domainsignatures

November 11, 2009

R topics documented:

`dataSource` *Contructor for annotation object*

Description

This function creates the necessary annotation object of class `ipDataSource` containing the pathway and InterPro ID mappings.

Usage

```
dataSource(mapping, type = "generic")
```

Arguments

`mapping`	A named list providing a mapping between enterzgene identifiers and arbitrary groupings of genes or pathways.
`type`	The type of pathway. A character skalar.

Details

For genes without pathway membership, `NA` list items need to be included in `mapping`. The names of the list comprise the gene universe to test against. The function will access the ensembl `biomaRt` database in order to retrieve the necessary InterPro domain information.

Value

Object of class `ipDataSource`

Author(s)

Florian Hahne

See Also

`gseDomain`, `getKEGGdata`

Examples

```
grouping <- list("653361"=c("pw1", "pw2"), "729230"="pw1",
                 "415117"="pw3")
dataSource(grouping)
```

`domainsignatures-package` *Geneset enrichment based on InterPro domain signatures.*

Description

Compute similarities to pathways for a set of entrezgene identifiers based on the InterPro domain signature

Details

Package:	domainsignatures
Type:	Package
Version:	1.0
Date:	2007-07-02
License:	LPGL?

see help for `gseDomain` for details

Author(s)

Florian Hahne, Tim Beissbarth

Maintainer: Florian Hahne <f.hahne@dkfz.de>

`getKEGGdata` *Fetch KEGG annotations and InterPro domains*

Description

Get all available KEGG annotations and InterPro domains for a set of entrezgene identifiers from the `KEGG` annotation package and from the ensembl `biomaRt`.

Usage

```
getKEGGdata(universe, pathways)
```

Arguments

`universe`	Character vector of entrezgene identifiers. This is the global universe of genes to test against.
`pathways`	Optional character vector of KEGG pathway identifiers. This can be used in order to test for over-representation of only a subset of all the available KEGG pathways.

Details

This function is a wrapper around the `KEGG` annotation package and a customized query of the ensembl `biomaRt` database. For the gene identifiers in `universe` and all KEGG pathways it will fetch the necessary information and create an object of class `ipDataSource` which can later be used as input to `gseDomain`

Value

An object of class `ipDataSource`.

Author(s)

Florian Hahne

See Also

`gseDomain`

Examples

```
getKEGGdata()
```

`getKEGGdescription` *KEGG description from ID*

Description

Get description of KEGG pathways from a list of KEGG identifiers.

Usage

```
getKEGGdescription(ids)
```

Arguments

`ids`	Character vector of KEGG identifiers

Value

Character vector of KEGG descriptions

Author(s)

Florian Hahne

Examples

```
getKEGGdescription("hsa03050")
```

`gseDomain` *Geneset enrichment based on InterPro domain signatures*

Description

Compute the similarity to pathways specified trough `dataSource` for a set of entrezgene identifiers.

Usage

```
gseDomain(dataSource, geneset, n=10000, verbose=TRUE, samples=FALSE)
```

Arguments

`dataSource`	Object of class `ipDataSource` containing pathway and InterPro domain mappings
`geneset`	Character vector of entrezgene identifiers
`n`	Number of subsampling iterations
`verbose`	Toggle progress report
`samples`	Logical indicating whether to return the similarity measures for all the resamples.

Details

Use this function to compute p-values for similarity of the domain signature of a gene set to all signatures of the pathways defined in `dataSource`. You should have created `dataSource` using either function `dataSource` or `getKEGGdata`.

Value

A list with items

`similarity`	Named vector of similarity measures for each pathway
`pvalue`	The p-values of similarity to each pathway. A named vector.
`dist`	A named list containing similarity measures for all the resamples

Author(s)

Florian Hahne

See Also

`gseDomain`

Examples

```
## see Vignette of this package for examples how to use this function
```

`ipDataSource-class` *A class to store mapping information between genes, pathways and interPro domains*

Description

This class represents the data necessary to run `gseDomain`

Details

You should always create these objects using either one of the functions `dataSource` or `getKEGGdata`.

Creating Objects

Objects can be created using

```
new('ipDataSource',
genes = ...., # Object of class character
pathways = ....  # Object of class character
domains = ....  # Object of class character
gene2Domains = ....  # Object of class environment
path2Domains = ....  # Object of class environment
type = ...., # Object of class character
)
```

or the functions `dataSource` or `getKEGGdata`.

Slots

`genes`: Vector of unique entrezgene identifiers

`pathways`: Vector of unique pathway identifiers

`domains`: Vector of unique InterPro identifiers

`gene2Domains`: Hash table mapping entrezgene IDs to Interpro IDs

`path2Domains`: Hash table mapping pathway IDs to InterPro IDs

`type`: Type of pathway. A character skalar

Methods

show display summary.

Author(s)

Florian Hahne

See Also

`dataSource` or `getKEGGdata`

Index