Package 'KSEAapp'

May 7, 2026

Title Kinase-Substrate Enrichment Analysis

Version 2.0

Description Infers relative kinase activity from phosphoproteomics data using the method described by Casado et al. (2013) <doi:10.1126/scisignal.2003573>.

License MIT + file LICENSE

Encoding UTF-8

Depends R (>= 2.10)

Imports gplots, graphics, stats

Suggests knitr, rmarkdown

VignetteBuilder knitr

LazyData true

RoxygenNote 7.3.3

NeedsCompilation no

Author Danica Wiredja [aut, cre]

Maintainer Danica Wiredja <`dwiredja@gmail.com`>

Repository CRAN

Date/Publication 2026-03-13 20:20:02 UTC

Contents

`KSData` *Kinase-Substrate (K-S) Relationship Dataset*

Description

K-S annotations from PhosphoSitePlus and NetworKIN predictions; This is an abbreviated version of the full dataset used purely for demonstration; please go to the GitHub page for access to the complete file: github.com/casecpb/KSEA/

Usage

```
data(KSData)
```

Format

abbreviated dataframe containing the kinase-substrate annotations and source

References

Hornbeck et al. (2015) Nucleic Acids Res. 43:D512-20

Horn et al. (2014) Nature Methods 11(6):603-4

`KSEA.Barplot` *The KSEA App Analysis (KSEA Bar Plot Only)*

Description

Takes a formatted phoshoproteomics data input and returns just the summary bar plot of kinase scores This function no longer offers the option to export results as a tiff file

Usage

```
KSEA.Barplot(KSData, PX, NetworKIN, NetworKIN.cutoff, m.cutoff, p.cutoff)
```

Arguments

`KSData`	the Kinase-Substrate dataset uploaded from the file prefaced with "PSP&NetworKIN_" available from github.com/casecpb/KSEA/
`PX`	the experimental data file formatted as described in the documentation
`NetworKIN`	a binary input of TRUE or FALSE, indicating whether or not to include NetworKIN predictions; NetworKIN = TRUE means inclusion of NetworKIN predictions
`NetworKIN.cutoff`	a numeric value between 1 and infinity setting the minimum NetworKIN score (can be left out if NetworKIN = FALSE)

`m.cutoff`	a numeric value between 0 and infinity indicating the min. # of substrates a kinase must have to be included in the bar plot output
`p.cutoff`	a numeric value between 0 and 1 indicating the p-value cutoff for indicating significant kinases in the bar plot

Value

creates the bar plot output highlighting key kinase results

References

Wiredja et al. (2017) Bioinformatics 33(21):3489-3491

Casado et al. (2013) Sci Signal. 6(268):rs6

Hornbeck et al. (2015) Nucleic Acids Res. 43:D512-20

Horn et al. (2014) Nature Methods 11(6):603-4

Examples

```
KSEA.Barplot(KSData, PX, NetworKIN=TRUE, NetworKIN.cutoff=5,
             m.cutoff=8, p.cutoff=0.05)
KSEA.Barplot(KSData, PX, NetworKIN=FALSE, m.cutoff=2, p.cutoff=0.05)
```

`KSEA.Heatmap` *The KSEA App Analysis (KSEA Heatmap Only)*

Description

Takes a list of the KSEA kinase score outputs from KSEA.Scores() and creates a merged heatmap (only applicable for multi-treatment studies)

Usage

```
KSEA.Heatmap(
  score.list,
  sample.labels,
  stats,
  m.cutoff,
  p.cutoff,
  sample.cluster
)
```

Arguments

`score.list`	the data frame outputs from the KSEA.Scores() function, compiled in a list format
`sample.labels`	a character vector of all the sample names for heatmap annotation; the names must be in the same order as the data in score.list; please avoid long names, as they may get cropped in the final image
`stats`	character string of either "p.value" or "FDR" indicating the data column to use for marking statistically significant scores
`m.cutoff`	a numeric value between 0 and infinity indicating the min. # of substrates a kinase must have to be included in the heatmap
`p.cutoff`	a numeric value between 0 and 1 indicating the p-value/FDR cutoff for indicating significant kinases in the heatmap
`sample.cluster`	a binary input of TRUE or FALSE, indicating whether or not to perform hierarchical clustering of the sample columns

Value

creates a heatmap highlighting the merged datasets; heatmap was generated using the heatmap.2() function (gplots package); asterisks mark scores that met the statistical cutoff, as defined by p.cutoff; blue color indicates negative kinase score, and red indicates positive kinase score

References

Wiredja et al. (2017) Bioinformatics 33(21):3489-3491

Casado et al. (2013) Sci Signal. 6(268):rs6

Hornbeck et al. (2015) Nucleic Acids Res. 43:D512-20

Horn et al. (2014) Nature Methods 11(6):603-4

Examples

```
#The score.list input must be a list of the data frame outputs from KSEA.Scores() function
#KSEA.Scores.1, KSEA.Scores.2, and KSEA.Scores.3 are all
#sample datasets provided within this package

KSEA.Heatmap(score.list=list(KSEA.Scores.1, KSEA.Scores.2, KSEA.Scores.3),
             sample.labels=c("Tumor.A", "Tumor.B", "Tumor.C"),
             stats="p.value", m.cutoff=3, p.cutoff=0.05, sample.cluster=TRUE)
```

`KSEA.KS_table` *The KSEA App Analysis (K-S Dataset Only)*

Description

Takes a formatted phoshoproteomics data input and returns just the kinase-subtrate (K-S) annotations used for KSEA calculations

Usage

```
KSEA.KS_table(KSData, PX, NetworKIN, NetworKIN.cutoff)
```

Arguments

`KSData`	the Kinase-Substrate dataset uploaded from the file prefaced with "PSP&NetworKIN_" available from github.com/casecpb/KSEA/
`PX`	the experimental data file formatted as described in the documentation
`NetworKIN`	a binary input of TRUE or FALSE, indicating whether or not to include NetworKIN predictions; NetworKIN = TRUE means inclusion of NetworKIN predictions
`NetworKIN.cutoff`	a numeric value between 1 and infinity setting the minimum NetworKIN score (can be left out if NetworKIN = FALSE)

Value

creates a new data frame in R with all kinase-substrate relationships used for the KSEA calculations

References

Wiredja et al. (2017) Bioinformatics 33(21):3489-3491

Casado et al. (2013) Sci Signal. 6(268):rs6

Hornbeck et al. (2015) Nucleic Acids Res. 43:D512-20

Horn et al. (2014) Nature Methods 11(6):603-4

Examples

```
KSData.dataset = KSEA.KS_table(KSData, PX, NetworKIN=TRUE, NetworKIN.cutoff=3)
KSData.dataset = KSEA.KS_table(KSData, PX, NetworKIN=FALSE)
```

`KSEA.Scores` *The KSEA App Analysis (KSEA Kinase Scores Only)*

Description

Takes a formatted phoshoproteomics data input and returns just the KSEA kinase scores and statistics

Usage

```
KSEA.Scores(KSData, PX, NetworKIN, NetworKIN.cutoff)
```

Arguments

`KSData`	the Kinase-Substrate dataset uploaded from the file prefaced with "PSP&NetworKIN_" available from github.com/casecpb/KSEA/
`PX`	the experimental data file formatted as described in the documentation
`NetworKIN`	a binary input of TRUE or FALSE, indicating whether or not to include NetworKIN predictions; NetworKIN = TRUE means inclusion of NetworKIN predictions
`NetworKIN.cutoff`	a numeric value between 1 and infinity setting the minimum NetworKIN score (can be left out if NetworKIN = FALSE)

Value

creates a new data frame in R with all the KSEA kinase scores, along with each one's statistical assessment

References

Wiredja et al. (2017) Bioinformatics 33(21):3489-3491

Casado et al. (2013) Sci Signal. 6(268):rs6

Hornbeck et al. (2015) Nucleic Acids Res. 43:D512-20

Horn et al. (2014) Nature Methods 11(6):603-4

Examples

```
scores = KSEA.Scores(KSData, PX, NetworKIN=TRUE, NetworKIN.cutoff=3)
scores = KSEA.Scores(KSData, PX, NetworKIN=FALSE)
```

`KSEA.Scores.1` *One of the 3 datasets for heatmap plotting*

Description

A sample KSEA.Scores output generated from the KSEA.Scores() function

Usage

```
data(KSEA.Scores.1)
```

Format

dataframe containing 7 columns in the exact order as listed below.

"KinaseGene" the HUGO gene name of the kinase

"mS" the mean log2FC of all the kinase's identified substrates

"Enrichment" the enrichment score (refer to Casado et al. (2013) Sci. Signal., 6, rs6-rs6)

"m" the number of experimentally-identifed substrates annotating to that kinase

"z.score" the normalized kinase score

"p.value" the statistical assessment of the kinase score

"FDR" the p-value adjusted for multiple hypothesis testing by the Benjamin-Hochberg method

References

unpublished data

`KSEA.Scores.2` *One of the 3 datasets for heatmap plotting*

Description

A sample KSEA.Scores output generated from the KSEA.Scores() function

Usage

```
data(KSEA.Scores.2)
```

Format

dataframe containing 7 columns in the exact order as listed below.

"KinaseGene" the HUGO gene name of the kinase

"mS" the mean log2FC of all the kinase's identified substrates

"Enrichment" the enrichment score (refer to Casado et al. (2013) Sci. Signal., 6, rs6-rs6)

"m" the number of experimentally-identifed substrates annotating to that kinase

"z.score" the normalized kinase score

"p.value" the statistical assessment of the kinase score

"FDR" the p-value adjusted for multiple hypothesis testing by the Benjamin-Hochberg method

References

unpublished data

`KSEA.Scores.3` *One of the 3 datasets for heatmap plotting*

Description

A sample KSEA.Scores output generated from the KSEA.Scores() function

Usage

```
data(KSEA.Scores.3)
```

Format

dataframe containing 7 columns in the exact order as listed below.

"KinaseGene" the HUGO gene name of the kinase

"mS" the mean log2FC of all the kinase's identified substrates

"Enrichment" the enrichment score (refer to Casado et al. (2013) Sci. Signal., 6, rs6-rs6)

"m" the number of experimentally-identifed substrates annotating to that kinase

"z.score" the normalized kinase score

"p.value" the statistical assessment of the kinase score

"FDR" the p-value adjusted for multiple hypothesis testing by the Benjamin-Hochberg method

References

unpublished data

`PX` *PX dataset for KSEA calculations*

Description

A sample PX dataset of the experimental phosphoproteomics input

Usage

```
data(PX)
```

Format

the experimental data file must be formatted exactly as described below; must have 6 columns in the exact order: Protein, Gene, Peptide, Residue.Both, p, FC; cannot have NA values, or else the entire peptide row is deleted; Description of each column in PX:

"Protein" the Uniprot ID for the parent protein

"Gene" the HUGO gene name for the parent protein

"Peptide" the peptide sequence

"Residue.Both" all phosphosites from that peptide, separated by semicolons if applicable; must be formatted as the single amino acid abbrev. with the residue position (e.g. S102)

"p" the p-value of that peptide (if none calculated, please write "NULL", cannot be NA)

"FC" the fold change (not log-transformed); usually the control sample is the denominator

References

unpublished data

Index