

Package ‘sfinx’

July 19, 2017

Type Package

Title Straightforward Filtering Index for AP-MS Data Analysis (SFINX)

Version 1.7.99

Date 2017-07-19

Description The straightforward filtering index (SFINX) identifies true positive protein interactions in a fast, user-friendly, and highly accurate way. It is not only useful for the filtering of affinity purification - mass spectrometry (AP-MS) data, but also for similar types of data resulting from other co-complex interactomics technologies, such as TAP-MS, Virotrap and BioID. SFINX can also be used via the website interface at <<http://sfinx.ugent.be>>.

License Apache License 2.0

LazyData TRUE

Depends R (>= 3.2.3)

URL <http://sfinx.ugent.be>

RoxygenNote 5.0.1

Suggests knitr, rmarkdown, testthat

VignetteBuilder knitr

NeedsCompilation no

Author Kevin Titeca [aut, cre],
Jan Tavernier [ths],
Sven Eyckerman [ths]

Maintainer Kevin Titeca <sfinxinteractomics@gmail.com>

Repository CRAN

Date/Publication 2017-07-19 20:55:19 UTC

R topics documented:

BaitIdentityExampleFile	2
DataInputExampleFile	2
sfinx	3

BaitIdentityExampleFile

A vector with proteins of interest (baits) for the TIP49 dataset.

Description

A character vector with all the bait proteins of interest. These proteins are all present as the exact same rownames in the DataInputExampleFile, and these are also the original bait proteins that were used in the original publication of Sardu et al. (see below).

Usage

BaitIdentityExampleFile

Format

A character vector with 26 bait protein entries of interest.

Source

M.E. Sardu, Y. Cai, J. Jin, S.K. Swanson, R.C. Conaway, J.W. Conaway, L. Florens, M.P. Washburn Probabilistic assembly of human protein interaction networks from label-free quantitative proteomics. Proc. Natl. Acad. Sci. USA, 105 (2008), pp. 1454 to 1459.

DataInputExampleFile *The TIP49 dataset of protein interactions (AP-MS).*

Description

A strictly numeric input matrix with unique proteins as rownames and unique projects as colnames. The cells of the matrix are filled with the associated **peptide counts**. Cells that have no associated peptide counts are filled with a zero. This specific example dataset is derived from the publication of Sardu et al. (see below). It contains complexes involved in chromatin remodeling and consists of 35 bait-specific projects and 35 negative controls.

Usage

DataInputExampleFile

Format

A matrix with 1581 rows (proteins) and 70 variables (projects).

Source

M.E. Sardi, Y. Cai, J. Jin, S.K. Swanson, R.C. Conaway, J.W. Conaway, L. Florens, M.P. Washburn Probabilistic assembly of human protein interaction networks from label-free quantitative proteomics. Proc. Natl. Acad. Sci. USA, 105 (2008), pp. 1454 to 1459.

sfinx

SFINX (Straightforward Filtering INdeX).

Description

sfinx identifies the true-positive protein interactions in affinity purification - mass spectrometry data sets and in similar co-complex interactomics data sets. It is highly accurate, fast and independent of external data input.

It is also available via the Web interface at <http://sfinx.ugent.be>, which has extra analysis and visualization features.

Usage

```
sfinx(InputData, BaitVector, BackgroundRatio = 5,
      BackgroundIdentity = "automatic", BaitInfluence = FALSE,
      ConstantLimit = TRUE, FWERType = "B")
```

Arguments

- | | |
|--------------------|--|
| InputData | A strictly numeric matrix with unique proteins as rownames and unique projects as colnames. The cells of the matrix are filled with the associated peptide counts . Cells that have no associated peptide counts have to be filled with a zero. |
| BaitVector | A character vector with all the bait proteins of interest. These proteins should all be present as the exact same rownames in InputData. sfinx will control this, and it will report possible deviations. |
| BackgroundRatio | Advanced. A natural number equal or bigger than 2, that specifies the maximal ratio of total considered projects over the amount of bait projects. If this parameter equals for example 5, it will take into account 4 times as much non-bait projects as it uses bait projects. sfinx will preferably first select the non-bait projects with most peptide counts as negative controls. |
| BackgroundIdentity | Deprecated. A character string or character vector describing the background projects. "automatic" is the advised default entry. However, all extra or alternative entries will be matched to the column headers and taken into account when possible. |
| BaitInfluence | Advanced. A logical. When TRUE, sfinx uses only the non-bait negative control projects with the biggest amount of data for the calculation of the background, but no negative control projects associated with other baits in the analysis. When FALSE, sfinx uses both. |

- ConstantLimit** Advanced. A logical. When TRUE, an internal cut-off is used that is a simplified constant for the actual complete calculation of the binomial equivalent. This is the version of *sfinx* that was used in the article (Titeca et al., J. Proteome Res., 2016). When FALSE, the complete calculation of the binomial equivalent is done. Some datasets with many highly abundant proteins can benefit from having this parameter FALSE.
- FWERType** Advanced. A character string that equals "B", "HolmB" or "Sidak". "B" gives the Bonferroni correction for the family wise error rate (FWER), "HolmB" gives Holm-Bonferroni correction, and "Sidak" gives Sidak correction. However, note that these options will only very rarely yield different output.

Details

For most standard applications of *sfinx*, the arguments `InputData` and `BaitVector` should be sufficient. Any optimization of the other parameters is discouraged and should be explicitly reported upon communication of the results.

Value

sfinx returns a list with two elements. The first element of the list contains a dataframe with the true-positive protein interactions that were identified by *sfinx* in `InputData` for the proteins of interest in `BaitVector`. The second element of the list contains a string with comments about the output and the underlying data.

Examples

```
sfinx(DataInputExampleFile, BaitIdentityExampleFile)
```

```
sfinx(InputData = DataInputExampleFile, BaitVector =  
BaitIdentityExampleFile, ConstantLimit = FALSE, FWERType = "Sidak")
```

Index

*Topic **datasets**

BaitIdentityExampleFile, [2](#)

DataInputExampleFile, [2](#)

BaitIdentityExampleFile, [2](#)

DataInputExampleFile, [2](#)

sfinx, [3](#)