSSCLDSLFLPVHEETPSTVVNDPAWKAAVRSLYQKIVREWIIISLSYAPCTTQGLLQDKLCKSNTWQRAHPTTDMVSLLSEIKISTGKNDIWTGVRTANIPAVIAAAAAASGANLKVSEAFNLEVLSTGVVSATVKCNHAGEIAGMRRLYNSIGGFQGGVGPTGIGGGLQRLISGAFSQPLQTENDSFNEMLVSKFVRLLQQFVNTAEKGGKVDKSQFRETCSQATALLLSNLGCESKSNVDGFSQLLRLLCWCPAYISTPDSMETGVFIWTWLVSAAPKLGSLVLAELVDAWLWTIDTKRGLFASHVRYSGPVAKLRPHLSPGEPEDPPEVDPVDQIIAHRLWLGFFIDRFEIVRQNSVEQMLLLGRMLQGTTKPPWNFSRHPAATGTFFTVMLLGLKFCSCQSQNNLQNFKTGLQLLEDRVYRASLVWFAHDPEWFDVNIANFAQSEAQSLSVFVHYLMNERLDTPTELKGRSRENGNILVDGNDEYHPVWGHMDNYALGREKRKQLLLMLCQHEADRLDVWAQTTITKESPYPRQKLSAEKWVEYARTAFSVDPRIALSLASRYYRRIEHTYDKQVIFLTIRHSYILDLRCIPEALPFFVTPKAVEENSVLLQQLPHWGACSITQALEFLTPAYKGHPRVMAYVLRVLESYPPEKVTFFMPQLVQALRYDEGKLVEGYLLRAAQRSDIFAHILIWHLQGESAPEPGKDASSEKNAAFQALLPLVRRHIIDGFTPKALDLFRREFAFFDKVTSISGVLFPIPKEERRAGIKRELEKIEVEGDDLYLPTAHTKLVRGIRVDSGIPLQSAAKVPIMITFNVVDRDGDPSDVKPQACIFKVGDDCRQDVLALQVIALLRDIFQAVGLNLYLFPYGVLPTGPERGIIEVVPNTRSRSQMGETTDGGLYEIFQQDYGPVGSPSFEAARENFLISSAGYAVASLILQPKDRHNGNLLFDNVGRLVHIDFGFILETSPGGNMRFESAHFKLSHEMTQLLDPSGVMKSDTWNQFVSLCVKGYLAARRYMDGIINTVLMMLDSGLPCFSRGDPIGNLRKRFHPEMSDREAANFMIRVCTDAYNKWTTAGYDLIQYLQQGIEK",

"blast\_predicted\_prot" : [

{

"\_id" : ObjectId("634ce47e7a4274e9d107732c"),

"blast\_prot\_pred\_id" : "g1026.t1\_1084",

"query\_accession" : "g1026.t1",

"target\_accession" : "XP\_010527597.1",

"sequence\_identity" : 81.6,

"length" : 2040,

"mismatches" : 360,

"gap\_openings" : 10,

"query\_start" : 126,

"query\_end" : 2161,

"target\_start" : 1,

"target\_end" : 2028,

"e\_value" : 0,

"bit\_score" : 3275.3

},

{

"\_id" : ObjectId("634ce47e7a4274e9d107732d"),

"blast\_prot\_pred\_id" : "g1026.t1\_1085",

"query\_accession" : "g1026.t1",

"target\_accession" : "XP\_010527598.1",

"sequence\_identity" : 81.5,

"length" : 2040,

"mismatches" : 361,

"gap\_openings" : 10,

"query\_start" : 126,

"query\_end" : 2161,

"target\_start" : 1,

"target\_end" : 2028,

"e\_value" : 0,

"bit\_score" : 3271.9

},

{

"\_id" : ObjectId("634ce47e7a4274e9d107732e"),

"blast\_prot\_pred\_id" : "g1026.t1\_1086",

"query\_accession" : "g1026.t1",

"target\_accession" : "KAG7599921.1",

"sequence\_identity" : 80.2,

"length" : 2042,

"mismatches" : 390,

"gap\_openings" : 10,

"query\_start" : 126,

"query\_end" : 2161,

"target\_start" : 1,

"target\_end" : 2033,

"e\_value" : 0,

"bit\_score" : 3226

},

{

"\_id" : ObjectId("634ce47e7a4274e9d107732f"),

"blast\_prot\_pred\_id" : "g1026.t1\_1087",

"query\_accession" : "g1026.t1",

"target\_accession" : "XP\_006306575.1",

"sequence\_identity" : 80.8,

"length" : 2040,

"mismatches" : 376,

"gap\_openings" : 9,

"query\_start" : 126,

"query\_end" : 2161,

"target\_start" : 1,

"target\_end" : 2029,

"e\_value" : 0,

"bit\_score" : 3216.4

},

{

"\_id" : ObjectId("634ce47e7a4274e9d1077330"),

"blast\_prot\_pred\_id" : "g1026.t1\_1088",

"query\_accession" : "g1026.t1",

"target\_accession" : "GAV80522.1",

"sequence\_identity" : 81.1,

"length" : 2049,

"mismatches" : 360,

"gap\_openings" : 12,

"query\_start" : 126,

"query\_end" : 2161,

"target\_start" : 1,

"target\_end" : 2035,

"e\_value" : 0,

"bit\_score" : 3216

}

],

"protein\_domains" : [

{

"\_id" : ObjectId("634e26f42271a3ed9894630e"),

"prot\_domain\_id" : "g1026.t1\_1227",

"seqid" : "g1026.t1",

"alignment\_start" : 421,

"alignment\_end" : 1597,

"envelope\_start" : 421,

"envelope\_end" : 1598,

"hmm\_acc" : "PF19274.2",

"hmm\_name" : "PI4K\_N",

"type" : "Family",

"hmm\_start" : 1,

"hmm\_end" : 1178,

"hmm\_length" : 1179,

"bit\_score" : 2111.2,

"e\_value" : 0,

"significance" : 1,

"clan" : "No\_clan"

},

{

"\_id" : ObjectId("634e26f42271a3ed9894630f"),

"prot\_domain\_id" : "g1026.t1\_1228",

"seqid" : "g1026.t1",

"alignment\_start" : 1648,

"alignment\_end" : 1778,

"envelope\_start" : 1638,

"envelope\_end" : 1800,

"hmm\_acc" : "PF00613.23",

"hmm\_name" : "PI3Ka",

"type" : "Repeat",

"hmm\_start" : 34,

"hmm\_end" : 157,

"hmm\_length" : 185,

"bit\_score" : 77,

"e\_value" : 1.3e-21,

"significance" : 1,

"clan" : "CL0020"

},

{

"\_id" : ObjectId("634e26f42271a3ed98946310"),

"prot\_domain\_id" : "g1026.t1\_1229",

"seqid" : "g1026.t1",

"alignment\_start" : 1901,

"alignment\_end" : 1960,

"envelope\_start" : 1899,

"envelope\_end" : 1966,

"hmm\_acc" : "PF00454.30",

"hmm\_name" : "PI3\_PI4\_kinase",

"type" : "Family",

"hmm\_start" : 3,

"hmm\_end" : 66,

"hmm\_length" : 250,

"bit\_score" : 33.4,

"e\_value" : 4.3e-8,

"significance" : 1,

"clan" : "CL0016"

},

{

"\_id" : ObjectId("634e26f42271a3ed98946311"),

"prot\_domain\_id" : "g1026.t1\_1230",

"seqid" : "g1026.t1",

"alignment\_start" : 1961,

"alignment\_end" : 2108,

"envelope\_start" : 1958,

"envelope\_end" : 2109,

"hmm\_acc" : "PF00454.30",

"hmm\_name" : "PI3\_PI4\_kinase",

"type" : "Family",

"hmm\_start" : 107,

"hmm\_end" : 249,

"hmm\_length" : 250,

"bit\_score" : 66,

"e\_value" : 4.8e-18,

"significance" : 1,

"clan" : "CL0016"

}

]

}

1. g1134.t1

{

"\_id" : ObjectId("634e34882271a3ed98947449"),

"gen\_id" : "g1134.t1",

"seq" : "MKTRLLRLLVSAPFAMPEVFVGYTPPGIARDPSHATCKIRKCKRCYNLPYVCPQFFPDSCTVECAYCKYTSPPEEPTTPSLSPLPTTPSTMSPSPTLSTPSLSPSLSTASPSLPTHPLMPSTPSPSPSFTLIIKTIYVAQIKKKCLLHKRWCLKHHLQSPTHRYKKRNPAQGTMPMGFKSVPPSVLDMPAQDHNFIGESELRLGRCSAGLCLFPRLLGFNAERVVVVSDRAEVVRMAGVMDITPNKFSLGTCRQTLFLQKNIPIKRSQLLWGSIGSSIPLVSRKSVSMRRCARSEPRAIASGQAATPVSEKSSLTEKPASEVIHFYRVPLVQESASAELLKLIQVKVSNHIVGLRTEQCFNIELRSDISNEKLLALKWLLQETYEPENLGTESFLEKKKQEGCFATIVEVGPRLSFTTAWSTNAVSICHACGLTEVSRLERSRRYLLFSKDSLSESQIHEFTAMVHDRMTECVYPQKLTSFSTSVVPEEVRFVSVMERGREALEEINQEMGLAFDEQDLEYYTRLFGDEIKRNPTNVELFDIAQSNSEHSRHWFFNGKIVIDGQPMSRTLMQIVKSTLKANPNNSVIGFKDNSSAIRGFPVKQMHPAQPGLTSPLNEVACDLDILFTAETHNFPCAVAPYPGAETGAGGRIRDTHATGRGSLVVASTAGYCVGNLNMEGAYAPWEDSSFKYPSNLASPLQILIDSSNGASDYGNKFGEPLIQGYTRTFGMRLPSGERREWLKPIMFSGGIGQIDHNHIHKGEPDIGMLVVKIGGPAYRIGMGGGAASSVVSGTNDAELDFNAVQRGDAEMAQKLYRVVRACVEMGEENPIISIHDQGAGGNCNVVKEIIDPKGAVIDIRAIVLGDNTMSVLEIWGAEYQEQDAILVKPESREILESICERERVSMAVIGTISGEGRIVLVDREAVEKSRESGLPSPPPAVDLELDKVLGDMPQKTFEFKRVVNVQEGLDIAPGITVMECLKRVLRLPSVCSKRFLTTKVDRCVTGLVAQQQTVGPLQITLSDVAVIAQTYTDVTGAACAIGEQPIKGLLDPKAMARLAVGEALTNLVWAKVTSLSDVKASGNWMYAAKLEGEGAAMYDAATALSEAMIELGIAIDGGKDSLSMAAHAGGEVVKAPGNLVISVYVTCPDITKTVTPDLKLGDDGVLLHLDLAKGKRRLGGSALAQVFDQIGNDFPDLEDVPYLKTVFEAVQDLIAEQLISAGHDISDGGLLVCALEMAFAGNKGLTLSLSSKRNYSLLETLLAEELGLILEVGKKNLDTVIRKLTSVGVDAEIIGTGADSPSVEVKVDGITHLNEKMYVLRDMWEDTSFQLEKLQRLASCVESEKEGLKSRREPSWALSFTPSFTDQKYMSAASKPKVAIIREEGSNGDREMSGAFYAAGFEPWDIAMSDLLRGAITLDEFRGVAFVGGFSYADVLDSAKGWSASIRFNGPLLNQFQEFYRRPDTFSLGVCNGCQLMALLGWVPGPQVGGASGGGGDTSQPRFIHNKSGRFECRFTSVTIKDSPAIMFKGMEGSTLGVWAAHGEGRAYFPDGGVADRMLQSGLAPVRYCDDDGSPTEVYPFNLNGSPLGIAAICSPDGRHLAMMPHPERCFLMWQFPWYPSHWDVEKKGPSPWLRMFQNAREWCS",

"blast\_predicted\_prot" : [

{

"\_id" : ObjectId("634ce47e7a4274e9d1077370"),

"blast\_prot\_pred\_id" : "g1134.t1\_1152",

"query\_accession" : "g1134.t1",

"target\_accession" : "KAH7548137.1",

"sequence\_identity" : 82.8,

"length" : 1409,

"mismatches" : 238,

"gap\_openings" : 2,

"query\_start" : 236,

"query\_end" : 1644,

"target\_start" : 1,

"target\_end" : 1405,

"e\_value" : 0,

"bit\_score" : 2368.2

},

{

"\_id" : ObjectId("634ce47e7a4274e9d1077371"),

"blast\_prot\_pred\_id" : "g1134.t1\_1153",

"query\_accession" : "g1134.t1",

"target\_accession" : "GAY59081.1",

"sequence\_identity" : 82.3,

"length" : 1416,

"mismatches" : 243,

"gap\_openings" : 3,

"query\_start" : 235,

"query\_end" : 1644,

"target\_start" : 14,

"target\_end" : 1428,

"e\_value" : 0,

"bit\_score" : 2364.7

},

{

"\_id" : ObjectId("634ce47e7a4274e9d1077372"),

"blast\_prot\_pred\_id" : "g1134.t1\_1154",

"query\_accession" : "g1134.t1",

"target\_accession" : "GAY59084.1",

"sequence\_identity" : 82.3,

"length" : 1416,

"mismatches" : 243,

"gap\_openings" : 3,

"query\_start" : 235,

"query\_end" : 1644,

"target\_start" : 8,

"target\_end" : 1422,

"e\_value" : 0,

"bit\_score" : 2364.7

},

{

"\_id" : ObjectId("634ce47e7a4274e9d1077373"),

"blast\_prot\_pred\_id" : "g1134.t1\_1155",

"query\_accession" : "g1134.t1",

"target\_accession" : "XP\_031251215.1",

"sequence\_identity" : 82,

"length" : 1412,

"mismatches" : 250,

"gap\_openings" : 2,

"query\_start" : 236,

"query\_end" : 1644,

"target\_start" : 1,

"target\_end" : 1411,

"e\_value" : 0,

"bit\_score" : 2362.8

},

{

"\_id" : ObjectId("634ce47e7a4274e9d1077374"),

"blast\_prot\_pred\_id" : "g1134.t1\_1156",

"query\_accession" : "g1134.t1",

"target\_accession" : "XP\_015388130.1",

"sequence\_identity" : 82.7,

"length" : 1406,

"mismatches" : 239,

"gap\_openings" : 2,

"query\_start" : 242,

"query\_end" : 1644,

"target\_start" : 10,

"target\_end" : 1414,

"e\_value" : 0,

"bit\_score" : 2362.8

}

],

"protein\_domains" : [

{

"\_id" : ObjectId("634e26f42271a3ed989463ac"),

"prot\_domain\_id" : "g1134.t1\_1385",

"seqid" : "g1134.t1",

"alignment\_start" : 359,

"alignment\_end" : 476,

"envelope\_start" : 357,

"envelope\_end" : 481,

"hmm\_acc" : "PF18076.4",

"hmm\_name" : "FGAR-AT\_N",

"type" : "Domain",

"hmm\_start" : 3,

"hmm\_end" : 107,

"hmm\_length" : 115,

"bit\_score" : 72.5,

"e\_value" : 2.9e-20,

"significance" : 1,

"clan" : "No\_clan"

},

{

"\_id" : ObjectId("634e26f42271a3ed989463ad"),

"prot\_domain\_id" : "g1134.t1\_1386",

"seqid" : "g1134.t1",

"alignment\_start" : 503,

"alignment\_end" : 552,

"envelope\_start" : 503,

"envelope\_end" : 552,

"hmm\_acc" : "PF18072.4",

"hmm\_name" : "FGAR-AT\_linker",

"type" : "Domain",

"hmm\_start" : 1,

"hmm\_end" : 49,

"hmm\_length" : 50,

"bit\_score" : 44.9,

"e\_value" : 1.3e-11,

"significance" : 1,

"clan" : "No\_clan"

},

{

"\_id" : ObjectId("634e26f42271a3ed989463ae"),

"prot\_domain\_id" : "g1134.t1\_1387",

"seqid" : "g1134.t1",

"alignment\_start" : 766,

"alignment\_end" : 919,

"envelope\_start" : 764,

"envelope\_end" : 922,

"hmm\_acc" : "PF02769.25",

"hmm\_name" : "AIRS\_C",

"type" : "Domain",

"hmm\_start" : 3,

"hmm\_end" : 153,

"hmm\_length" : 156,

"bit\_score" : 80.1,

"e\_value" : 2e-22,

"significance" : 1,

"clan" : "No\_clan"

},

{

"\_id" : ObjectId("634e26f42271a3ed989463af"),

"prot\_domain\_id" : "g1134.t1\_1388",

"seqid" : "g1134.t1",

"alignment\_start" : 1176,

"alignment\_end" : 1306,

"envelope\_start" : 1164,

"envelope\_end" : 1307,

"hmm\_acc" : "PF02769.25",

"hmm\_name" : "AIRS\_C",

"type" : "Domain",

"hmm\_start" : 13,

"hmm\_end" : 155,

"hmm\_length" : 156,

"bit\_score" : 59,

"e\_value" : 6.2e-16,

"significance" : 1,

"clan" : "No\_clan"

},

{

"\_id" : ObjectId("634e26f42271a3ed989463b0"),

"prot\_domain\_id" : "g1134.t1\_1389",

"seqid" : "g1134.t1",

"alignment\_start" : 1374,

"alignment\_end" : 1643,

"envelope\_start" : 1374,

"envelope\_end" : 1643,

"hmm\_acc" : "PF13507.9",

"hmm\_name" : "GATase\_5",

"type" : "Domain",

"hmm\_start" : 1,

"hmm\_end" : 260,

"hmm\_length" : 260,

"bit\_score" : 344.2,

"e\_value" : 4e-103,

"significance" : 1,

"clan" : "CL0014"

}

]

}

1. g216.t1

{

"\_id" : ObjectId("634e34882271a3ed989470b3"),

"gen\_id" : "g216.t1",

"seq" : "MATPLATSSGENDAAGQPTKMEGRGDQQKPQRVPLLKLFSFADSYDGLLMTLGSIGACIHGASVPVFFVFFGKMINIIGLAYLFPKEASHKVAKYSLDFLYLSVAILFSSWLEVACWMHTGERQAAKMRMAYLKSMLNQDISLFDTEATTGEVISAITSDIIVVQDALSEKFGNFLHYISRFIVGFTIGFASVWQISLVTLSIVPLIAFAGGLYAYVATGLIARVRKSYVKAGEIAEEVIGNVRTVQAFSGEDRAVKLYREALVNTYTYGRKAGLTKGLGLGSMHCVLFLSWALLVWFTSIIVHKNIANGGDSFTTMLNVVISGLRPGVLIALSLGQAAPDISAFVRAKAAAYPIFDMIGRNTVSKTSAKTGRKLNKVEGHIEFKDVFFSYPSRPNVTIFENLNLEIPSGKIVALVGGSGSGKSTVISLIERFYDPISGQILLDGNDIRGLDLQWLRQQIGLVNQEPALFATTIRENILYGKDDATLDEITRAAKLSEAISFINNLPEKFETQVGERGIQLSGGQKQRIAIARAIVKDPSILLLDEATSALDAESEKSVQEALDRAVVGRTTVVVAHRLSTIRSADVIAVVQNGKIVETGSHDELISNPNSAYSSLVQLQEAASLQRHPSEGPSLSRPHSIRYSRELSRTRTSFNTSFCSDRESVGRPGIDATEPMKKKPVSPKRLYSMIGPDWSYGVVGTIGALIAGALMPLFALGVTQALVAYYMDWDTTRHEVKKIATLFCGAACLTLIVHAIEHLSFGIMGERLTLRVRERMFSAILRNEIGWFDDVNNTSSMLASRLESDATLLRTIVVDRSTILLQNVGLVVGAFIISFILNWRLTLVVLAMYPLIISGHISEKLFMQGYGGNLSQAYLKANMLAAEAVSNIRTVAAFCSEEKVLDLYARELVEPSKRSFTRGQIAGLFYGVSQFFIFSSYGLALWYGSVLMEKEISSFKSVMKTFMVLIVTALAMGETLALAPDLLKGNQMVASVFEVMDRASQVSGDTGQELANVEGNIEFRNVQFSYPSRPDVIIFRDFNLKVHAGKSLALVGQSGSGKSSVLSLILRFYDPTAGNVMIDGRDIKKMNLKSLRKHIGLVQQEPALFATTIYENILYGKEGASESEVMEAAKLANAHTFISSLPDGYSTRVGERGVQLSGGQRQRVAIARAVLKNPETLLLDEATSALDVESERVVQQALDRLMRSRTSIVVAHRLSTIKNADQIAVIQSGKIVEQGTHSGLMENRNGAYYKLICLQQQEGK",

"blast\_predicted\_prot" : [

{

"\_id" : ObjectId("634ce47e7a4274e9d1076fa5"),

"blast\_prot\_pred\_id" : "g216.t1\_181",

"query\_accession" : "g216.t1",

"target\_accession" : "KJB21124.1",

"sequence\_identity" : 82.5,

"length" : 1258,

"mismatches" : 207,

"gap\_openings" : 2,

"query\_start" : 1,

"query\_end" : 1258,

"target\_start" : 1,

"target\_end" : 1245,

"e\_value" : 0,

"bit\_score" : 2011.9

},

{

"\_id" : ObjectId("634ce47e7a4274e9d1076fa6"),

"blast\_prot\_pred\_id" : "g216.t1\_182",

"query\_accession" : "g216.t1",

"target\_accession" : "XP\_022738803.1",

"sequence\_identity" : 83.6,

"length" : 1250,

"mismatches" : 193,

"gap\_openings" : 2,

"query\_start" : 9,

"query\_end" : 1258,

"target\_start" : 8,

"target\_end" : 1245,

"e\_value" : 0,

"bit\_score" : 2011.5

},

{

"\_id" : ObjectId("634ce47e7a4274e9d1076fa7"),

"blast\_prot\_pred\_id" : "g216.t1\_183",

"query\_accession" : "g216.t1",

"target\_accession" : "XP\_016728624.1",

"sequence\_identity" : 82.6,

"length" : 1258,

"mismatches" : 206,

"gap\_openings" : 2,

"query\_start" : 1,

"query\_end" : 1258,

"target\_start" : 1,

"target\_end" : 1245,

"e\_value" : 0,

"bit\_score" : 2010

},

{

"\_id" : ObjectId("634ce47e7a4274e9d1076fa8"),

"blast\_prot\_pred\_id" : "g216.t1\_184",

"query\_accession" : "g216.t1",

"target\_accession" : "XP\_012472187.1",

"sequence\_identity" : 84,

"length" : 1227,

"mismatches" : 188,

"gap\_openings" : 1,

"query\_start" : 32,

"query\_end" : 1258,

"target\_start" : 59,

"target\_end" : 1277,

"e\_value" : 0,

"bit\_score" : 2009.2

},

{

"\_id" : ObjectId("634ce47e7a4274e9d1076fa9"),

"blast\_prot\_pred\_id" : "g216.t1\_185",

"query\_accession" : "g216.t1",

"target\_accession" : "TYI34377.1",

"sequence\_identity" : 82.7,

"length" : 1257,

"mismatches" : 204,

"gap\_openings" : 3,

"query\_start" : 3,

"query\_end" : 1258,

"target\_start" : 2,

"target\_end" : 1245,

"e\_value" : 0,

"bit\_score" : 2006.9

}

],

"protein\_domains" : [

{

"\_id" : ObjectId("634e26f42271a3ed98945f59"),

"prot\_domain\_id" : "g216.t1\_278",

"seqid" : "g216.t1",

"alignment\_start" : 52,

"alignment\_end" : 322,

"envelope\_start" : 50,

"envelope\_end" : 325,

"hmm\_acc" : "PF00664.26",

"hmm\_name" : "ABC\_membrane",

"type" : "Family",

"hmm\_start" : 3,

"hmm\_end" : 271,

"hmm\_length" : 274,

"bit\_score" : 166,

"e\_value" : 1.5e-48,

"significance" : 1,

"clan" : "CL0241"

},

{

"\_id" : ObjectId("634e26f42271a3ed98945f5a"),

"prot\_domain\_id" : "g216.t1\_279",

"seqid" : "g216.t1",

"alignment\_start" : 401,

"alignment\_end" : 549,

"envelope\_start" : 400,

"envelope\_end" : 549,

"hmm\_acc" : "PF00005.30",

"hmm\_name" : "ABC\_tran",

"type" : "Domain",

"hmm\_start" : 2,

"hmm\_end" : 137,

"hmm\_length" : 137,

"bit\_score" : 122.9,

"e\_value" : 1.4e-35,

"significance" : 1,

"clan" : "CL0023"

},

{

"\_id" : ObjectId("634e26f42271a3ed98945f5b"),

"prot\_domain\_id" : "g216.t1\_280",

"seqid" : "g216.t1",

"alignment\_start" : 699,

"alignment\_end" : 970,

"envelope\_start" : 698,

"envelope\_end" : 970,

"hmm\_acc" : "PF00664.26",

"hmm\_name" : "ABC\_membrane",

"type" : "Family",

"hmm\_start" : 3,

"hmm\_end" : 274,

"hmm\_length" : 274,

"bit\_score" : 168.7,

"e\_value" : 2.2e-49,

"significance" : 1,

"clan" : "CL0241"

},

{

"\_id" : ObjectId("634e26f42271a3ed98945f5c"),

"prot\_domain\_id" : "g216.t1\_281",

"seqid" : "g216.t1",

"alignment\_start" : 1036,

"alignment\_end" : 1184,

"envelope\_start" : 1035,

"envelope\_end" : 1184,

"hmm\_acc" : "PF00005.30",

"hmm\_name" : "ABC\_tran",

"type" : "Domain",

"hmm\_start" : 2,

"hmm\_end" : 137,

"hmm\_length" : 137,

"bit\_score" : 120.3,

"e\_value" : 8.8e-35,

"significance" : 1,

"clan" : "CL0023"

}

]

}

1. g2485.t1

{

"\_id" : ObjectId("634e34882271a3ed98947990"),

"gen\_id" : "g2485.t1",

"seq" : "MGAAGEEKLKFCIDRGGTFTDVYAEIPGHPDGRVLKLLSVDPANYEDAPVEGIRRILEEYTGKEIPRAFKFPTDKIQWIRMGTTVATNALLERKGERIALCVTKGFKDLLQIGNQARPNIFDLTVSKPSNLYEAVIEVDERVELVMENEKVSMDSAALFEGVSGELLRISKPLNKEALKPLLKGLLDKGISSLAVVLMHSYTYPEHENAVEKLALELGFRHVSLSSALTPMVRAVPRGFTATVDAYLTPVIKEYLSGFLAKFDEGLGKVNVLFMQSDGGLAPESRFSGHKAVLSGPAGGVVGYSQTLFRLETEKPLIGFDMGGTSTDVSRYAGNYEQVLETQIAGAIIQAPQLDINTVAAGGGSKLKFQFGAFRVGPESVGAHPGPVCYRKGGELAVTDANLILGYVIPDYFPSIFGPKEDQPLDIKATREEFEKLSKEINTYRKSQDPSAKDMTVEEIAMGFINVANETMCRPIRQLTEMRGHETKNHALACFGGAGPQHACAIARSLGMKEVFIHRFCGILSAYGMGLADVVEEAQEPYSAVYTPESMLEVSRREALLQKEVKQKLQEQGFREENITTETYVNLRYEGTDTAIMVKGQVSGDGSGADYATEFVKLFQQEYGFKLQNRNIMICDVRVRGVGVTNIFKQRALEPATGASKIEGHYYVFFGDGWHNTPLFKLENLGFGHVIPGPAIIMNGNSTVIIEPKSKAIITKYGNIKIEVESEVTSVQMAENAADVVQLSIFNHRFMGIAEQMGRTLQRTSISTNIKERLDFSCALFSPDGGLVANAPHVPVHLGAMSSTVRWQLMHWGENLYEGDVLVTNHPCAGGSHLPDITVITPVFNNGKLVFFVASRGHHAEIGGITPGSMPPFSKSIREEGAAIKAFKLVEKGIFQEEGITKLLQFPCSDDPVHKIPGTRKLQDNLSDLRAQVAANQRGISLIKELIEHYGLKTVQAYMKYVQLNAEEAVREMLKSVAARVQAQWPNTIKSNPVVIEEEDYMDDGSIIHLKLTIDSTKGEATFDFSGTSPEVYGNWNAPEAVTAAAIIYCLRCLVDVDIPLNQGCLAPVKLHIPADSFLSPSEKAAVVGGNVLTSQRITDVVLLAFQACACSQGCMNNLTFGDNTFGYYETIGGGSGAGPSWDGTSAVQCHMTNTRMTDPEIFEQRYPVILHKFGLRKDSGGAGLHRGGDGLLREIEFKKPVMVSILSERRVHAPRGLKGGKDGACGANYLFTKDRRKIYLGGKNTVEVQPGEVLQILTPGGGGWGTPP",

"blast\_predicted\_prot" : [

{

"\_id" : ObjectId("634ce47e7a4274e9d10776f4"),

"blast\_prot\_pred\_id" : "g2485.t1\_2052",

"query\_accession" : "g2485.t1",

"target\_accession" : "XP\_021627540.1",

"sequence\_identity" : 84.4,

"length" : 1268,

"mismatches" : 196,

"gap\_openings" : 1,

"query\_start" : 1,

"query\_end" : 1266,

"target\_start" : 1,

"target\_end" : 1268,

"e\_value" : 0,

"bit\_score" : 2197.9

},

{

"\_id" : ObjectId("634ce47e7a4274e9d10776f5"),

"blast\_prot\_pred\_id" : "g2485.t1\_2053",

"query\_accession" : "g2485.t1",

"target\_accession" : "XP\_002262987.1",

"sequence\_identity" : 84.8,

"length" : 1267,

"mismatches" : 192,

"gap\_openings" : 1,

"query\_start" : 1,

"query\_end" : 1266,

"target\_start" : 1,

"target\_end" : 1267,

"e\_value" : 0,

"bit\_score" : 2194.9

},

{

"\_id" : ObjectId("634ce47e7a4274e9d10776f6"),

"blast\_prot\_pred\_id" : "g2485.t1\_2054",

"query\_accession" : "g2485.t1",

"target\_accession" : "KAH7571537.1",

"sequence\_identity" : 85,

"length" : 1268,

"mismatches" : 189,

"gap\_openings" : 1,

"query\_start" : 1,

"query\_end" : 1267,

"target\_start" : 92,

"target\_end" : 1359,

"e\_value" : 0,

"bit\_score" : 2194.5

},

{

"\_id" : ObjectId("634ce47e7a4274e9d10776f7"),

"blast\_prot\_pred\_id" : "g2485.t1\_2055",

"query\_accession" : "g2485.t1",

"target\_accession" : "XP\_010553896.1",

"sequence\_identity" : 84.9,

"length" : 1268,

"mismatches" : 190,

"gap\_openings" : 2,

"query\_start" : 1,

"query\_end" : 1267,

"target\_start" : 1,

"target\_end" : 1267,

"e\_value" : 0,

"bit\_score" : 2193.7

},

{

"\_id" : ObjectId("634ce47e7a4274e9d10776f8"),

"blast\_prot\_pred\_id" : "g2485.t1\_2056",

"query\_accession" : "g2485.t1",

"target\_accession" : "XP\_034693899.1",

"sequence\_identity" : 84.6,

"length" : 1268,

"mismatches" : 194,

"gap\_openings" : 1,

"query\_start" : 1,

"query\_end" : 1267,

"target\_start" : 1,

"target\_end" : 1268,

"e\_value" : 0,

"bit\_score" : 2193.7

}

],

"protein\_domains" : [

{

"\_id" : ObjectId("634e26f42271a3ed9894682d"),

"prot\_domain\_id" : "g2485.t1\_2538",

"seqid" : "g2485.t1",

"alignment\_start" : 10,

"alignment\_end" : 217,

"envelope\_start" : 10,

"envelope\_end" : 218,

"hmm\_acc" : "PF05378.16",

"hmm\_name" : "Hydant\_A\_N",

"type" : "Family",

"hmm\_start" : 1,

"hmm\_end" : 177,

"hmm\_length" : 178,

"bit\_score" : 216.4,

"e\_value" : 2.5e-64,

"significance" : 1,

"clan" : "CL0108"

},

{

"\_id" : ObjectId("634e26f42271a3ed9894682e"),

"prot\_domain\_id" : "g2485.t1\_2539",

"seqid" : "g2485.t1",

"alignment\_start" : 237,

"alignment\_end" : 536,

"envelope\_start" : 237,

"envelope\_end" : 537,

"hmm\_acc" : "PF01968.21",

"hmm\_name" : "Hydantoinase\_A",

"type" : "Family",

"hmm\_start" : 1,

"hmm\_end" : 290,

"hmm\_length" : 291,

"bit\_score" : 368.9,

"e\_value" : 1.8e-110,

"significance" : 1,

"clan" : "CL0108"

},

{

"\_id" : ObjectId("634e26f42271a3ed9894682f"),

"prot\_domain\_id" : "g2485.t1\_2540",

"seqid" : "g2485.t1",

"alignment\_start" : 559,

"alignment\_end" : 713,

"envelope\_start" : 548,

"envelope\_end" : 722,

"hmm\_acc" : "PF19278.2",

"hmm\_name" : "Hydant\_A\_C",

"type" : "Domain",

"hmm\_start" : 20,

"hmm\_end" : 182,

"hmm\_length" : 193,

"bit\_score" : 33.9,

"e\_value" : 2.1e-8,

"significance" : 1,

"clan" : "No\_clan"

},

{

"\_id" : ObjectId("634e26f42271a3ed98946830"),

"prot\_domain\_id" : "g2485.t1\_2541",

"seqid" : "g2485.t1",

"alignment\_start" : 738,

"alignment\_end" : 1267,

"envelope\_start" : 738,

"envelope\_end" : 1267,

"hmm\_acc" : "PF02538.17",

"hmm\_name" : "Hydantoinase\_B",

"type" : "Family",

"hmm\_start" : 1,

"hmm\_end" : 517,

"hmm\_length" : 517,

"bit\_score" : 720,

"e\_value" : 1.4e-216,

"significance" : 1,

"clan" : "No\_clan"

}

]

}

1. g2565.t1

{

"\_id" : ObjectId("634e34882271a3ed989479e0"),

"gen\_id" : "g2565.t1",

"seq" : "MLELPRNPQERVELGIDVTSCLTCSLPSLGHSMENYGYLFTLAIPPLKLSAALHFWLYSVQLDFVVALIDCSSSCTSHRLFESLVERSDKYVFFCVLFCSGTSTFMAKSQQYFYSKDGSASPTSARSREWGGPSRWTEYLGPDVASSFSSRNARNVGPDGQVRGSGAGKSLNIQWVLQLVEVGEGLMAKMYRLNQILDYPDPINHVFSPAFWKAGVFPNHPKICLLLSKKFPEHFSKLQLERVDKAALDALSDSAEIHLQSLEPWVQLLVDLMAFREQALRLILDLSSTVITLLPHQNSLILHAFMDLFCSFVRVNLLSEKIPRKMMLQIYNLLHAMSRSDRDCDFYHRLVQFIDSYDPPLKGLQEDLNFVSPRIGEVLEAVGPSIFLSTDTRKLRNEGFLSPYHPRYPDILTNSAHPMRAQDLANVTAYREWVLLGYLVCPDELLRVTSIDIALVVLKESLVLTLFRDEYVLLHEDYQLYVLPHILESKKMAKSGRTKQKEADLEYNVAKQWGYLENTGYGFSEVHEQALHSCGAIHRERRILLKQEIGRMVLFFTDQPSLLAPNIQMVFSALALAQSEVLWYFQHVGIASSKLKAANVISVDIDPSDPTIGFLLDGMDRLCCLVRKYIAAIRGYARSYLSSCAGRIRYLLGTPGIVALDLDANLKGLFQRIVQHLESIPKPQGENVSAMTCDLSDFRKDWFSILMIVTSSRSSINIRHLEKATVSTGKEGLLSEGNAAYNWSRCVDELESQLSKHGSLKKLYFYHQHLTAVFRNTMFGPEGRPQHCCAWLSVASSFPDCASLIMPEEVTKIGRDAVLYVESLIESIMGGLEGLINILDSEGGFGALESQLLPEQAAVYLNDSSRVPNLSMKSPKGIGGFPVPGHESYPENNNSVKMLEAAIQRLTNLCSVLNDMEPICVLNHVFVLREYMRECILGNFRKRFLVALKNENDLQRPSVLESLIRRHMSVVHLAEQHISMDLTQGIREILLQEAFSGPVSCLHLFDKPAEQPNGSAIEIVCNWYIDNIVKDVSSAGISFAPAHKCFKSTRPVGGYFAESVTDPKELQSFVRIFGGYGVDRLDRLMKEHTAALLNCIHTSLRANREVIEAIAGSMHSGDRIEREACLKQIVDLDTLIGFCVEAGQALAFDHLLAEATGAVLEDNAPLIYSLLSGIIKHIPEEVPQKSEITRIKGVANSIAASGDHDSGWIRSILEEVGGANDGSWGLLPYFFAAFMSSGIWNTTGFNIDTGGFTNNIHCLARCISAVIAGSEYVRLEREHKQKPLSIDPIDDSLDSDMHTRLSAEASIKSAMQLFVKFAAGIVLDSWGDANRTHLVAKLIFLDQLCEISPYLVRSTLESHVPYGILRSIYSQYYSNTPSTPLAMPNASPHHSPAVSLTHASPITKQPRGDSTPQHNIYDPVNTKGSSSHDHSYDADSGSVRSLRSSDARHRNTRRSGPLDSSSSRKLKIAEGSASSSTGPSPLPRFAASRSGPMSYK",

"blast\_predicted\_prot" : [

{

"\_id" : ObjectId("634ce47e7a4274e9d107774d"),

"blast\_prot\_pred\_id" : "g2565.t1\_2141",

"query\_accession" : "g2565.t1",

"target\_accession" : "OMO92344.1",

"sequence\_identity" : 83.1,

"length" : 1396,

"mismatches" : 220,

"gap\_openings" : 6,

"query\_start" : 106,

"query\_end" : 1496,

"target\_start" : 1,

"target\_end" : 1385,

"e\_value" : 0,

"bit\_score" : 2314.7

},

{

"\_id" : ObjectId("634ce47e7a4274e9d107774e"),

"blast\_prot\_pred\_id" : "g2565.t1\_2142",

"query\_accession" : "g2565.t1",

"target\_accession" : "OMO75005.1",

"sequence\_identity" : 82.7,

"length" : 1396,

"mismatches" : 226,

"gap\_openings" : 6,

"query\_start" : 106,

"query\_end" : 1496,

"target\_start" : 1,

"target\_end" : 1385,

"e\_value" : 0,

"bit\_score" : 2304.2

},

{

"\_id" : ObjectId("634ce47e7a4274e9d107774f"),

"blast\_prot\_pred\_id" : "g2565.t1\_2143",

"query\_accession" : "g2565.t1",

"target\_accession" : "XP\_021901589.1",

"sequence\_identity" : 82.8,

"length" : 1395,

"mismatches" : 227,

"gap\_openings" : 5,

"query\_start" : 106,

"query\_end" : 1496,

"target\_start" : 1,

"target\_end" : 1386,

"e\_value" : 0,

"bit\_score" : 2302.3

},

{

"\_id" : ObjectId("634ce47e7a4274e9d1077750"),

"blast\_prot\_pred\_id" : "g2565.t1\_2144",

"query\_accession" : "g2565.t1",

"target\_accession" : "XP\_021285978.1",

"sequence\_identity" : 83,

"length" : 1396,

"mismatches" : 222,

"gap\_openings" : 6,

"query\_start" : 106,

"query\_end" : 1496,

"target\_start" : 1,

"target\_end" : 1385,

"e\_value" : 0,

"bit\_score" : 2301.9

},

{

"\_id" : ObjectId("634ce47e7a4274e9d1077751"),

"blast\_prot\_pred\_id" : "g2565.t1\_2145",

"query\_accession" : "g2565.t1",

"target\_accession" : "EOX96917.1",

"sequence\_identity" : 83,

"length" : 1396,

"mismatches" : 222,

"gap\_openings" : 6,

"query\_start" : 106,

"query\_end" : 1496,

"target\_start" : 1,

"target\_end" : 1385,

"e\_value" : 0,

"bit\_score" : 2298.9

}

],

"protein\_domains" : [

{

"\_id" : ObjectId("634e26f42271a3ed9894689c"),

"prot\_domain\_id" : "g2565.t1\_2649",

"seqid" : "g2565.t1",

"alignment\_start" : 179,

"alignment\_end" : 1374,

"envelope\_start" : 176,

"envelope\_end" : 1375,

"hmm\_acc" : "PF09735.12",

"hmm\_name" : "Nckap1",

"type" : "Family",

"hmm\_start" : 12,

"hmm\_end" : 1118,

"hmm\_length" : 1119,

"bit\_score" : 1168,

"e\_value" : 0,

"significance" : 1,

"clan" : "CL0715"

}

]

}

1. g2724.t1

{

"\_id" : ObjectId("634e34882271a3ed98947a7f"),

"gen\_id" : "g2724.t1",

"seq" : "MGKNVVYDDEDEAELDEEEGEPVERDDVGDRDEDEDEEDEGPDDYEDDGFIVDDVDEEEVEEQEEEERQESDEERQKKKKKRKKKESEDVLDEDDYLLLQDNNINIQKKKFKRLKKARKDSEDGYGLSDEEFDGSGKDGLTAEEKIKYSLFGDEGPLEDMPEEEEPIEEEEDADIGEEDEMADFIVDEEMDANGEPVRRRRVKGKKPRSTYGASSYALQEAQDIFGDVENLLLSRKQELESSEWKERRLEDEFEPTILSEKYMTEKDEQIKTIDIPERMQISEESTGSPPADEISIEEESNWLYGQLSSLLKDPAQQGLDAGEGRDLCIKKDDIVKFLELHHVQKLDIPFIAMYRKEQCLSLLKDPDNDGTDDDEDNLDRKPTIKWHKVLWMIQDLDRKWLLLQKRKATLRSYYTKRFEEESRRIYEETRLNLNQHLFESILKSLKAAETEREVDDVDAKFNLHFPPGEVGIDEGQYKRPKRKSLYSICSKAGLWEVASKFGYSAEQLGLALSLEKLVDELEDAKETPEEMASNFTCAMFETPQGVLKGARHMAAVEISCEPSVKKYVRGIYMENAVVSTNPTPDGNVAIDSFHQFAGVKWLREKPLRKFQDAQWLLIQKAEEEKLLQVTFKLPENQLNRLNSDCNEHYLSVGVSKYAQLWNEQRKEILKDALNNFILPSMEKEARALLSSRAKNWLSFEYGKVLWNKVSVGPYQKKETDMTSDEEAAPRVMACCWGPGKPATTFVMLDSSGEVLDVLYAGSLTLRSQNINDQQRKKSDQQRVLKFMMDHQPHIVVLGAVNLSCTRLKDDIYEVIFKMVEENPRDVGQEMDELSIVYGDESLPQLYEHSRISSEQLLQQPGIVKRAVALGRYLQNPLAMVATLCGPGREILSWKLHPLENFLLPDEKYAMIEQVMVDATNQVGIDINLAAGHEWLFAPLQFISGLGPRKAANLQRSLVRAGAIFTRKDFVTTHGLGKKVFVNAVGFLRIRRSGLAANSSQFIDLLDDTRIHPESYGLAQELAKDVYDEDVRGDTNDDEDALEMAIEHVRDRPNLLKSLDLDAYLASKKRENKREIFNDIRRELMNGFQDWRKQYKEPSQDEEFLMISGETEETLAEGRIVQATVRRIQSGRAICVLDSGLTGILMKEDYSDDWRDIVELSDRLHEGDILTCKIKSILKNRYQIFLVCKESELRNNRFQHTQNVDPYYHEDPSSLQTEQDRARKEKELVKKHFKPRMIVHPRFQNITADEAMEYLADKEPGESIIRPSSRGPSFLTLTLKVYDGVYAHKDIVEGGKEHKDITSLLRIGKTLKIGEDTFEDLDEVMDRYVDPLVTHLKAMLSYRKFKRGKKAEVDELLRIEKAENPSRIVYSFGISHEHPGTFILTYIRSTNPHHEYIGLYPKGFKFRKRMFEDLDRLVAYFQRHIDDPVHESVPSIRSVAAMVPMRSPATGGSSGSGWGNQTNEGGWRGHSSDRDCSSTPRSRTGRNDYRNGGGRDGHPSGLPRPYGGRGRGRGSYSSNRGSGANDDGTPRWDSGGGWGSGGTVSGSGWGSGSGSGGGSGGGSGGGSGGGGGWGSSGVSGGGGGWGSSGVSGGGSGWGSSGGSGGGGGWSGGGGGWSSGGGSESGSGWGGGSGWGSESGGRKGGWGGESTDRKSGGW",

"blast\_predicted\_prot" : [

{

"\_id" : ObjectId("634ce47e7a4274e9d1077817"),

"blast\_prot\_pred\_id" : "g2724.t1\_2343",

"query\_accession" : "g2724.t1",

"target\_accession" : "XP\_010528633.1",

"sequence\_identity" : 80.9,

"length" : 1390,

"mismatches" : 249,

"gap\_openings" : 9,

"query\_start" : 121,

"query\_end" : 1503,

"target\_start" : 122,

"target\_end" : 1502,

"e\_value" : 0,

"bit\_score" : 2216.4

},

{

"\_id" : ObjectId("634ce47e7a4274e9d1077818"),

"blast\_prot\_pred\_id" : "g2724.t1\_2344",

"query\_accession" : "g2724.t1",

"target\_accession" : "XP\_007010711.2",

"sequence\_identity" : 80.5,

"length" : 1367,

"mismatches" : 252,

"gap\_openings" : 7,

"query\_start" : 150,

"query\_end" : 1508,

"target\_start" : 150,

"target\_end" : 1509,

"e\_value" : 0,

"bit\_score" : 2164.8

},

{

"\_id" : ObjectId("634ce47e7a4274e9d1077819"),

"blast\_prot\_pred\_id" : "g2724.t1\_2345",

"query\_accession" : "g2724.t1",

"target\_accession" : "EOY19521.1",

"sequence\_identity" : 80.5,

"length" : 1367,

"mismatches" : 252,

"gap\_openings" : 7,

"query\_start" : 150,

"query\_end" : 1508,

"target\_start" : 150,

"target\_end" : 1509,

"e\_value" : 0,

"bit\_score" : 2164.4

},

{

"\_id" : ObjectId("634ce47e7a4274e9d107781a"),

"blast\_prot\_pred\_id" : "g2724.t1\_2346",

"query\_accession" : "g2724.t1",

"target\_accession" : "XP\_021620863.1",

"sequence\_identity" : 80.5,

"length" : 1351,

"mismatches" : 257,

"gap\_openings" : 6,

"query\_start" : 114,

"query\_end" : 1458,

"target\_start" : 117,

"target\_end" : 1466,

"e\_value" : 0,

"bit\_score" : 2155.2

},

{

"\_id" : ObjectId("634ce47e7a4274e9d107781b"),

"blast\_prot\_pred\_id" : "g2724.t1\_2347",

"query\_accession" : "g2724.t1",

"target\_accession" : "XP\_021296325.1",

"sequence\_identity" : 80.1,

"length" : 1367,

"mismatches" : 257,

"gap\_openings" : 7,

"query\_start" : 150,

"query\_end" : 1508,

"target\_start" : 150,

"target\_end" : 1509,

"e\_value" : 0,

"bit\_score" : 2144

}

],

"protein\_domains" : [

{

"\_id" : ObjectId("634e26f42271a3ed98946983"),

"prot\_domain\_id" : "g2724.t1\_2880",

"seqid" : "g2724.t1",

"alignment\_start" : 32,

"alignment\_end" : 120,

"envelope\_start" : 30,

"envelope\_end" : 120,

"hmm\_acc" : "PF14632.9",

"hmm\_name" : "SPT6\_acidic",

"type" : "Family",

"hmm\_start" : 2,

"hmm\_end" : 89,

"hmm\_length" : 89,

"bit\_score" : 52.2,

"e\_value" : 6.5e-14,

"significance" : 1,

"clan" : "No\_clan"

},

{

"\_id" : ObjectId("634e26f42271a3ed98946984"),

"prot\_domain\_id" : "g2724.t1\_2881",

"seqid" : "g2724.t1",

"alignment\_start" : 728,

"alignment\_end" : 884,

"envelope\_start" : 725,

"envelope\_end" : 884,

"hmm\_acc" : "PF14639.9",

"hmm\_name" : "YqgF",

"type" : "Domain",

"hmm\_start" : 4,

"hmm\_end" : 150,

"hmm\_length" : 150,

"bit\_score" : 56.4,

"e\_value" : 3.1e-15,

"significance" : 1,

"clan" : "CL0580"

},

{

"\_id" : ObjectId("634e26f42271a3ed98946985"),

"prot\_domain\_id" : "g2724.t1\_2882",

"seqid" : "g2724.t1",

"alignment\_start" : 888,

"alignment\_end" : 989,

"envelope\_start" : 888,

"envelope\_end" : 990,

"hmm\_acc" : "PF14635.9",

"hmm\_name" : "HHH\_7",

"type" : "Domain",

"hmm\_start" : 1,

"hmm\_end" : 103,

"hmm\_length" : 104,

"bit\_score" : 76.6,

"e\_value" : 1.7e-21,

"significance" : 1,

"clan" : "CL0198"

},

{

"\_id" : ObjectId("634e26f42271a3ed98946986"),

"prot\_domain\_id" : "g2724.t1\_2883",

"seqid" : "g2724.t1",

"alignment\_start" : 1004,

"alignment\_end" : 1091,

"envelope\_start" : 1003,

"envelope\_end" : 1091,

"hmm\_acc" : "PF17674.4",

"hmm\_name" : "HHH\_9",

"type" : "Domain",

"hmm\_start" : 2,

"hmm\_end" : 70,

"hmm\_length" : 70,

"bit\_score" : 25.9,

"e\_value" : 0.000015,

"significance" : 1,

"clan" : "CL0198"

},

{

"\_id" : ObjectId("634e26f42271a3ed98946987"),

"prot\_domain\_id" : "g2724.t1\_2884",

"seqid" : "g2724.t1",

"alignment\_start" : 1206,

"alignment\_end" : 1423,

"envelope\_start" : 1206,

"envelope\_end" : 1423,

"hmm\_acc" : "PF14633.9",

"hmm\_name" : "SH2\_2",

"type" : "Domain",

"hmm\_start" : 1,

"hmm\_end" : 215,

"hmm\_length" : 215,

"bit\_score" : 248.3,

"e\_value" : 5.300000000000001e-74,

"significance" : 1,

"clan" : "CL0541"

}

]

}

1. g2847.t1

{

"\_id" : ObjectId("634e34882271a3ed98947afa"),

"gen\_id" : "g2847.t1",

"seq" : "MSRLDTSCFLLESDRFDPNAAFRSSELPVLTGNLSWNGTEERGLNLRRTRIEGIFTARVRENHRHGSDAEISSTNSCTFVSVVPIPPRASRHYRSFQSLLGRSVAENLSFVSFIICYLRLFSTNDVLCVKRHWILRQYHYGRPDLRNKTQKRVLAPNRELPPQNEHFLADFEQLKSQYTDQDQFRAMTESVLISLVVQCSSHAPRAEFLLFALQRLCSIGYINWDSFLPSLLSAVSSAEISLNQSSQATPSLPSTNLSQPGILPSSSGVPSSSNFHSTNPTSPLPSVHCIGSPVQSAVEPTSGTSFSPLRSSDISSTGLRTTRVSSSIRDAAISSLRQLCCKIISIGLEFHVKPGTHAEIFHHMINWLVNADQREQGMDECDGKLWRPEKTICKWLHSCIDVIWLLVDEDKSRVPFYELLRSNLQFIDNIPDDEALFTLIMEIHRRRDMIAMHMQMLDQHLHCPTFGTHRILSQTMANVNVEAVANLRYSPITYPSVLGEPLHGEDLAGAIPKGSLDWERAVRCIRHAIRTTPSPDWWKRVLVVAPCYRNTTQGPTPGAVFTSDMICEAIIDRIVELLKLTNSDTNCWHEWLVFSDIFFFLIKSGCTDFVDFVDKLVSRLLGGDNHILQTNHVAWLLAQIIRVELVINALTSDSRKVETTRKILSFHREDRNSDPNNPQSILLDFISSCQYLRIWSLNTSTREYVNNEQLQKGKQIDEWWRVSKGDRILDYMNMDDRSIGMFWVVSYTMAQPAGENVMNWLSSAGVAEVPGLNLQPNERLMMIREVSPLPMSLLSGFSINLCLKLALQMEDSLFVSQAVPSIAMVETYTRLLLLAPHSMFRSHFTHLAQRSPSLLSKPGVTLLVLEILNYRLLPLYRYQGKSKTLMYDVTKIVSTLKGKRGDHRVFRLAENLCMNLILSLRDFFSVKREGKGPTEFTETLNRVTIITLAITIKTRGIADADHVGYLQTMLEQILATSQHTWSENTLRYFPSLLRDALGRRVDKRGSTIQAWQQNETTVINQCKQLLSPSAEPHYFMTYISSSFPQHRQYLCAGAWILMQAHPENINNANLARVLREFSPEEVTANIYKMVDVLLHHIQMELQRVHSVQEIALKTCANLAFFFWTHELLPLDILLLALIDRDDNPHALRIVMSLLDRQELQQRVKSYCMNRGPPEHWLYTGIFKRNELQKALGNHLSGKDRYPTFFDDIAARLLPVIPLIIYRLIENDAIDAADRVLAMYSNFLAYHPLRFTFVRDILAYFYGHLPGKLIVRILKVLDLSKIPFSESFPQHINSTNPVICPPLDYFATLLLGLVNNVIPPLSSNSRSGSMVDTAHNPARAPHTKTPGASQPGPTTVSEGQKAFYQIQDPGTYTQLVLETAVIEILSLPVPASQIVSSLIQIIVNIQPTLVQSSNGPHGSSSGAGQGSVLPTSPSGGSADSISASRSTPCGSGINTASFVSRSGYTCQQLSCLLIQACGLLLAQLPPDFHMQLYMEASRIIKDTWWLTDEKRSLGELDSAVGYALLDPTWAAQDNTSTAIGNVVALLHSFFSNLPQEWLDSTHIIVKNLSPVTSVAMLRVAFRIMGPLLPRLANAHMLFNNTLSLLLNAMVDVFGKNIQPRTPVEASEIKDLIDFLHHVVHYEGQGGPVQPNSKPRPEVMVLIGRAAEKLRPDVQHLLSHLKPDVNSSIYAATHPNLARNAS",

"blast\_predicted\_prot" : [

{

"\_id" : ObjectId("634ce47e7a4274e9d1077898"),

"blast\_prot\_pred\_id" : "g2847.t1\_2472",

"query\_accession" : "g2847.t1",

"target\_accession" : "XP\_021293610.1",

"sequence\_identity" : 81.2,

"length" : 1560,

"mismatches" : 289,

"gap\_openings" : 4,

"query\_start" : 144,

"query\_end" : 1700,

"target\_start" : 48,

"target\_end" : 1606,

"e\_value" : 0,

"bit\_score" : 2543.8

},

{

"\_id" : ObjectId("634ce47e7a4274e9d1077899"),

"blast\_prot\_pred\_id" : "g2847.t1\_2473",

"query\_accession" : "g2847.t1",

"target\_accession" : "KDO64560.1",

"sequence\_identity" : 81.1,

"length" : 1559,

"mismatches" : 289,

"gap\_openings" : 4,

"query\_start" : 145,

"query\_end" : 1700,

"target\_start" : 84,

"target\_end" : 1639,

"e\_value" : 0,

"bit\_score" : 2541.5

},

{

"\_id" : ObjectId("634ce47e7a4274e9d107789a"),

"blast\_prot\_pred\_id" : "g2847.t1\_2474",

"query\_accession" : "g2847.t1",

"target\_accession" : "KDO64561.1",

"sequence\_identity" : 81.1,

"length" : 1559,

"mismatches" : 289,

"gap\_openings" : 4,

"query\_start" : 145,

"query\_end" : 1700,

"target\_start" : 2,

"target\_end" : 1557,

"e\_value" : 0,

"bit\_score" : 2541.5

},

{

"\_id" : ObjectId("634ce47e7a4274e9d107789b"),

"blast\_prot\_pred\_id" : "g2847.t1\_2475",

"query\_accession" : "g2847.t1",

"target\_accession" : "KDO64559.1",

"sequence\_identity" : 81.1,

"length" : 1559,

"mismatches" : 289,

"gap\_openings" : 4,

"query\_start" : 145,

"query\_end" : 1700,

"target\_start" : 79,

"target\_end" : 1634,

"e\_value" : 0,

"bit\_score" : 2541.5

},

{

"\_id" : ObjectId("634ce47e7a4274e9d107789c"),

"blast\_prot\_pred\_id" : "g2847.t1\_2476",

"query\_accession" : "g2847.t1",

"target\_accession" : "XP\_006469228.1",

"sequence\_identity" : 81.1,

"length" : 1559,

"mismatches" : 289,

"gap\_openings" : 4,

"query\_start" : 145,

"query\_end" : 1700,

"target\_start" : 79,

"target\_end" : 1634,

"e\_value" : 0,

"bit\_score" : 2540.8

}

],

"protein\_domains" : [

{

"\_id" : ObjectId("634e26f42271a3ed98946a39"),

"prot\_domain\_id" : "g2847.t1\_3062",

"seqid" : "g2847.t1",

"alignment\_start" : 747,

"alignment\_end" : 1309,

"envelope\_start" : 565,

"envelope\_end" : 1321,

"hmm\_acc" : "PF11573.11",

"hmm\_name" : "Med23",

"type" : "Family",

"hmm\_start" : 489,

"hmm\_end" : 1070,

"hmm\_length" : 1304,

"bit\_score" : 96.5,

"e\_value" : 1e-27,

"significance" : 1,

"clan" : "No\_clan"

}

]

}

1. g3154.t1

{

"\_id" : ObjectId("634e34882271a3ed98947c2d"),

"gen\_id" : "g3154.t1",

"seq" : "MAEGAPEPNAVPEAEKKTEQSLPFFKLFSFADRYDWLLMVCGSVGAVIHGSSMPAFFLLFGEMVNGFGKNQFDLQKMTNEVSKAIPSLIIFVPGYALYFVYLGLVVCFSSYAEIACWMYTGERQVSTLRKKYLEAVLKQDVGFFDTDARTGDIVFSVSTDTLLVQDAISEKVGNFIHYLSTFLAGLVVGFVSAWKLALLSVAVIPGIAFAGGLYAYTLTGITSKSRESYANAGIIAEQAIAQVRTVYSYVGERKTLDAYSDAIQYTLKLGYKAGMAKGLGLGCTYGIACMSWALVFWYAGVFIRNGQTDGGKAFTAIFSAIVGGMSLGQSFSNLGAFSKGKAAGYKLMEIISQQPSIIQDPRDGKYLSGVNGNIEFKDVTFSYPSRPDFMIFRKFSMYFPAGKTVAVVGGSGSGKSTVVALIERFYDPNEGQILLDDVDLKTLQLKWLRDQIGLVNQEPALFATTILENILYGKPDAIMAEVESAASAANAHSFITLLPNGYHTQVGERGVQLSGGQKQRIAIARAMLKNPKILLLDEATSALDAGSESIVQEALDRLMVGRTTVVVAHRLSTIRNVDSIAVIQQGQVVETGTHEQLISKSGGAYASLIRFQEMVGTRDFSNPSTKRSRSSRLSNSLSTRSISLRSGSLRNLSYSYSTGADGRIEMVSNAETGRKTRVPDGYFLRLLKLNAPEWPYSIMGAIGSVLSGFIGPTFAIVMGNMIEVFYYSNYASMERKTKEYVFVYIGAGIYAVVAYLIQHYFFSIMGENLTTRVRRMMLSAILRNEVGWFDEEEHNSSLLAARLATDAADVKSAMAERISVILQNMTSLLTSFIVAFIVEWRVSVLILGTFPLLVLANFAQSVSSCLVSAELTGSLFFTCIGAGALIALLTGCIRAARLKLQVSVVIGYLVLLSWRSEIGFAGDTAKAHAKTSMIAGEGVSNIRTVAAFNAQNKILSLFSNELRVPRLGSLRRSQTAGSLFGLSQLALYASEALILWYGAHLVSKGLSTFSKVIKVFVVLVVTANSVAETVSLAPEIIRGGEAVGSVFSILDRFTRIDPDDPESEPAENLRGEIELRHVDFSYPSRPEVMVFKDFNLRIRAGKSQALVGASGSGKSSVIALIERFYDPTAGKVMIDGKDIRRFNLKSLRLRVGLVQQEPALFATTIFENIAYGKEGSTESEVIEAARASNAHSFISALPGGYKTPVGERGVQLSGGQKQRIAIARAVLKNPTILLLDEATSALDAESECVLQEALERLMRGRTTVLVAHRLSTIRGVDSIGVVQDGRILEQGSHSELISRPDGAYSRLLQLQTHRV",

"blast\_predicted\_prot" : [

{

"\_id" : ObjectId("634ce47e7a4274e9d10779c8"),

"blast\_prot\_pred\_id" : "g3154.t1\_2776",

"query\_accession" : "g3154.t1",

"target\_accession" : "XP\_010528330.1",

"sequence\_identity" : 84.7,

"length" : 1317,

"mismatches" : 135,

"gap\_openings" : 5,

"query\_start" : 1,

"query\_end" : 1315,

"target\_start" : 5,

"target\_end" : 1256,

"e\_value" : 0,

"bit\_score" : 2097

},

{

"\_id" : ObjectId("634ce47e7a4274e9d10779c9"),

"blast\_prot\_pred\_id" : "g3154.t1\_2777",

"query\_accession" : "g3154.t1",

"target\_accession" : "XP\_022741367.1",

"sequence\_identity" : 84.7,

"length" : 1315,

"mismatches" : 135,

"gap\_openings" : 5,

"query\_start" : 1,

"query\_end" : 1315,

"target\_start" : 1,

"target\_end" : 1249,

"e\_value" : 0,

"bit\_score" : 2093.9

},

{

"\_id" : ObjectId("634ce47e7a4274e9d10779ca"),

"blast\_prot\_pred\_id" : "g3154.t1\_2778",

"query\_accession" : "g3154.t1",

"target\_accession" : "XP\_031404499.1",

"sequence\_identity" : 84.1,

"length" : 1305,

"mismatches" : 142,

"gap\_openings" : 4,

"query\_start" : 11,

"query\_end" : 1315,

"target\_start" : 18,

"target\_end" : 1257,

"e\_value" : 0,

"bit\_score" : 2088.9

},

{

"\_id" : ObjectId("634ce47e7a4274e9d10779cb"),

"blast\_prot\_pred\_id" : "g3154.t1\_2779",

"query\_accession" : "g3154.t1",

"target\_accession" : "KAH7554605.1",

"sequence\_identity" : 85,

"length" : 1304,

"mismatches" : 130,

"gap\_openings" : 4,

"query\_start" : 12,

"query\_end" : 1315,

"target\_start" : 13,

"target\_end" : 1251,

"e\_value" : 0,

"bit\_score" : 2083.5

},

{

"\_id" : ObjectId("634ce47e7a4274e9d10779cc"),

"blast\_prot\_pred\_id" : "g3154.t1\_2780",

"query\_accession" : "g3154.t1",

"target\_accession" : "XP\_022741366.1",

"sequence\_identity" : 82.6,

"length" : 1352,

"mismatches" : 143,

"gap\_openings" : 6,

"query\_start" : 1,

"query\_end" : 1315,

"target\_start" : 1,

"target\_end" : 1297,

"e\_value" : 0,

"bit\_score" : 2083.1

}

],

"protein\_domains" : [

{

"\_id" : ObjectId("634e26f42271a3ed98946bd8"),

"prot\_domain\_id" : "g3154.t1\_3477",

"seqid" : "g3154.t1",

"alignment\_start" : 42,

"alignment\_end" : 324,

"envelope\_start" : 39,

"envelope\_end" : 325,

"hmm\_acc" : "PF00664.26",

"hmm\_name" : "ABC\_membrane",

"type" : "Family",

"hmm\_start" : 4,

"hmm\_end" : 273,

"hmm\_length" : 274,

"bit\_score" : 182.8,

"e\_value" : 1.0999999999999999e-53,

"significance" : 1,

"clan" : "CL0241"

},

{

"\_id" : ObjectId("634e26f42271a3ed98946bd9"),

"prot\_domain\_id" : "g3154.t1\_3478",

"seqid" : "g3154.t1",

"alignment\_start" : 394,

"alignment\_end" : 541,

"envelope\_start" : 392,

"envelope\_end" : 541,

"hmm\_acc" : "PF00005.30",

"hmm\_name" : "ABC\_tran",

"type" : "Domain",

"hmm\_start" : 3,

"hmm\_end" : 137,

"hmm\_length" : 137,

"bit\_score" : 116.5,

"e\_value" : 1.3000000000000001e-33,

"significance" : 1,

"clan" : "CL0023"

},

{

"\_id" : ObjectId("634e26f42271a3ed98946bda"),

"prot\_domain\_id" : "g3154.t1\_3479",

"seqid" : "g3154.t1",

"alignment\_start" : 698,

"alignment\_end" : 1022,

"envelope\_start" : 697,

"envelope\_end" : 1024,

"hmm\_acc" : "PF00664.26",

"hmm\_name" : "ABC\_membrane",

"type" : "Family",

"hmm\_start" : 2,

"hmm\_end" : 272,

"hmm\_length" : 274,

"bit\_score" : 143.1,

"e\_value" : 1.3999999999999998e-41,

"significance" : 1,

"clan" : "CL0241"

},

{

"\_id" : ObjectId("634e26f42271a3ed98946bdb"),

"prot\_domain\_id" : "g3154.t1\_3480",

"seqid" : "g3154.t1",

"alignment\_start" : 1092,

"alignment\_end" : 1240,

"envelope\_start" : 1091,

"envelope\_end" : 1240,

"hmm\_acc" : "PF00005.30",

"hmm\_name" : "ABC\_tran",

"type" : "Domain",

"hmm\_start" : 2,

"hmm\_end" : 137,

"hmm\_length" : 137,

"bit\_score" : 112.4,

"e\_value" : 2.5e-32,

"significance" : 1,

"clan" : "CL0023"

}

]

}

1. g3393.t1

{

"\_id" : ObjectId("634e34882271a3ed98947d1c"),

"gen\_id" : "g3393.t1",

"seq" : "MGSKFKPSRSNLSESSDATDSQKPPMPPTVTFARRTSSGRYVSYSRDDLDSEIGSVDFMNYTVHIPPTPDNQPMDPSISQKVEEQYVSNSLFSGGFNNVTRAHLMDKVIDSDTRHHPQMAGAKGSSCAIPGCDGKVMSDERGQDILPCECDFKICRDCYIDAVKSGGQTCPGCKEPYKNPDLNDLSGDGRPPLTLPLTSGNSRMERRLSLMKSTKSGIIRNQAGEFDHNRWLFETRGTYGYGNAMWTNDDYGDGKDDDIVEPAELMSKPWRPLTRKLKVPAGIISPYRQVKHGTYIPSERRRSSGIYNLALHEHLDLENSTVGRVHLRPKLLIVIRIVVLALFLTWRVTHPNTDAVWLWGMSVVCEVWFAFSWLLDQLPKLCPVMRATDLDVLEEKFETPTPSNPTGKSDLPGIDIFVSTADPEKEPPLVTANTILSILAADYPVEKLSCYVSDDGGALLTFEAMAEAASFANMWVPFCRKHAIEPRNPESYFSLKKDPYKNKVKTDFVKDRRRIKREYDEFKVRINSLPDSIRRRSDAYHAREEIKAMKMQRENMRDDETVEPIKIPKATWMADGTHWPGTWLHPSADHSRGDHAGIIQVMLKPPSDEQLYGTPEGFLDFSEVDIRLPLLVYVSREKRPGYDHNKKAGAMNALVRASAVMSNGPFILNLDCDHYIYNSKALREGMCFMMDRGGDRICYVQFPQRFEGIDPSDRYANHNTVFFDVNMRALDGLMGPFYVGTGCLFRRIALYGFDPPRSKDHHPSFCSCCCPRRKKRTHVPEENRALRMGDSDDDGMELNLHPKKFGNSTFLIESIPVAEFQGRPLADHPAVKNGRPPGALTIPRQLLDASTVAEAIGVVSCWYEDKTEWGTRIGWIYGSVTEDVVTGYRMHNRGWKSVYCVTKRDAFRGTAPINLTDRLHQVLRWATGSVEIFFSRNNAFLASPKMKILQRIAYLNVGIYPFTSIFLIVYCFLPALSLFSGQFIVQTLSVTFLVYLLAISVTLCLLALLEIKWSKISLEEWWRNEQFWLIGGTSAHLAAVLQGLLKVIAGIEISFTLTSKSAGDDMDDEFADLYVVRWTSLMLPPVTIIMVNLIAIAVGFSRAIYAVIPQWSRLLGGVFFAFWVLAHLYPFAKGLMGRRGRTPTIVYVWSGLIAITISLLWVAINPPEGGRNNNEIGGSFTFP",

"blast\_predicted\_prot" : [

{

"\_id" : ObjectId("634ce47e7a4274e9d1077aef"),

"blast\_prot\_pred\_id" : "g3393.t1\_3071",

"query\_accession" : "g3393.t1",

"target\_accession" : "XP\_010558640.1",

"sequence\_identity" : 85,

"length" : 1188,

"mismatches" : 126,

"gap\_openings" : 8,

"query\_start" : 1,

"query\_end" : 1183,

"target\_start" : 1,

"target\_end" : 1141,

"e\_value" : 0,

"bit\_score" : 2073.5

},

{

"\_id" : ObjectId("634ce47e7a4274e9d1077af0"),

"blast\_prot\_pred\_id" : "g3393.t1\_3072",

"query\_accession" : "g3393.t1",

"target\_accession" : "XP\_010549830.1",

"sequence\_identity" : 84.9,

"length" : 1188,

"mismatches" : 128,

"gap\_openings" : 7,

"query\_start" : 1,

"query\_end" : 1183,

"target\_start" : 1,

"target\_end" : 1142,

"e\_value" : 0,

"bit\_score" : 2061.6

},

{

"\_id" : ObjectId("634ce47e7a4274e9d1077af1"),

"blast\_prot\_pred\_id" : "g3393.t1\_3073",

"query\_accession" : "g3393.t1",

"target\_accession" : "EOY10080.1",

"sequence\_identity" : 82.7,

"length" : 1191,

"mismatches" : 182,

"gap\_openings" : 10,

"query\_start" : 1,

"query\_end" : 1183,

"target\_start" : 1,

"target\_end" : 1175,

"e\_value" : 0,

"bit\_score" : 2017.7

},

{

"\_id" : ObjectId("634ce47e7a4274e9d1077af2"),

"blast\_prot\_pred\_id" : "g3393.t1\_3074",

"query\_accession" : "g3393.t1",

"target\_accession" : "XP\_002520032.1",

"sequence\_identity" : 82.8,

"length" : 1190,

"mismatches" : 151,

"gap\_openings" : 8,

"query\_start" : 1,

"query\_end" : 1183,

"target\_start" : 1,

"target\_end" : 1143,

"e\_value" : 0,

"bit\_score" : 2014.6

},

{

"\_id" : ObjectId("634ce47e7a4274e9d1077af3"),

"blast\_prot\_pred\_id" : "g3393.t1\_3075",

"query\_accession" : "g3393.t1",

"target\_accession" : "OMP08846.1",

"sequence\_identity" : 82.6,

"length" : 1190,

"mismatches" : 153,

"gap\_openings" : 9,

"query\_start" : 1,

"query\_end" : 1183,

"target\_start" : 1,

"target\_end" : 1143,

"e\_value" : 0,

"bit\_score" : 2011.5

}

],

"protein\_domains" : [

{

"\_id" : ObjectId("634e26f42271a3ed98946d2f"),

"prot\_domain\_id" : "g3393.t1\_3820",

"seqid" : "g3393.t1",

"alignment\_start" : 129,

"alignment\_end" : 178,

"envelope\_start" : 129,

"envelope\_end" : 178,

"hmm\_acc" : "PF14570.9",

"hmm\_name" : "zf-RING\_4",

"type" : "Domain",

"hmm\_start" : 1,

"hmm\_end" : 48,

"hmm\_length" : 48,

"bit\_score" : 59.4,

"e\_value" : 2.3e-16,

"significance" : 1,

"clan" : "CL0229"

},

{

"\_id" : ObjectId("634e26f42271a3ed98946d30"),

"prot\_domain\_id" : "g3393.t1\_3821",

"seqid" : "g3393.t1",

"alignment\_start" : 414,

"alignment\_end" : 1168,

"envelope\_start" : 414,

"envelope\_end" : 1177,

"hmm\_acc" : "PF03552.17",

"hmm\_name" : "Cellulose\_synt",

"type" : "Family",

"hmm\_start" : 1,

"hmm\_end" : 711,

"hmm\_length" : 722,

"bit\_score" : 1246.6,

"e\_value" : 0,

"significance" : 1,

"clan" : "CL0110"

}

]

}

1. g3633.t1

{

"\_id" : ObjectId("634e34882271a3ed98947e0c"),

"gen\_id" : "g3633.t1",

"seq" : "MLISVGDRSTNANEFLARDGELCEKDLAGRPGMARKRRRIKHPPPSMHGPLHPHLSLLFLAPFLSRFLLLSHASPIPASSSAVLALPIASCRAVVSVLELGWQAFRNKGIASRTPDVNRLALFEMRDSQWGTPVNIIVGSHVWVEDPGVAFIDGQVKSITGQQAEVRTTDGRSVTANLSNLYPKDMEAPAGGVDDMTKLSYLHEPGVLENLKTRYELNDIYTYTGNILIAINPFQKLPHIYDAHMMQQYKGAPFGELSPHVFAVADQAFRAMVNEGKSNSILVSGESGAGKTETTKMLMRYLAYLGGRTGTEGRTVEQQVLESNPVLEAFGNAKTVRNNNSSRFGKFVEIQFNKQERISGAAIRTYLLERSRVCQISTPERNYHCFYLLCAAPQEEIERYKLGNPKTFHYLNQSNCLELAGVSDAHDYLATRRAMDIVGISEKEQEAIFRVVAAILHLGNIEFTKGKEVDSSVLKDDKAKFHLETTAELLMCDPQALEDALCKRVMVTPEEVIKRSLDPQSAAINRDGLAKTIYSRLFDWLVDKINDSIGQDSSSKHLIGVLDIYGFESFKNNSFEQFCINFTNEKLQQHFNQHVFKMEQEEYTKEEIDWSYIEFVDNQDVLDLIEKKPGGIIALLDEACMFPKSTHETFSNKLYQTFKSHKRFIKPKLSRTDFTISHYAGEVTYQSDLFLDKNKDYVVPEHQDLLSASKCSFVSGLFPPPPEETSKSSKFSSIGSRFKLQLQQLMDTLNSTEPHYIRCVKPNNLLMPAIFENVNIMQQLRCGGVLEAIRISCAGYPTRKPFYEFINRFGLLAPDALEGNYNEKAACRKILDKVGLKGYQIGKTKVFLRAGQMAELDARRAQVLNGAAKIIQRRIRTHQARRQFIKLLDAAISLQSLCRGRLACTQYMNLRRQAAAVKIQKHARRHQARKKYRCLHLSALAIQTALRAMAAHKEYRFRKQTQAATIIQARWRCHKKTMFYKRLKKGVVAAQTRWRGRLARRELRKLKMAARETGALKEAKDKLEKKVEELTYRVQLEKRLRLDLEEAKAREIAKYQSSLEEMQKKVDEANALVVKEREAAKKAIEEAPPVIKETEVLVEDTKKVESLTKEVEILKASLASEKEKAEAATNKYKEAQESVEEKMKKLEETQKRVQQLQESLTRLEEKANNLESENKVLRQQALSMAPNKMLSGRSRSILQRGGESGHFSVDARHNVDLHSHSMSHRDVSEVEDKPQKSLNEKQQENQELLIRCITQHLGFQGGRPIAACIIYKCLLHWRSFEVERTSIFDKIIQTIGQAIETQDNNEILAYWLSNASTLLLLLQRTLKASGAAGMAPQRRRSSASLFGRMTQSFRGTPQGVNLAMINGGISGTDTLRQVEAKYPALLFKQQLTAYVEKIYGMIRDNLKKEISPLLGLCIQAPRTSRASLVKGASRSVANTAAQQALIAHWQGIEKGLTNFLNTLKANHVPPFLVRKVFTQIFSFINVQLFNSLLLRRECCSFSNGEYVKAGLAELEHWCYKATDEYAGSAWDELKHIRQAIGFLVIHQKPKKTLDEISHDLCPVLSIQQLYRISTMYWDDKYGTHSVSPDVIANMRVLMTEDSNNAVSSSFLLDDDSSIPFSVDDLSKSMSQIDISDIEPPALIRDHSGFSFLLPCSDNH",

"blast\_predicted\_prot" : [

{

"\_id" : ObjectId("634ce47e7a4274e9d1077bb5"),

"blast\_prot\_pred\_id" : "g3633.t1\_3269",

"query\_accession" : "g3633.t1",

"target\_accession" : "XP\_010553755.1",

"sequence\_identity" : 89,

"length" : 1529,

"mismatches" : 166,

"gap\_openings" : 2,

"query\_start" : 131,

"query\_end" : 1657,

"target\_start" : 2,

"target\_end" : 1530,

"e\_value" : 0,

"bit\_score" : 2262.6

},

{

"\_id" : ObjectId("634ce47e7a4274e9d1077bb6"),

"blast\_prot\_pred\_id" : "g3633.t1\_3270",

"query\_accession" : "g3633.t1",

"target\_accession" : "XP\_010531810.1",

"sequence\_identity" : 88.4,

"length" : 1530,

"mismatches" : 174,

"gap\_openings" : 3,

"query\_start" : 131,

"query\_end" : 1657,

"target\_start" : 2,

"target\_end" : 1531,

"e\_value" : 0,

"bit\_score" : 2251.5

},

{

"\_id" : ObjectId("634ce47e7a4274e9d1077bb7"),

"blast\_prot\_pred\_id" : "g3633.t1\_3271",

"query\_accession" : "g3633.t1",

"target\_accession" : "XP\_024015736.1",

"sequence\_identity" : 89.1,

"length" : 1528,

"mismatches" : 163,

"gap\_openings" : 2,

"query\_start" : 131,

"query\_end" : 1657,

"target\_start" : 2,

"target\_end" : 1526,

"e\_value" : 0,

"bit\_score" : 2248.4

},

{

"\_id" : ObjectId("634ce47e7a4274e9d1077bb8"),

"blast\_prot\_pred\_id" : "g3633.t1\_3272",

"query\_accession" : "g3633.t1",

"target\_accession" : "XP\_006442753.2",

"sequence\_identity" : 89.3,

"length" : 1529,

"mismatches" : 161,

"gap\_openings" : 2,

"query\_start" : 131,

"query\_end" : 1657,

"target\_start" : 3,

"target\_end" : 1531,

"e\_value" : 0,

"bit\_score" : 2231.8

},

{

"\_id" : ObjectId("634ce47e7a4274e9d1077bb9"),

"blast\_prot\_pred\_id" : "g3633.t1\_3273",

"query\_accession" : "g3633.t1",

"target\_accession" : "XP\_015388647.1",

"sequence\_identity" : 89.3,

"length" : 1529,

"mismatches" : 162,

"gap\_openings" : 2,

"query\_start" : 131,

"query\_end" : 1657,

"target\_start" : 3,

"target\_end" : 1531,

"e\_value" : 0,

"bit\_score" : 2231.4

}

],

"protein\_domains" : [

{

"\_id" : ObjectId("634e26f42271a3ed98946e8d"),

"prot\_domain\_id" : "g3633.t1\_4170",

"seqid" : "g3633.t1",

"alignment\_start" : 138,

"alignment\_end" : 178,

"envelope\_start" : 137,

"envelope\_end" : 181,

"hmm\_acc" : "PF02736.22",

"hmm\_name" : "Myosin\_N",

"type" : "Domain",

"hmm\_start" : 2,

"hmm\_end" : 42,

"hmm\_length" : 45,

"bit\_score" : 40.3,

"e\_value" : 2e-10,

"significance" : 1,

"clan" : "CL0010"

},

{

"\_id" : ObjectId("634e26f42271a3ed98946e8e"),

"prot\_domain\_id" : "g3633.t1\_4171",

"seqid" : "g3633.t1",

"alignment\_start" : 193,

"alignment\_end" : 849,

"envelope\_start" : 193,

"envelope\_end" : 849,

"hmm\_acc" : "PF00063.24",

"hmm\_name" : "Myosin\_head",

"type" : "Domain",

"hmm\_start" : 1,

"hmm\_end" : 677,

"hmm\_length" : 677,

"bit\_score" : 846.3,

"e\_value" : 1.6e-254,

"significance" : 1,

"clan" : "CL0023"

},

{

"\_id" : ObjectId("634e26f42271a3ed98946e8f"),

"prot\_domain\_id" : "g3633.t1\_4172",

"seqid" : "g3633.t1",

"alignment\_start" : 867,

"alignment\_end" : 884,

"envelope\_start" : 865,

"envelope\_end" : 885,

"hmm\_acc" : "PF00612.30",

"hmm\_name" : "IQ",

"type" : "Motif",

"hmm\_start" : 3,

"hmm\_end" : 20,

"hmm\_length" : 21,

"bit\_score" : 13.4,

"e\_value" : 0.054,

"significance" : 1,

"clan" : "CL0220"

},

{

"\_id" : ObjectId("634e26f42271a3ed98946e90"),

"prot\_domain\_id" : "g3633.t1\_4173",

"seqid" : "g3633.t1",

"alignment\_start" : 914,

"alignment\_end" : 933,

"envelope\_start" : 913,

"envelope\_end" : 933,

"hmm\_acc" : "PF00612.30",

"hmm\_name" : "IQ",

"type" : "Motif",

"hmm\_start" : 2,

"hmm\_end" : 21,

"hmm\_length" : 21,

"bit\_score" : 18.3,

"e\_value" : 0.0015,

"significance" : 1,

"clan" : "CL0220"

},

{

"\_id" : ObjectId("634e26f42271a3ed98946e91"),

"prot\_domain\_id" : "g3633.t1\_4174",

"seqid" : "g3633.t1",

"alignment\_start" : 961,

"alignment\_end" : 975,

"envelope\_start" : 961,

"envelope\_end" : 976,

"hmm\_acc" : "PF00612.30",

"hmm\_name" : "IQ",

"type" : "Motif",

"hmm\_start" : 1,

"hmm\_end" : 15,

"hmm\_length" : 21,

"bit\_score" : 12.2,

"e\_value" : 0.14,

"significance" : 1,

"clan" : "CL0220"

},

{

"\_id" : ObjectId("634e26f42271a3ed98946e92"),

"prot\_domain\_id" : "g3633.t1\_4175",

"seqid" : "g3633.t1",

"alignment\_start" : 1476,

"alignment\_end" : 1579,

"envelope\_start" : 1475,

"envelope\_end" : 1580,

"hmm\_acc" : "PF01843.22",

"hmm\_name" : "DIL",

"type" : "Repeat",

"hmm\_start" : 2,

"hmm\_end" : 104,

"hmm\_length" : 105,

"bit\_score" : 85.8,

"e\_value" : 2.1e-24,

"significance" : 1,

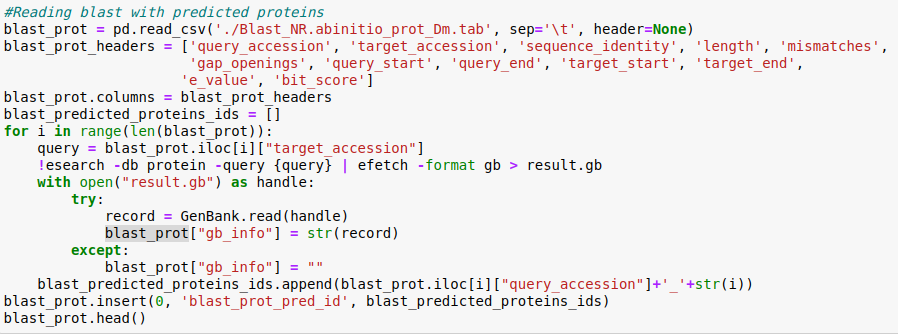
"clan" : "CL0020"

}

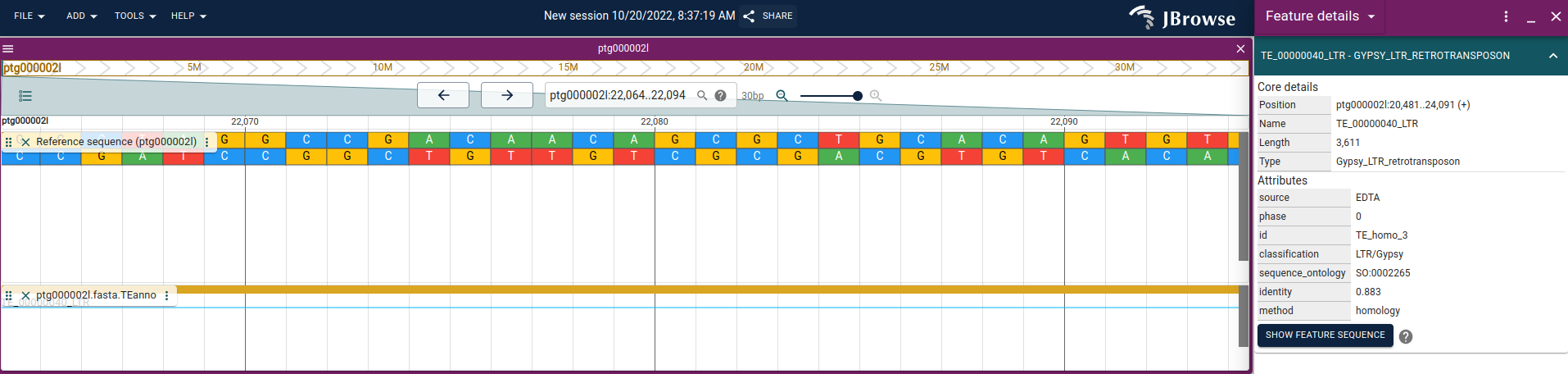
]

}

1. We take each target accession value from the blast file and search all of them using entrez direct via unix command, thanks to the esearch utility, we get a genbank file, save it as a temp file and converted the content to string to finally save it as a field in the collection as shown in the code bellow:

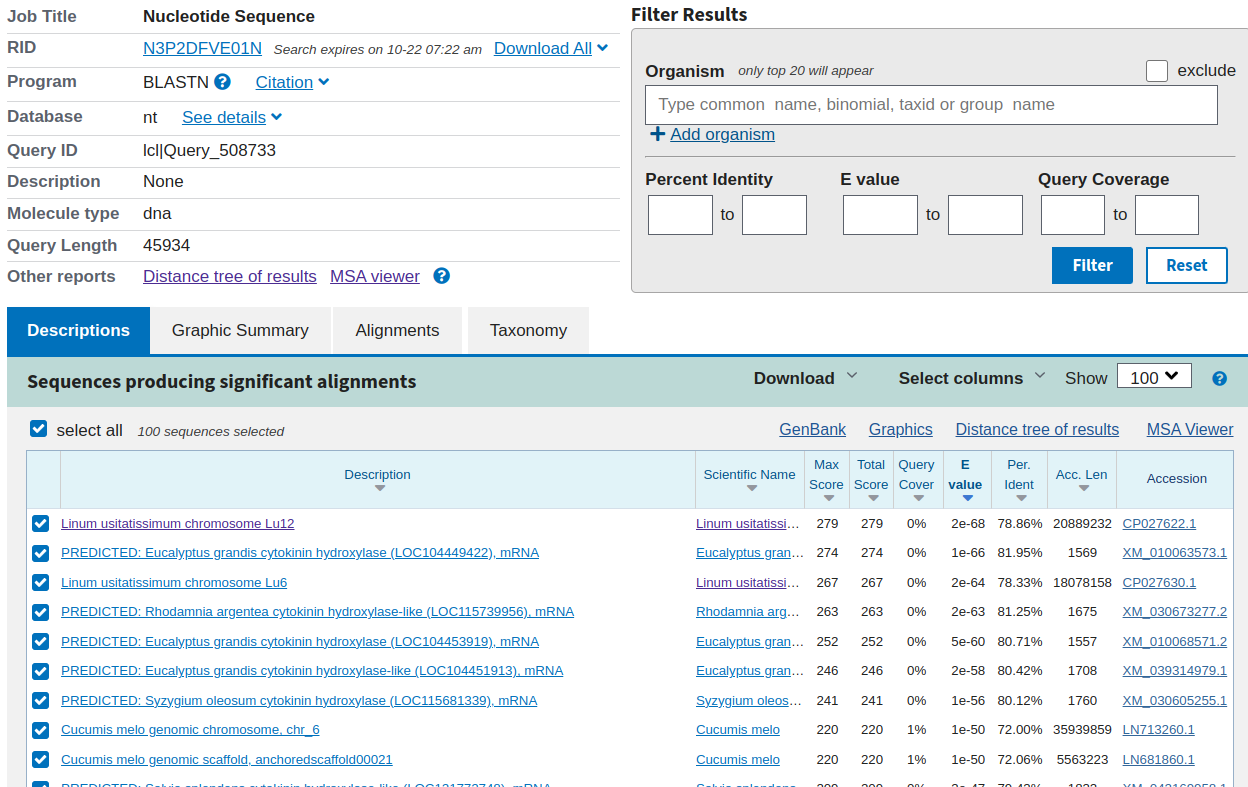


1. Showing a region of TEs in JBrowse



1. Chromosome extra information:

To get this information we made a blastn of a part of the chromosome on the NCBI web page, in which we get the following results:



Finally we searched for the organism in the NCBI web site and obtained the following extra information.

1. Order: Malpighiales
2. Family: Linaceae
3. Gender: Linum
4. Species: Linum usitatissimum