Package ‘SMAD’

May 30, 2024

Type Package

Title Statistical Modelling of AP-MS Data (SMAD)

Version 1.20.0

Description Assigning probability scores to protein interactions captured in affinity purification mass spectrometry (AP-MS) expriments to infer protein-protein interactions. The output would facilitate non-specific background removal as contaminants are commonly found in AP-MS data.

Depends R (>= 3.6.0), RcppAlgos

Imports magrittr (>= 1.5), dplyr, stats, tidyr, utils, Rcpp (>= 1.0.0)

LinkingTo Rcpp

License MIT + file LICENSE

Encoding UTF-8

LazyData true

RoxygenNote 7.1.0

Suggests knitr, rmarkdown, testthat, BiocStyle

VignetteBuilder knitr

biocViews MassSpectrometry, Proteomics, Software

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

git_branch RELEASE_3_19

git_last_commit 9ac2533

git_last_commit_date 2024-04-30

Repository Bioconductor 3.19

Date/Publication 2024-05-29

Author Qingzhou Zhang [aut, cre]

Maintainer Qingzhou Zhang <zqzneptune@hotmail.com>
CompPASS Comparative Proteomic Analysis Software Suite (CompPASS) is based on spoke model. This algorithm was developed by Dr. Mathew Sowa for defining the human deubiquitinating enzyme interaction landscape (Sowa, Mathew E., et al., 2009). The implementation of this algorithm was inspired by Dr. Sowa's online tutorial (http://besra.hms.harvard.edu/ipmsmsdbs/cgi-bin/tutorial.cgi). The output includes Z-score, S-score, D-score and WD-score. This function also computes entropy and normalized WD-score. The source code for this function was based on the source code. https://github.com/dnusinow/cRomppass

Usage

CompPASS(datInput)

Arguments

datInput A dataframe with column names: idRun, idBait, idPrey, countPrey. Each row represent one unique protein captured in one pull-down experiment

Value

A data frame consists of unique bait-prey pairs with Z-score, S-score, D-score and WD-score indicating interacting probabilities.

Author(s)

Qingzhou Zhang, <zqzneptune@hotmail.com>

References


Examples

data(TestDatInput)
datScore <- CompPASS(TestDatInput)
head(datScore)

Description

DICE The Dice coefficient is used to score the interaction affinity between two proteins.

Usage

DICE(datInput)

Arguments

datInput A dataframe with column names: idRun, idPrey. Each row represent one unique protein captured in one pull-down experiment.

Value

A dataframe consists of pairwise combination of preys identified in the input with DICE scores.

Author(s)

Qingzhou Zhang, <zqzneptune@hotmail.com>

References


Examples

data(TestDatInput)
datScore <- DICE(TestDatInput)
head(datScore)
Description
Hart Scoring algorithm based on a hypergeometric distribution error model (Hart et al., 2007).

Usage
Hart(datInput)

Arguments
datInput A dataframe with column names: idRun, idPrey. Each row represents one unique protein captured in one pull-down experiment.

Value
A dataframe consists of pairwise combinations of preys identified in the input with Hart scores indicating interacting probabilities computed from negative log transformed Hypergeometric test P-values.

Author(s)
Qingzhou Zhang, <zqzneptune@hotmail.com>

References

Examples
data(TestDatInput)
datScore <- Hart(TestDatInput)
head(datScore)
HGScore Scoring algorithm based on a hypergeometric distribution error model (Hart et al., 2007) with incorporation of NSAF (Zybailov, Boris, et al., 2006). This algorithm was first introduced to predict the protein complex network of Drosophila melanogaster (Guruharsha, K. G., et al., 2011). This scoring algorithm was based on matrix model.

Usage
HG(datInput)

Arguments
datInput A dataframe with column names: idRun, idPrey, countPrey, lenPrey. Each row represent one unique protein captured in one pull-down experiment.

Value
A dataframe consists of pairwise combindation of preys identified in the input with HG scores indicating interacting probabilities computed from negative log transformed Hypergeometric test P-values.

Author(s)
Qingzhou Zhang, <zqzneptune@hotmail.com>

References

Examples
data(TestDatInput)
datScore <- HG(TestDatInput)
head(datScore)
Description

PE Incorporated both spoke and matrix model.

Usage

PE(datInput, rBait = 0.37, cntPseudo = 1)

Arguments

datInput: A dataframe with column names: idRun, idBait, idPrey. Each row represents one unique protein captured in one pull-down experiment.

rBait: The value of the 'r' parameter as described in the publication.

cntPseudo: The value of the 'pseudo count' parameter.

Value

A dataframe consists of protein-protein interactions from both Spoke and Matrix model.

Author(s)

Qingzhou Zhang, <zqzneptune@hotmail.com>

References


Examples

data(TestDatInput)
datScore <- PE(TestDatInput, 0.37, 1)
head(datScore)
TestDatInput

Test data for SMAD

Description

It is a subset of unfiltered BioPlex 2.0 consisting of apoptosis as bait proteins

Usage

data(TestDatInput)

Format

A data frame with 5000 rows and 5 variables

Details

• idRun A unique identifier for one AP-MS run
• idBait A unique identifier for the bait protein
• idPrey A unique identifier for the prey protein
• countPrey Spectra/Peptider count for the prey protein
• lenPrey Protein length for the prey protein
Index

* datasets
  TestDatInput, 7
CompPASS, 2
DICE, 3
Hart, 4
HG, 5
PE, 6
TestDatInput, 7