Package 'PathQuant'

April 8, 2016

Type Package

Title Pathway Analysis with distance between genes and metablites using KEGG maps.

Version 0.1.1

Date 2015-07-22

Author Sarah Cherkaoui <cherkaos@gmail.com>
Sandra Therrien-Laperriere <sandra.therrien.laperriere@gmail.com>

Maintainer Sandra Therrien-Laperriere <sandra.therrien.laperriere@gmail.com>

Description Modeling of KEGG metabolism pathway maps to calculate shortest distance between gene-metabolite pairs to asses their biological relevance with tables, statistical and graphical outputs.

License GNU-3

Imports igraph, XML, RUnit, ggplot2, data.table, KEGGREST,knitr,rmarkdown,scales

VignetteBuilder knitr

Suggests testthat, knitr, rmarkdown

LazyData TRUE

RoxygenNote 5.0.1

R topics documented:

`distanceGeneToAllMetabolite` *Function calculating shortest distance between a selected gene of a gene-metabolite pairs and every metabolites.*

Description

Function calculating shortest distance between a selected gene of a gene-metabolite pairs and every metabolites (in a pair or not) on a graph model of KEGG map selected, where nodes are metabolites and reactions are edges.

Usage

```
distanceGeneToAllMetabolite(pathwayId, data, metabolite, gene)
```

Arguments

`pathwayId`	KEGG Id of selected map
`data`	is a dataFrame with 2 columns. Where each line reprensents an associations with the first column as gene KEGG Ids and the sencond column as metabolite KEGG Ids.
`metabolite`	is a dataframe of 1 column with the KEGG Ids of all measured metabolites.
`gene`	is the KEGG Id of the selected gene

Details

If a gene or a metabolite is present on multiple edges or nodes, then shortest distance are calculated for every combinaison possible and the shortest distance is selected.

Output: barplot showing the distribution of the calculated distance, where red bars represent a distance with an associated metabolite and grey bars distance with no associated metabolites.

Examples

```
distanceGeneToAllMetabolite(metabolismOverviewMapKEGGId,
shinAndAlDF, completeMetaboDF, "hsa:1373")
```

`getDistanceAll` *Function calculating shortest distance between every gene in a gene-metabolite pairs and all metabolite.*

Description

Function calculating shortest distance between every gene in a gene-metabolite pairs of your data parameter and every metabolite (in a pair or not) on a graph model of KEGG map selected, where nodes are metabolites and reactions are edges.

Usage

```
getDistanceAll(pathwayId, data, metabolite)
```

Arguments

pathwayId	KEGG Id of selected map
data	is a dataFrame with 2 columns. Where each line reprensents an associations with the first column as gene KEGG Ids and the sencond column as metabolite KEGG Ids.
metabolite	is a dataframe of 1 column with the KEGG Ids of all measured metabolites.

Details

If a gene or a metabolite is present on multiple edges or nodes, then shortest distance are calculated for every combinaison possible and the shortest distance is selected.

Output : dataframe with metabolite in columns and gene in rows with shortest distance values.

Examples

```
getDistanceAll(metabolismOverviewMapKEGGId,shinAndAlDF,
          completeMetaboDF)
```

---

getDistanceAsso *Function calculating shortest distance between each gene-metabolite pairs.*

Description

Function calculating shortest distance between each gene-metabolite pairs of your data parameter on a graph model of KEGG map selected, where nodes are metabolites and reactions are edges.

Usage

```
getDistanceAsso(pathwayId, data, ordered = FALSE)
```

Arguments

pathwayId	KEGG Id of selected map
data	is a dataFrame with 2 columns. Where each line reprensents an associations with the first column as gene KEGG Ids and the sencond column as metabolite KEGG Ids.
ordered	[option] ascendent ordering of distance

Details

If a gene or a metabolite is present on multiple edges or nodes, then shortest distance are calculated for every combinaison possible and the shortest distance is selected.

Output : dataframe with the following columns : geneCommonName, geneKEGGId, isGeneInMap, metaboliteCommonName, metaboliteKEGGId, isMetaboliteInMap, distance

Examples

```
getDistanceAsso(metabolismOverviewMapKEGGId,shinAndAlDF)
```

`heatmapAsso` *Function that output a heatmap to visualize distance calculated between every gene-metabolite associations in input.*

Description

Function calculting shortest distance between every genes and metabolites in a gene-metabolite pairs of your data parameter on a graph model of KEGG map selected, where nodes are metabolites and reactions are edges.

Usage

```
heatmapAsso(pathwayId, data)
```

Arguments

`pathwayId`	KEGG Id of selected map
`data`	is a dataFrame with 2 columns. Where each line reprensents an associations with the first column as gene KEGG Ids and the sencond column as metabolite KEGG Ids.

Details

If a gene or a metabolite is present on multiple edges or nodes, then shortest distance are calculated for every combinaison possible and the shortest distance is selected.

Output : Heatmap graphic where black countoured distances represent an association.

Examples

```
heatmapAsso(metabolismOverviewMapKEGGId, shinAndAlDF)
```

`permutationTest` *Statistical permutation test to asses the relevance of associated gene-metabolite pairs vs. randomly selected pairs*

Description

Permutation test to evauluate if gene-metabolite associations pairs of your data parameter are significantly closer than randomly selected gene-metabolite pairs by calculating shortest distance between assocatied gene-metabolite pairs vs. randomly selected pairs.

Usage

```
permutationTest(pathwayId, data, gene, metabolite, permutation,
  output = c("medians", "pvalue", "histogram"))
```

Arguments

pathwayId	KEGG Id of selected map
data	is a dataframe with 2 columns. Where each line reprensents an associations with the first column as gene KEGG Ids and the sencond column as metabolite KEGG Ids.
gene	is a dataframe of 1 column with the KEGG Ids of all measured genes.
metabolite	is a dataframe of 1 column with the KEGG Ids of all measured metabolites.
permutation	is the number desired permutations
output	medians (list of median of every permutation), pvalue of the permutation test, histogram (a histrogram representing the distribution of every permutation median)

Details

If a gene or a metabolite is present on multiple edges or nodes, then shortest distance are calculated for every combinaison possible and the shortest distance is selected.

Examples

```
permutationFunction(metabolismOverviewMapKEGGId, shinAndAlDF,
            completeGeneDF,completeMetaboDF, 1000, "histogram")
```

Index