Package 'rmimp'

March 30, 2015

Type Package

Title Predicting the impact of mutations on kinase-substrate phosphorylation

Version 1.0

Date 2013-10-29

Author Omar Wagih

Maintainer Omar Wagih <wagih@ebi.ac.uk>

Description No description

License LGPL

R topics documented:

`mimp` *Predict the impact of single variants on phosphorylation.*

Description

This function takes in mutation, sequence and phosphorylation data to predict the impact the mutation has on phosphorylation.

Usage

```
mimp(muts, seqs, psites, prob.thresh = 0.5, log2.thresh = 1,
  display.results = T, include.cent = F, model.data = "hconf")
```

Arguments

`muts`	Mutation data file: a space delimited text file OR data frame containing two columns (1) gene and (1) mutation. Example:

TP53	R282W
CTNNB1	S33C
CTNNB1	S37F

`seqs` Sequence data file containing protein sequences in FASTA format OR named list of sequences where each list element is the uppercase sequence and the name of each element is that of the protein. Example: list(TP53="ABCXYZ", CDK2="HJKEWR")

`psites` Phosphorylation data file (optional): a space delimited text file OR data frame containing two columns (1) gene and (1) positions of phosphorylation sites. Example:

TP53	280
CTNNB1	29
CTNNB1	44

`prob.thresh` Probability threshold of gains and losses. This value should be between 0.5 and 1.

`log2.thresh` Threshold for the absolute value of log ratio between wild type and mutant scores. Anything less than this value is discarded (default: 1).

`include.cent` If TRUE, gains and losses caused by mutation in the central STY residue are kept. Scores of peptides with a non-STY central residue is given a score of 0 (default: FALSE).

`model.data` Name of specifcity model data to use, can be "hconf" : individual experimental kinase specificity models used to scan for rewiring events. For experimental kinase specificity models, grouped by family, set to "hconf-fam". Both are considered high confidence. For lower confidence predicted specificity models , set to "lconf". NOTE: Predicted models are purely speculative and should be used with caution

Value

The data is returned in a `data.frame` with the following columns:

`gene` Gene with the rewiring event

`mut` Mutation causing the rewiring event

`psite_pos` Position of the central residue of the phosphosite

`mut_dist` Distance of the mutation relative to the central phosphosite

`wt` Sequence of the wildtype phosphosite (before the mutation)

`mt` Sequence of the mutated phosphosite (after the mutation)

`score_wt` Matrix similarity score of the wildtype phosphosite

`score_mt` Matrix similarity score of the mutated phosphosite

`log_ratio` Log2 ratio between mutant and wildtype scores. A high positive log ratio represents a high confidence gain-of-signaling event. A high negative log ratio represents a high confidence loss-of-signaling event. This ratio is NA for mutations that affect the central phosphorylation sites

`pwm` Name of the kinase being rewiried

`prob` Joint probability of wild type sequence belonging to the foreground distribution and mutated sequence belonging to the background distribution, for loss and vice versa for gain

`effect` Type of rewiring event, can be "loss" or "gain"

nseqs	Number of sequences used to construct the PWM. PWMs constructed with a higher number of sequences are generally considered of better quality.
pwm_fam	Family/subfamily of kinase being rewired. If a kinase subfamily is available the family and subfamily will be seprated by an underscore e.g. "DMPK_ROCK". If no subfamily is available, only the family is shown e.g. "GSK"

Examples

```
# Get the path to example mutation data
mut.file = system.file("extdata", "mutation_data.txt", package = "rmimp")

# Get the path to example FASTA sequence data
seq.file = system.file("extdata", "sequence_data.txt", package = "rmimp")

# View the files in a text editor
browseURL(mut.file)
browseURL(seq.file)

# Run rewiring analysis
results = mimp(mut.file, seq.file, display.results=TRUE)

# Show head of results
head(results)
```

---

results2html *Display MIMP results interactively in browser*

Description

Display MIMP results interactively in browser

Usage

```
results2html(x, max.rows = 5000)
```

Arguments

x	Data frame resulting from mimp call.
max.rows	If data contains more rows than this value, results won't be displayed.