Package 'Path2PPI'

October 18, 2017

Туре	Package
Title	Prediction of pathway-related protein-protein interaction networks
Versi	on 1.6.0
Date	2016-03-02
Main	tainer Oliver Philipp <contact@oliverphilipp.info></contact@oliverphilipp.info>
	ription Package to predict protein-protein interaction (PPI) networks in target organisms for which only a view information about PPIs is available. Path2PPI predicts PPI networks based on sets of proteins which can belong to a certain pathway from well-established model organisms. It helps to combine and transfer information of a certain pathway or biological process from several reference organisms to one target organism. Path2PPI only depends on the sequence similarity of the involved proteins.
Licen	se GPL (>= 2)
URL	http://www.bioinformatik.uni-frankfurt.de/
Depe	nds R ($>= 3.2.1$), igraph ($>= 1.0.1$), methods
Sugge	ests knitr, rmarkdown, RUnit, BiocGenerics, BiocStyle
Vigne	etteBuilder knitr
	or Oliver Philipp [aut, cre], Ina Koch [ctb]
	Tiews NetworkInference, SystemsBiology, Network, Proteomics, Pathways
Needs	sCompilation no
R to	opics documented: addReference 2 ai 3 getHybridNetwork 4 getPI 4 homologyScore 6
	Path2PPI-class plot.Path2PPI predictPPI

2 addReference

addReference	Add reference species	

Description

Adds reference species to an object from the class Path2PPI.

Usage

addReference(path2ppi, taxName, taxId, proteins, irefindex, homologs)

Arguments

path2ppi	An object of the class Path2PPI.
taxName	A character string giving the taxonomy name.
taxId	A character string giving the taxonomy identifier.
proteins	Either a character vector with the identifiers of the proteins which are involved in the corresponding pathway or a character vector with the protein names or aliases, respectively, named by the protein identifiers.
irefindex	Either a data frame, representing the iRefIndex table of the current reference species, e.g. loaded previously via read.table, or the corresponding file name of the iRefIndex file.
homologs	Either a data frame representing the results of the BLAST search (e.g. loaded previously via read. table) or the corresponding file name of the BLAST result file.

Details

This method searches for all relevant interactions in the data frame or file defined in iRefIndex. There are different and often ambiguous protein identifiers defined in an iRefIndex file, and the putative "major" identifiers are not necessarily those defined in the corresponding "major" columns "uidA" and "uidB". Furthermore, iRefIndex also contains protein complexes. Hence, Path2PPI applies an advanced search algorithm to automatically find relevant interactions associated with the pathway or the proteins of interest, respectively. The user does not have to predefine the identifiers types (Uniprot, Swissprot, Ensemble etc.), since these types are often unambiguously assigned. The algorithm searches for each identifier in 10 columns where any type of identifier or accession number is defined ("uidA", "altA", "OriginalReferenceA", "FinalReferenceA", "aliasA", "uidB", "altB", "OriginalReferenceB", "FinalReferenceB" and "aliasB"). Additionally, it searches for each complex which contains one or more of the predefined proteins. Subsequently, each homologous relationship which is not relevant for the previously found interactions is declined. The results of these searches are centralized in the Path2PPI object and can be visualized using the appropriate methods (e.g. showReferences)

Value

An object from the class Path2PPI with attached reference species.

ai 3

Author(s)

Oliver Philipp <MolBI-software@bioinformatik.uni-frankfurt.de>

See Also

showReferences, removeReference

Examples

ai

Data set to predict autophagy induction in Podospora anserina

Description

This data set consists of all data files necessary to predict the putative interactions of the induction step of autophagy in *Podospora anserina* by means of the corresponding PPIs in human and yeast.

Usage

```
data("ai")
```

Format

human.ai.irefindex: Data frames with 1694 observations of 54 variables. yeast.ai.irefindex: Data frames with 3840 observations of 54 variables. pa2human.ai.homologs: Data frames with 261 observations of 12 variables. pa2yeast.ai.homologs: Data frames with 98 observations of 12 variables. human.ai.proteins: Named character vector with 5 elements. yeast.ai.proteins: Named character vector with 7 elements.

Details

Data frames human.ai.irefindex and yeast.ai.irefindex consists of all relevant interactions of the corresponding iRefIndex files. The two data frames pa2human.ai.homologs and pa2yeast.ai.homologs are the necessary parts of the result files from the BLAST searches of the *P. anserina* proteom against the proteoms of human and yeast. The named character vectors human.ai.proteins and yeast.ai.proteins consists of the proteins involved in the induction process of autophagy in human and yeast.

Value

Four data frames and two named character vectors (see above).

4 getHybridNetwork

References

Camacho, C. et al. (2009). BLAST+: architecture and applications. BMC Bioinformatics, 10(1), 421.

Razick, S. et al. (2008). iRefIndex: a consolidated protein interaction database with provenance. BMC Bioinformatics, 9(1), 405.

Examples

data(ai)

getHybridNetwork

Get hybrid network of the predicted PPI

Description

Get the hybrid network of the previously predicted PPI. The hybrid network consists of all relevant interactions from the reference species, the predicted interactions in the target species and all relevant homologous relationships.

Usage

```
getHybridNetwork(path2ppi, igraph = FALSE)
```

Arguments

path2ppi An object of the class Path2PPI.

igraph Logical; if TRUE then the hybrid network is given as igraph-object. Otherwise a

data frame, consisting of each interaction and homologous relationship, will be

returned.

Value

See igraph argument.

Author(s)

Oliver Philipp <MolBI-software@bioinformatik.uni-frankfurt.de>

See Also

getPPI

Examples

getPPI 5

```
pa2yeast.ai.homologs)
ppi <- predictPPI(ppi)</pre>
```

```
#Return the hybrid network as data frame hybrid <- getHybridNetwork(ppi)
```

#Return the hybrid network as igraph object
hybrid <- getHybridNetwork(ppi,igraph=TRUE)</pre>

getPPI

Get predicted PPI

Description

Get the predicted PPI of an Path2PPI object consisting of each predicted interaction and protein in the target species.

Usage

```
getPPI(path2ppi, raw=FALSE, igraph=FALSE)
```

Arguments

path2ppi An object of the class Path2PPI.

raw Logical; if TRUE then the detailed view of the predicted PPI will be returned.

That means that each predicted interaction deduced from each reference species is given. In contrast, FALSE leads to the actually predicted and combined PPI

where no redundancies occur.

igraph Logical; if TRUE then the returned PPI is given as igraph-object. Otherwise a

data frame with each predicted interaction will be returned.

Value

See igraph argument.

Author(s)

Oliver Philipp <MolBI-software@bioinformatik.uni-frankfurt.de>

See Also

getHybridNetwork

Examples

6 homologyScore

yeast.ai.proteins, yeast.ai.irefindex,

```
pa2yeast.ai.homologs)

ppi <- predictPPI(ppi)

#Get the predicted PPI as data frame.
network <- getPPI(ppi)

#Get the detailed predicted PPI as data frame.</pre>
```

homologyScore

Computes homology scores

Description

Computes the homology scores based on the BLAST E-value. This function is used by the predictPPI method to compute homology scores to decide whether an interaction in a reference species is adopted to the target species (see package vignette for a detailed description). It can be used to test which E-values lead to which scores given a predefined E-value range.

Usage

homologyScore(e.value, h.range)

network.raw <- getPPI(ppi,raw=TRUE)</pre>

Arguments

e.value One BLAST E-value or a numeric vector with different BLAST E-values

h.range Numeric vector consisting of two values. The first value indicates the lower

bound (smallest E-value). Each E-value which is equal or less than this bound is scored with 1. The second value indicates the upper bound (biggest E-value).

Each E-value which is equal or greater than this bound is scored with 0.

Details

Uses a linear function to map the E-value v to the range [l,u] where l is the lower and u the upper bound:

$$s(v) = |m \log_{10}(v) + b|$$

$$m = \frac{1}{\log_{10}(l) - \log_{10}(u)}$$

$b = -(m\log_{10}(u))$

Value

Numeric vector containing the scores.

Author(s)

Oliver Philipp <MolBI-software@bioinformatik.uni-frankfurt.de>

Path2PPI-class 7

See Also

```
predictPPI
```

Examples

```
l <- 1e-100 #lower bound u <- 1e-20 #upper bound h.range <- c(1,u) #define range e.values <- c(1e-20,1e-40,1e-60,1e-80,1e-100) #some BLAST E-values homologyScore(e.values,h.range)
```

Path2PPI-class

Class "Path2PPI"

Description

An instance of the class Path2PPI is the major object in the Path2PPI package. It manages all reference species and the target species. The prediction algorithm is implemented in this class as well.

Usage

```
Path2PPI(...)
```

Arguments

... Argument list (see Note below).

Value

An instance of the class Path2PPI.

Slots

```
pathway: Object of class "character"
targetSpecies: Object of class ".TargetSpecies"
referenceContainer: Object of class ".ReferenceContainer"
h.thresh: Object of class "numeric"
h.range: Object of class "numeric"
i.thresh: Object of class "numeric"
consider.complexes: Object of class "logical"
max.complex.size: Object of class "numeric"
raw.ppi: Object of class "data.frame"
ppi: Object of class "data.frame"
```

8 plot.Path2PPI

Methods

```
addReference signature(path2ppi = "Path2PPI")
getHybridNetwork signature(path2ppi = "Path2PPI")
getPPI signature(path2ppi = "Path2PPI")
initialize signature(.Object = "Path2PPI")
plot.Path2PPI signature(x = "Path2PPI")
predictPPI signature(path2ppi = "Path2PPI")
removeReference signature(path2ppi = "Path2PPI")
show signature(object = "Path2PPI")
showInteraction signature(path2ppi = "Path2PPI")
showReferences signature(path2ppi = "Path2PPI")
```

Note

Arguments to Path2PPI() and the new method are obligatory and must be named if they differ from this order:

pathway A character string with the name of the pathway which has to be predicted.
 targetName A character string giving the taxonomy name of the target species.
 A character string giving the taxonomy identifier of the target species.

Author(s)

Oliver Philipp <MolBI-software@bioinformatik.uni-frankfurt.de>

Examples

```
ppi <- Path2PPI("Autophagy induction", "Podospora anserina", "5145")
ppi</pre>
```

plot.Path2PPI

Plots the predicted PPI

Description

Plots the predicted PPI in three different ways. Depending on the type argument it manages the specific layout settings and finally uses the plot function of the igraph package.

plot.Path2PPI 9

Arguments

x An object from the class Path2PPI where the PPI network already has been

predicted.

type Character string. Which graph type to plot. "ppi": plots only the predicted PPI.

"hybrid": plots the hybrid network which consists of all relevant interactions from the reference species, the predicted interactions in the target species and

all relevant homologous relationships.

multiple.edges Logical. Is only considered if type="ppi". If TRUE then each reference inter-

action is depicted in the species-specific color (raw mode), in contrast, if set to

FALSE only the finalized / combined interactions are depicted.

scores Logical. If TRUE the edge scores will be shown.

species.colors Named vector, to specify the species colors. If no value is given then default

colors are used.

vertices.opacity

Numeric value between 0 and 1 defining the opacity of the vertices.

use.identifiers

Logical. If TRUE then only the proteins identifiers are used as the vertex labels.

protein.labels Named vector to define the labels of the vertices. If no value is given then the

protein identifiers are used. The vector does not have to be complete, i.e. not

each protein has to be defined.

show. legend Logical. If TRUE then a legend is depicted.

vertices.coordinates

Data frame containing the coordinates of the vertices. If no value is given then

coordinates are computed using the layout.auto function.

return.coordinates

Logical. If TRUE the coordinates of the vertices are returned.

tkplot Logical. If TRUE the graph is drawn in the interactive graph drawing facility

tkplot.

... Additional plotting parameters.

Details

The argument return. coordinates only works correctly if tkplot=FALSE. If you want to get the coordinates of the tkplot device use tkplot.getcoords.

Value

If return.coordinates=TRUE the coordinates of the vertices are returned.

Note

If you want to export the plotted graph to postscript you have to consider that the default font family is set to *sans* for vertex and edge labels. Please change the default font family of postscript to *sans* before you call the plot method: ps.options(family="sans"). Additionally, you have to consider that the default value for vertices.opacity is set to 0.8 in order to enhance the visibility of the graph, since some edges may be hidden by the vertices. Postscript does not support semitransparencies. Hence, please change the vertices.opacity argument to 1 if you want to export the graph using postscript.

10 predictPPI

Author(s)

Oliver Philipp <MolBI-software@bioinformatik.uni-frankfurt.de>

See Also

```
predictPPI, igraph for other plotting parameters
```

Examples

```
data(ai) #Load test data set
ppi <- Path2PPI("Autophagy induction", "Podospora anserina", "5145")</pre>
ppi <- addReference(ppi, "Homo sapiens", "9606", human.ai.proteins,</pre>
                    human.ai.irefindex, pa2human.ai.homologs)
ppi <- addReference(ppi, "Saccharomyces cerevisiae (S288c)", "559292",
                    yeast.ai.proteins, yeast.ai.irefindex,
                    pa2yeast.ai.homologs)
ppi <- predictPPI(ppi,h.range=c(1e-60,1e-20))</pre>
#Plot the predicted PPI with the default settings and return
#the coordinates of the vertices
set.seed(12)
coordinates <- plot(ppi, return.coordinates=TRUE)</pre>
#Plot the predicted PPI and show each underlying reference interaction.
#Use different species specific colors. To compare both graphs,
#use the coordinates computed before
plot(ppi,multiple.edges=TRUE,vertices.coordinates=coordinates)
#Plot the corresponding hybrid network with predefined species colors.
#Also define some labels for the proteins of the target species.
#Keep in mind: You can not use the data in "coordinates" since
#the hybrid network consists of more vertices than the default PPI
set.seed(40)
target.labels<-c("B2AE79"="PaTOR","B2AXK6"="PaATG1",</pre>
                "B2AUW3"="PaATG17", "B2AM44"="PaATG11",
                "B2AQV0"="PaATG13", "B2B5M3"="PaVAC8")
species.colors <- c("5145"="red","9606"="blue","559292"="green")</pre>
plot(ppi,type="hybrid",species.colors=species.colors,
protein.labels=target.labels)
```

predictPPI

Prediction of the PPI

Description

Major method of the Path2PPI class to predict the final PPI in the target species using the information available from the stored reference species. Different values for the arguments of this method can lead to different PPI networks, differing in the degree of reliability and strictness.

predictPPI 11

Usage

Arguments

path2ppi An object of the class Path2PPI.

mode Which interaction from the reference species should be taken into account. "both":

both interactors of an interaction has to be in the initial protein list previously inserted by the user (recommended if it is a large network or many proteins were initially defined, respectively). "one": only one of the interactors of each reference interaction has to be in the initial protein list (may lead to very large

networks).

h. thresh E-value cutoff at which each homologous relationship definitely will be declined

(see also h.range argument).

h.range Numeric vector consisting of two values. The first value indicates the lower

border (smallest E-value). Each E-value which is equal or less than this border is scored with 1 (best). The second value indicates the upper border (biggest E-value). Each E-value which is equal or greater than this border is scored with

0 (worst).

i.thresh Numeric. Threshold for accepted interactions. If the computed prediction score

for an interaction is less than i. thresh it will be declined.

consider.complexes

Logical. If TRUE then interactions are also considered which actually indicate an association of the current protein to one bigger protein complex. This may lead to very large networks if mode="one" since all other proteins of this complex are considered as well, i.e., each protein in such complexes are considered to interact with each other protein of this complex. If mode="both" then each protein of an complex has to be in the initial protein list to consider each interaction (see

details).

max.complex.size

Numeric. Is only considered if consider.complexes=TRUE. The maximum size of complexes to be considered.

decline.self.interaction.ref

Logical. If TRUE then all self interactions from reference species are declined.

decline.self.interaction.tar

Logical. If TRUE then all predicted self interactions in target species are declined.

verbose Logical. FALSE hides messages in the output.

Details

Difference of h.thresh and h.range: If only one protein in the target species was found to be homologous to a current reference species protein and this homology was rated with an E-value which is equal or smaller than h.thresh it is scored with 1 (even if the E-value is larger than the upper border of h.range). See package vignette for more details.

12 removeReference

Use the complex arguments with care, since each complex may lead to a vast amount of interactions, i.e., each protein is considered to interact with each other of this complex; e.g. if there are 10 proteins involved in one complex, this would lead to 10 over 2 = 45 interactions.

Value

An object of the class Path2PPI with predicted PPI.

Author(s)

Oliver Philipp <MolBI-software@bioinformatik.uni-frankfurt.de>

See Also

```
plot.Path2PPI, homologyScore
```

Examples

```
data(ai) #Load test data set
ppi <- Path2PPI("Autophagy induction", "Podospora anserina", "5145")</pre>
ppi <- addReference(ppi, "Homo sapiens", "9606", human.ai.proteins,</pre>
                    human.ai.irefindex, pa2human.ai.homologs)
ppi <- addReference(ppi, "Saccharomyces cerevisiae (S288c)", "559292",</pre>
                    yeast.ai.proteins, yeast.ai.irefindex,
                    pa2yeast.ai.homologs)
#Using the default settings leads to 8 predicted interactions in the
#target species
ppi <- predictPPI(ppi)</pre>
#Consider complexes where each complex is allowed to be up to 10 proteins
#large. For this smaller pathway only one more interaction was predicted when
#considering larger complexes.
ppi <- predictPPI(ppi,consider.complexes=TRUE,max.complex.size=10)</pre>
#We can be less strict and decrease h.range what obviously increases the
#number of predicted interactions to 13
ppi <- predictPPI(ppi,h.range=c(1e-60,1e-20))</pre>
```

removeReference

Remove reference species

Description

Remove reference species previously attached to an object from the class Path2PPI.

```
removeReference(path2ppi, species)
```

showInteraction 13

Arguments

path2ppi An object from the class Path2PPI.

species Either a number between 1 and the number of stored reference species or a

character string with the taxonomy id of the reference species to remove.

Value

An object of the class Path2PPI with removed reference species species.

Author(s)

Oliver Philipp <MolBI-software@bioinformatik.uni-frankfurt.de>

See Also

```
showReferences, addReference
```

Examples

```
data(ai) #Load test data set

ppi <- Path2PPI("Autophagy induction", "Podospora anserina", "5145")

ppi <- addReference(ppi, "Homo sapiens", "9606", human.ai.proteins, human.ai.irefindex, pa2human.ai.homologs)

ppi <- addReference(ppi, "Saccharomyces cerevisiae (S288c)", "559292", yeast.ai.proteins, yeast.ai.irefindex, pa2yeast.ai.homologs)

#Remove second reference species

ppi <- removeReference(ppi,2)

#Remove reference species with taxonomy id "9606"

ppi <- removeReference(ppi,"9606")</pre>
```

showInteraction

Information about an interaction

Description

Use showInteraction to get detailed information about one interaction of the predicted PPI.

```
showInteraction(path2ppi, interaction, mode="default", verbose=TRUE)
```

14 showReferences

Arguments

path2ppi An object from the class Path2PPI.

interaction Character vector consisting of the identifiers of the two interactors.

mode Character string. Which information of this interaction is requested. "default":

only the predicted interaction and some major information are provided. "detailed": all interactions deduced from each reference species with this interaction is provided. "references": each reference interaction of the current interaction with some major information. "references.detailed": each reference interaction of the current interaction with all available information (extracted from the cor-

responding iRefIndex data set).

verbose Logical. FALSE hides messages in the output.

Value

Data frame with the requested information defined in mode.

Author(s)

Oliver Philipp <MolBI-software@bioinformatik.uni-frankfurt.de>

See Also

```
plot.Path2PPI,showReferences
```

Examples

showReferences

Information about reference species

Description

Get information about the currently stored reference species. If indicated by returnValue a data frame - containing information about each protein or interaction - is provided as well.

```
showReferences(path2ppi, species = NA, returnValue = NA)
```

showReferences 15

Arguments

path2ppi An object from the class Path2PPI.

species Either a number between 1 and the number of stored reference species or a

character string with the taxonomy id. If no value for species is given then

information about each stored reference species is provided.

returnValue Character value indicating whether to return a value. "proteins": a data frame

containing the proteins associated with the pathway of interest in the corresponding reference species. "interactions": a data frame containing all processed, relevant and non-redundant interactions. "irefindex": a data frame containing all relevant interactions in the raw irefindex format. Is only reasonable if species is defined. If no value for returnValue is given then only general

information is provided.

Value

See description for returnValue

Author(s)

Oliver Philipp <MolBI-software@bioinformatik.uni-frankfurt.de>

See Also

addReference, removeReference, showInteraction

Examples

```
data(ai) #Load test data set
ppi <- Path2PPI("Autophagy induction", "Podospora anserina", "5145")</pre>
ppi <- addReference(ppi, "Homo sapiens", "9606", human.ai.proteins,</pre>
                    human.ai.irefindex, pa2human.ai.homologs)
ppi <- addReference(ppi, "Saccharomyces cerevisiae (S288c)", "559292",
                    yeast.ai.proteins, yeast.ai.irefindex,
                    pa2yeast.ai.homologs)
#Get general information about each stored reference species
showReferences(ppi)
#Get general information about reference species with the taxonomy id "9606"
showReferences(ppi, species="9606")
#Get all proteins associated with the pathway of interest
#and previously given by the user
proteins <- showReferences(ppi, species="9606", returnValue="proteins")</pre>
#Get all processed and non-redundant interactions previously
#determined to be relevant for the pathway of interest
interactions <- showReferences(ppi, species="9606", returnValue="interactions")
#Get all relevant interactions in the detailed irefindex format
irefindex <- showReferences(ppi, species="9606", returnValue="irefindex")</pre>
```

Index

```
*Topic datasets
                                                 showInteraction, 8, 13, 15
    ai, 3
                                                 showInteraction, Path2PPI-method
                                                          (Path2PPI-class), 7
addReference, 2, 8, 13, 15
                                                 showInteraction-methods
addReference, Path2PPI-method
                                                          (showInteraction), 13
        (Path2PPI-class), 7
                                                 showReferences, 2, 3, 8, 13, 14, 14
ai, 3
                                                 showReferences, Path2PPI-method
                                                          (Path2PPI-class), 7
getHybridNetwork, 4, 5, 8
getHybridNetwork,Path2PPI-method
                                                 tkplot, 9
        (Path2PPI-class), 7
                                                 tkplot.getcoords, 9
getPPI, 4, 5, 8
                                                 yeast.ai.irefindex(ai), 3
getPPI,Path2PPI-method
        (Path2PPI-class), 7
                                                 yeast.ai.proteins(ai), 3
homologyScore, 6, 12
human.ai.irefindex(ai), 3
human.ai.proteins(ai), 3
igraph, 8, 10
initialize, 8
initialize, Path2PPI-method
        (Path2PPI-class), 7
layout.auto, 9
new, 8
pa2human.ai.homologs(ai), 3
pa2yeast.ai.homologs(ai), 3
Path2PPI (Path2PPI-class), 7
Path2PPI-class, 7
plot.Path2PPI, 8, 8, 12, 14
plot.Path2PPI,Path2PPI-method
        (Path2PPI-class), 7
predictPPI, 7, 8, 10, 10
predictPPI,Path2PPI-method
        (Path2PPI-class), 7
removeReference, 3, 8, 12, 15
removeReference,Path2PPI-method
        (Path2PPI-class), 7
show, 8
```

show, Path2PPI-method (Path2PPI-class), 7