Find a Gene

Jessica Diaz-Vigil

2023-06-06

```
library(bio3d)
```

```
aln.seq<-read.fasta("seqdump1.fasta")
pre.heat.data <- seqidentity(aln.seq)
```

```
pre.heat.data <- seqidentity(aln.seq)
pre.heat.data
```

```
##                         Equus_asinus Gallus_gallus Ornithorhynchus_anatinus
## Equus_asinus                   1.000         0.079                    0.079
## Gallus_gallus                  0.079         1.000                    0.993
## Ornithorhynchus_anatinus       0.079         0.993                    1.000
## Arvicola_amphibius             0.079         0.993                    0.987
## Macaca_fascicularis            0.073         0.987                    0.980
## Callorhinchus_milii            0.079         0.987                    0.980
## protein                        0.995         0.079                    0.079
## Cottoperca_gobio               0.086         0.967                    0.967
##                         Arvicola_amphibius Macaca_fascicularis
## Equus_asinus                         0.079               0.073
## Gallus_gallus                        0.993               0.987
## Ornithorhynchus_anatinus             0.987               0.980
## Arvicola_amphibius                   1.000               0.980
## Macaca_fascicularis                  0.980               1.000
## Callorhinchus_milii                  0.980               0.974
## protein                              0.079               0.073
## Cottoperca_gobio                     0.960               0.954
##                         Callorhinchus_milii protein Cottoperca_gobio
## Equus_asinus                          0.079   0.995            0.086
## Gallus_gallus                         0.987   0.079            0.967
## Ornithorhynchus_anatinus              0.980   0.079            0.967
## Arvicola_amphibius                    0.980   0.079            0.960
## Macaca_fascicularis                   0.974   0.073            0.954
## Callorhinchus_milii                   1.000   0.079            0.954
## protein                               0.079   1.000            0.086
## Cottoperca_gobio                      0.954   0.086            1.000
```

```
par(oma=c(6,2,1,4))
par(mar=c(5,4,4,2) + 0.1)
heatmap(pre.heat.data)
```
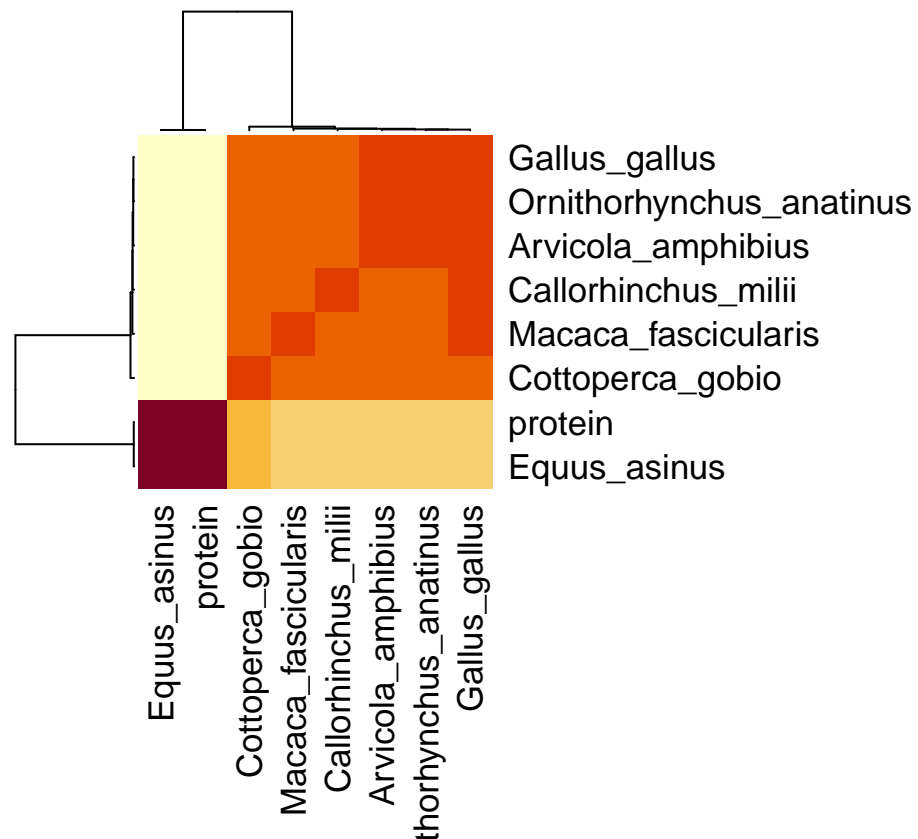
```
cons.seq<-consensus(aln.seq)

b.data<-blast.pdb(cons.seq$seq)

## Searching ... please wait (updates every 5 seconds) RID = 7XFXE3EC016
## .........
## Reporting 45 hits

plotb <- plot.blast(b.data)

##   * Possible cutoff values:    215 172 136 89 69 -2
##           Yielding Nhits:    9 20 32 38 43 45
##
##   * Chosen cutoff value of:    69
##           Yielding Nhits:    43
```
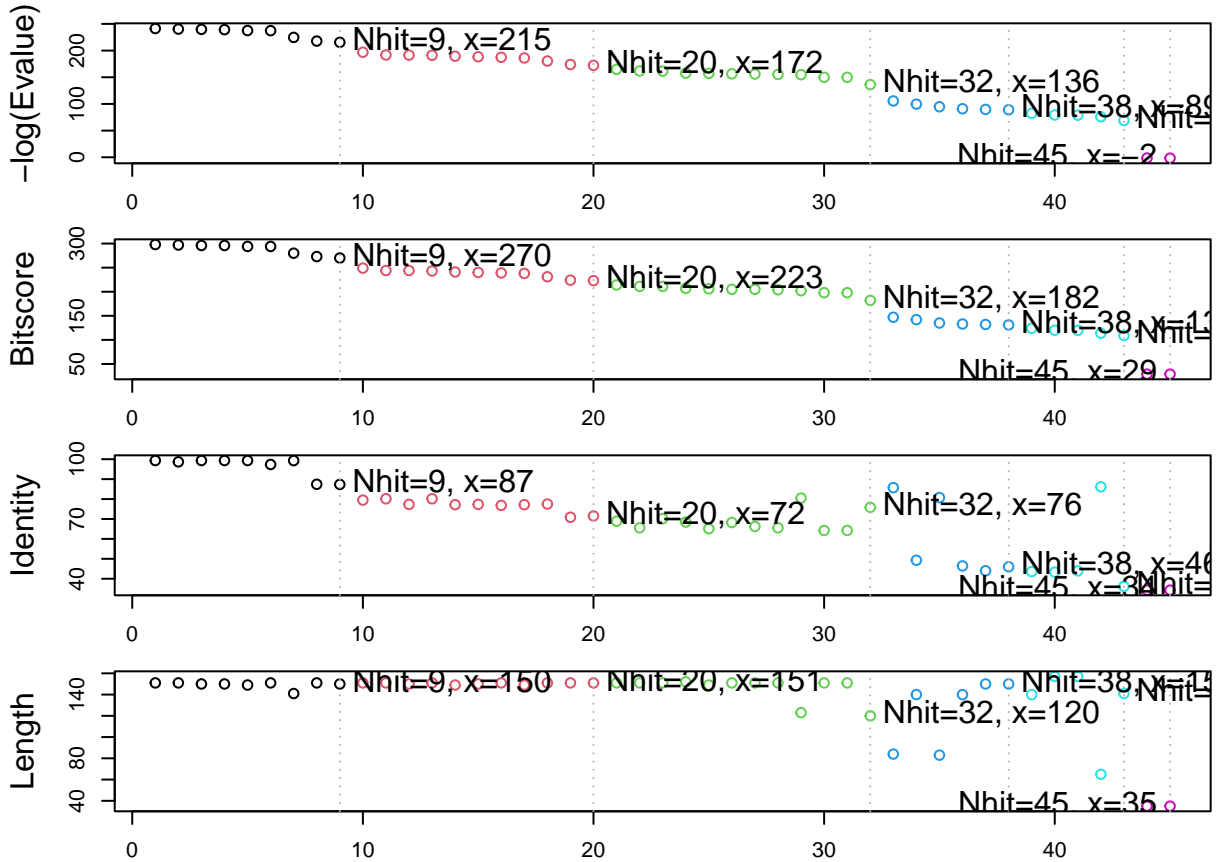
−log(Evalue) panel:
Nhit=9, x=215 Nhit=20, x=172 Nhit=32, x=136 Nhit=38, x=89 Nhit=
Nhit=45, x=−2

Bitscore panel:
Nhit=9, x=270 Nhit=20, x=223 Nhit=32, x=182 Nhit=38, x=11 Nhit=
Nhit=45, x=29

Identity panel:
Nhit=9, x=87 Nhit=20, x=72 Nhit=32, x=76 Nhit=38, x=46
Nhit=45, x=34 Nhit=

Length panel:
Nhit=9, x=150 Nhit=20, x=151 Nhit=38, x=1 Nhit=
Nhit=32, x=120
Nhit=45, x=35

```
head(pdb.annotate(plotb))
```

```
## Warning in .format_tbl(ret, ids, anno.terms, unique = unique): Annotation data could not be found for
##    3J7P_s, 7OYC_n, 3JAJ_s, 7OYA_n, 4V6W_a, 6XU6_a, 7R81_o, 3J6X_1, 4U3M_c, 6ZJ3_s, 4V7E_b, 8BTD_s, 4V

##          structureId chainId macromoleculeType chainLength experimentalTechnique
## 7NWH_n          7NWH       n           Protein          25                    EM
## 3JAG_n          3JAG       n           Protein          23                    EM
## 7TQL_N          7TQL       N           Protein         141                    EM
## 7PZY_O          7PZY       O           Protein         151                    EM
## 5OQL_s          5OQL       s           Protein         151                    EM
## 3J80_N          3J80       N           Protein         151                    EM
##          resolution scopDomain                                       pfam ligandId
## 7NWH_n         4.10       <NA> Ribosomal protein L41 (Ribosomal_L41)        <NA>
## 3JAG_n         3.65       <NA> Ribosomal protein L41 (Ribosomal_L41)        <NA>
## 7TQL_N         3.20       <NA>                                       <NA>    <NA>
## 7PZY_O         2.32       <NA>                                       <NA>    <NA>
## 5OQL_s         3.20       <NA> Ribosomal protein S15 (Ribosomal_S15)        <NA>
## 3J80_N         3.75       <NA> Ribosomal protein S15 (Ribosomal_S15)        <NA>
##          ligandName                                  source
## 7NWH_n         <NA>                     Oryctolagus cuniculus
## 3JAG_n         <NA>                     Oryctolagus cuniculus
## 7TQL_N         <NA>                              Homo sapiens
## 7PZY_O         <NA>                    Candida albicans SC5314
## 5OQL_s         <NA> Thermochaetoides thermophila DSM 1495
## 3J80_N         <NA>                      Kluyveromyces lactis
##
```

```
## 7NWH_n                                                   Mammalian pre-termination 80S ribosome with eRF1 and e
## 3JAG_n                          Structure of a mammalian ribosomal termination complex with ABCE1, eRF1(A
## 7TQL_N CryoEM structure of the human 40S small ribosomal subunit in complex with translation initiati
## 7PZY_O                                                             Structure of the vacant Ca
## 5OQL_s                                             Cryo-EM structure of the 90S pre-ribosome i
## 3J80_N                                                    CryoEM structure of 40S-eIF1-
##                                              citation rObserved rFree rWork
## 7NWH_n Powers, K.T., et al. Nucleic Acids Res (2021)        NA    NA    NA
## 3JAG_n             Brown, A., et al. Nature (2015)          NA    NA    NA
## 7TQL_N         Lapointe, C.P., et al. Nature (2022)         NA    NA    NA
## 7PZY_O          Zgadzay, Y., et al. Sci Adv (2022)          NA    NA    NA
## 5OQL_s  Cheng, J., et al. Nat Struct Mol Biol (2017)        NA    NA    NA
## 3J80_N             Hussain, T., et al. Cell (2014)          NA    NA    NA
##         spaceGroup
## 7NWH_n        <NA>
## 3JAG_n        <NA>
## 7TQL_N        <NA>
## 7PZY_O        <NA>
## 5OQL_s        <NA>
## 3J80_N        <NA>
```