Package ‘IMMAN’

April 15, 2020

Title Interlog protein network reconstruction by Mapping and Mining ANalysis

Version 1.6.0

Description Reconstructing Interlog Protein Network (IPN) integrated from several Protein protein Interaction Networks (PPINs). Using this package, overlaying different PPINs to mine conserved common networks between diverse species will be applicable.

Author Minoo Ashtiani,
Payman Nickchi,
Abdollah Safari,
Mehdi Mirzaie,
Mohieddin Jafari

Maintainer Minoo Ashtiani <`ashtiani.minoo@gmail.com`>

biocViews SequenceMatching, Alignment, SystemsBiology, GraphAndNetwork, Network, Proteomics

RoxygenNote 6.1.1

License Artistic-2.0

Encoding UTF-8

LazyData true

Imports STRINGdb, Biostrings, igraph, graphics, utils, seqinr

Suggests knitr, rmarkdown, testthat

VignetteBuilder knitr

git_url https://git.bioconductor.org/packages/IMMAN

git_branch RELEASE_3_10

git_last_commit f6243fb

git_last_commit_date 2019-10-29

Date/Publication 2020-04-14

R topics documented:

Celegance 2
FruitFly 2
H.sapiens 3
IMMAN 3
R.norvegicus 5

Index **7**

`Celegance` *Celegance*

Description

A sample collection of 49 UNIPROT_AC of Celegance species.

Usage

```
data("Celegance")
```

Format

a character vector

References

http://www.uniprot.org/

Examples

```
data("Celegance")
print(Celegance)
```

`FruitFly` *Fruit Fly*

Description

A sample collection of 56 UNIPROT_AC of Fruit Fly species.

Usage

```
data("FruitFly")
```

Format

a character vector

References

http://www.uniprot.org/

Examples

```
data("FruitFly")
print(FruitFly)
```

`H.sapiens` *Homo sapiens (Human)*

Description

A sample collection of 76 UNIPROT_AC of Homo sapiens species.

Usage

```
data("H.sapiens")
```

Format

a character vector

References

http://www.uniprot.org/

Examples

```
data("H.sapiens")
print(H.sapiens)
```

`IMMAN` *Interlog protein network reconstruction by Mapping and Mining ANalysis*

Description

A function for reconstructing Interlog Protein Network (IPN) integrated from Protein-protein Interaction Networks (PPIN) from different species. Users can overlay different PPINs to mine conserved common network between diverse species. It helps to retrieve IPN with different degrees of conservation to have better protein function prediction and PPIN analysis.

Usage

```
IMMAN(ProteinLists, fileNames = NULL, Species_IDs, identityU,
  substitutionMatrix, gapOpening, gapExtension, BestHit, coverage,
  NetworkShrinkage, score_threshold, STRINGversion,
  InputDirectory = getwd())
```

Arguments

`ProteinLists`	a list in which each element contains protein names of a species as a character vector. If it was NULL then the protein lists file name should be addressed in fileNames parameter.

`fileNames`	a character vector, containing names of text files containing protein list for each species. The protein list of each species must be in a column without header and rownames in seperate ".txt" files. The ProteinLists argument should be include at least two text file names addressing the protein list of each species which are in UniProt accession IDs format.
`Species_IDs`	a numeric vector; taxonomy ID for each organism which are provided in fileNames
`identityU`	numeric; value for selecting proteins whose alignment score is greater or equal than identityU
`substitutionMatrix`	a scoring substitution matrix to be used for alignment setting.
`gapOpening`	numeric; indicating the cost for opening a gap in the alignment
`gapExtension`	The incremental cost incurred along the length of the gap in the alignment
`BestHit`	logical; if TRUE describes a pair protein sequence among two different species which is the reciprocal best hit in sequence similarity analysis, whilst, if it is FALSE, indicates a nonreciprocal best hit
`coverage`	Number of connected proteins pairs in each Ortholog Protein Set (OPS) pair (termed as "coverage") to reconstruct an edge of OPS pair in the IPN (Interlog Protein Network)
`NetworkShrinkage`	logical; if TRUE OPSs that are similar to each other would be merged.
`score_threshold`	numeric; STRINGdb score for protein protein interaction (PPI) selection in STRING database
`STRINGversion`	character; indicating which version of STRING database should program search in for the score of PPIs.
`InputDirectory`	By default is getwd(). You can set this parameter to indicate where the downloaded file from STRING should be saved.

Value

a list contaning four elements:

IPNEdges : data.frame; Edges of resulted interlog protein network.

IPNNodes : data.frame; Nodes of resulted interlog protein network. Each node represents an OPS which is a set of ortholog proteins.

Network : list; Retrived PPINs of each input species.

maps : list; It includes data.frames indicating STRING_id data base matched to their corresponding UNIPROT_AC. The number of data.frames is according to the the number of species.

IPN : an igraph object representing the interlog protein network.

Author(s)

Minoo Ashtiani, Payman Nickchi, Abdollah Safari, Mehdi Mirzaie, Mohieddin Jafari

See Also

`pairwiseAlignment`

Examples

```
data(H.sapiens)
data(R.norvegicus)

ProteinLists = list(as.character(Celegance$V1), as.character(FruitFly$V1))

List1_Species_ID = 6239  # taxonomy ID List1 Celegance
List2_Species_ID = 7227  # taxonomy ID List2 FruitFly

Species_IDs  = c(List1_Species_ID, List2_Species_ID)

identityU = 30
substitutionMatrix = "BLOSUM62"
gapOpening = -8
gapExtension = -8
NetworkShrinkage = FALSE
coverage = 1
BestHit = TRUE
score_threshold = 400
STRINGversion="10"

# Run the IMMAN function for the parameters
output = IMMAN(ProteinLists, fileNames=NULL, Species_IDs,
              identityU, substitutionMatrix,
              gapOpening, gapExtension, BestHit,
              coverage, NetworkShrinkage,
              score_threshold, STRINGversion,
              InputDirectory = getwd())

output$IPNEdges
output$IPNNodes
output$Networks
output$Networks[[1]]
output$maps
output$maps[[2]]
```

`R.norvegicus` *Rattus norvegicus (Rat)*

Description

A sample collection of 70 UNIPROT_AC of Rattus norvegicus species.

Usage

```
data("R.norvegicus")
```

Format

a character vector

References

http://www.uniprot.org/

Examples

```
data("R.norvegicus")
print(R.norvegicus)
```

Index

∗Topic **datasets**
    `Celegance`, 2
    `FruitFly`, 2
    `H.sapiens`, 3
    `R.norvegicus`, 5

`Celegance`, 2

`FruitFly`, 2

`H.sapiens`, 3

`IMMAN`, 3

`pairwiseAlignment`, *4*

`R.norvegicus`, 5