Package ‘HD2013SGI’

August 15, 2024

Type Package

Title Mapping genetic interactions in human cancer cells with RNAi and multiparametric phenotyping

Version 1.44.0

Author Bernd Fischer

Maintainer Mike Smith <mike.smith@embl.de>


License Artistic-2.0

LazyLoad true

Depends R (>= 2.10.0), RColorBrewer, gplots, geneplotter, splots, limma, vcd, LSD,EBImage

Suggests BiocStyle

SystemRequirements GNU make

biocViews ExperimentData, CancerData, ColonCancerData, MicrotitrePlateAssayData, CellCulture, Homo_sapiens_Data, HighThroughputImagingData

git_url https://git.bioconductor.org/packages/HD2013SGI

git_branch RELEASE_3_19

git_last_commit 1014851

git_last_commit_date 2024-06-28

Repository Bioconductor 3.19

Date/Publication 2024-08-15
Contents

HD2013SGI-package ......................................................... 2
datamatrix ............................................................... 3
datamatrixfull ............................................................. 4
featuresPerWell .......................................................... 5
HD2013SGIHeatmapHuman .................................................. 6
HD2013SGImaineffects ..................................................... 7
HD2013SGIOrderDim ....................................................... 8
HD2013SGIselectByStability .............................................. 8
Interactions ............................................................... 10
mainEffects ............................................................. 11
nrOfInteractionsPerTarget ................................................ 12
QueryAnnotation .......................................................... 12
stabilitySelection ......................................................... 13
TargetAnnotation .......................................................... 14

Index 15

HD2013SGI-package  Experimental Data and Analysis of the HCT116 Genetic Interaction Matrix

Description

Experimental Data and Analysis of the HCT116 Genetic Interaction Matrix. This package contains the data and source code for the paper Laufer, Fischer, Billmann, Huber, Boutros, HD2013SGI, 2013.

Details

Package: HD2013SGI
Type: Package
Version: 0.0.3
License: Artistic-2.0
LazyLoad: true
Imports: rhdf5, RColorBrewer, gplots, geneplotter, MASS, grid, hwriter, splots, igraph, abind, limma, vcd, LSD
SystemRequirements: GNU make
biocViews: Infrastructure
Built: R 2.15.1; ; 2013-02-14 12:08:09 UTC; unix

The interaction matrix can be loaded by data(Interactions, package="HD2013SGI"). Type ?Interactions to see a documentation of the interaction data.
The vignette of the package can be seen by typing `library("HD2013SGI")` > `vignette("HD2013SGI")` It contains the complete documentation and R-code for the analysis of the data published in the original publication.

All intermediate results are precomputed and can be loaded. The following datasets are available:

- **featuresPerWell**
  - The screen data in screen order
- **datamatrixfull**
  - The phenotype data of all pairwise genetic perturbation experiments before quality control and feature selection
- **QueryAnnotation**
  - Annotation of all the query genes in the screen
- **TargetAnnotation**
  - Annotation of all target genes in the screen
- **stabilitySelection**
  - Results from the feature selection step
- **datamatrix**
  - The phenotype data of all pairwise genetic perturbation experiments after quality control and feature selection
- **mainEffects**
  - estimated main effects (single knock down effects)
- **nrOfInteractionsPerTarget**
  - number of interactions per target gene
- **Interactions**
  - The genetic interaction data (pi-scores, p-values)

A number of helper functions are defined in the package and used in the vignette.

- **HD2013SGIorderDim**
  - hclust on one out of three dimensions of an interaction matrix
- **HD2013SGIHeatmapHuman**
  - plotting a heatmap of a three dimensional array of pi-scores (target genes x query genes x features)
- **HD2013SGImaineffects**
  - estimation main effects
- **HD2013SGIselectByStability**
  - feature selection to select features most stable between replicated experiments

**Author(s)**

Bernd Fischer
Maintainer: Bernd Fischer <bernd.fischer@embl.de>

**References**


**Examples**

```r
data(Interactions, package="HD2013SGI")
```

<table>
<thead>
<tr>
<th>datamatrix</th>
<th>Phenotypic data after quality control and feature selection</th>
</tr>
</thead>
</table>

**Description**

Phenotypic features of pairwise genetic perturbation experiments **after** selection of non-redundant features and quality control. D is the 6-dimensional array of experimental measurements. Its dimensions are target genes x target siRNA designs x query genes x query siRNA designs x features x replicates. The array has a dimnames attribute, but there exists a more comprehensive annotation of target genes, query genes, and phenotypes in **Anno**.
Phenotypic data before quality control and feature selection

**Description**

Phenotypic features of pairwise genetic perturbation experiments **before** selection of non-redundant features and quality control. \( \mathbf{D} \) is the 6-dimensional array of experimental measurements. Its dimensions are target genes x target siRNA designs x query genes x query siRNA designs x features x replicates. The array has a dimnames attribute.

**Usage**

data(datamatrixfull)
featuresPerWell

Format

The format is: List of 1 $ D: num [1:345, 1:2, 1:20, 1:2, 1:353, 1:2] 2686 2573 2650 3000 2733 ...
...
.$ targetGene : chr [1:345] "B1" "B2" "B3" "B4" ...
...
.$ targetDesign: chr [1:2] "1" "2" ...
...
.$ queryGene : chr [1:20] "01" "02" "03" "04" ...
...
.$ queryDesign : chr [1:2] "1" "2" ...
...
.$ features : chr [1:353] "count" "nuc.0.m.cx" "nuc.0.m.cy" "nuc.0.m.majoraxis" ...
...
.$ replicate : chr [1:2] "1" "2"

References

Laufer, Fischer et al., 2013

See Also

HD2013SGI

Examples

data(datamatrixfull, package="HD2013SGI")
plot(datamatrixfull$D[,1,1,1,1,1])

featuresPerWell

Original phenotypic measurements in screen order

Description

Original phenotypic features in screen order. Anno contains the annotation for each experiment including the plate name, row, col, and field. data is a data.frame with a column for each phenotypic feature and rows as much as there are experiments in the screen.

Usage

data(featuresPerWell)

Format

The format is: List of 2 $ Anno:'data.frame': 231840 obs. of 4 variables: ..$ plate: chr [1:231840] "001CIQ01IRI" "001CIQ01IRI" "001CIQ01IRI" "001CIQ01IRI" ...
...
.$ row : chr [1:231840] "B" "B" "B" ...
...
.$ col : chr [1:231840] "1" "1" "1" "1" ...
...
.$ field: chr [1:231840] "1" "2" "3" "4" ...

$ data: num [1:231840, 1:353] 2780 3120 2242 2603 2170 ...
...
.$: NULL ...
.
.$: chr [1:353] "count" "nuc.0.m.cx" "nuc.0.m.cy" "nuc.0.m.majoraxis" ...

References

Laufer, Fischer et al., 2013

See Also

HD2013SGI
Examples

data(featuresPerWell, package="HD2013SGI")
plot(log2(featuresPerWell$data[,1]),pch="."

HD2013SGIHeatmapHuman

Plotting heatmaps of genetic interaction scores

Description

Plotting heatmaps of three-dimensional arrays of interaction scores. Two dimensions of the array will be flattened.

Usage

HD2013SGIHeatmapHuman(x, cuts, col, colnames = TRUE, rownames = FALSE, mrow = 10, mcol = 10, cexrow = 1, cexcol = 1, border = 0.1, space = 0.05)

Arguments

x
A three dimensional array to be plotted as a heatmap.
cuts
cuts on the values of x for color coding. length(cuts) has be one larger than length(col).
col
Values of x are mapped on color definitions as defined in col using the cuts argument.
colnames
Gene names for columns of the matrix.
rownames
Gene names for rows of the matrix.
mrow
row margin for printing gene names.
mcol
column margin for printing gene names.
cexrow
cex for the row names.
cexcol
cex for the column names.
border
line width of the border.
space
spacing between elements of the third array dimension of x after flattening.

Value

Nothing is returned.

Author(s)

Bernd Fischer

References

Laufer, Fischer et al., 2013

See Also

HD2013SGI
HD2013SGImaineffects  Estimation of main effects

Description

A function to estimate main effects (single knock-down effects) in genetic interaction screens.

Usage

HD2013SGImaineffects(x, TP, TargetNeg, QueryNeg, eps = 1e-04, maxiter = 100, na.rm = TRUE)

Arguments

- **x**: Two dimensional array.
- **TP**: Assignment of target genes to target plates. Used to compute target main effects for each target plate separately.
- **TargetNeg**: Negative controls within the set of target genes.
- **QueryNeg**: Negative controls within the set of query genes.
- **eps**: Real number greater than 0. A tolerance for convergence.
- **maxiter**: The maximum number of iterations.
- **na.rm**: Logical. Should missing values be removed?

Value

- **neg**: Effect of the negative control.
- **targetMainEffect**: Target main effects.
- **queryMainEffect**: Query main effects.
- **pi**: Pairwise interaction scores (pi-scores).

Author(s)

Changes applied by Bernd Fischer to the implementation of R stats function `medpolish`.

References

Laufer, Fischer et al., 2013

See Also

HD2013SGI
HD2013SGIorderDim  

hcclust on one out of three dimensions of an interaction matrix

Description

hcclust on one out of three dimensions of a three-dimensional array of interaction scores (target genes x query genes x features)

Usage

HD2013SGIorderDim(x, i)

Arguments

x  
A three dimensional array to be clustered.
i  
The dimension of the array along which the data is clustered.

Value

Returns a cluster hierarchy of class hclust.

Author(s)

Bernd Fischer

References

Laufer, Fischer et al., 2013

See Also

HD2013SGI

HD2013SGIselectByStability

Feature selection method

Description

A function to select features that are most stable across replicated experiments

Usage

HD2013SGIselectByStability(subsample, preselect = "count", Rdim = 40, verbose = TRUE)
Arguments

subsample This is the input data. Usually a subsample of the complete screen is enough to select the non-redundant features. subsample is a list with three elements: D is a three-dimensional array with dimensions samples x features x replicates. As samples usually 1000 to 5000 experiments are randomly selected. The function needs two replicates.

preselect Names of the features that should be preselected, e.g. count is preselected in this screen, because of its biological interpretability and comparability to other viability-based genetic interaction screens.

Rdim The maximum number of features to be selected.

verbose If TRUE information about the progress and the quality of the selected features is printed.

Value

(selected = selected, correlation = correlation, ratioPositive = ratioPositive, correlationAll = correlationAll)

selected The names of the selected features in the order as selected.

correlation The correlation of the residual features after fitting a linear function on the previously selected features. Correlations are in same order as selected.

ratioPositive The fraction of positively correlated features among all candidate features in each step. In same order as selected.

correlationAll The correlation of the residual features of all candidate features in each step of the selection process.

Author(s)

Bernd Fischer

References

Laufer, Fischer et al., 2013

See Also

HD2013SGI
Interactions

Description

The genetic interaction data. Pairwise interaction scores (\textit{piscore}) are presented in a 6-dimensional array with dimensions target genes x target siRNA designs x query genes x query siRNA designs x features x replicates. BH-corrected p-values (\textit{padj}) are presented in a 5-dimensional array with dimensions target genes x target siRNA designs x query genes x query siRNA designs x features. An annotation of target and query genes and of phenotypes can be found in (\textit{Anno}). \textit{scale} is the standard deviation measure used for normalization. At first standard deviations were computed between replicates for each experiment and afterwards the median of standard deviations was computed.

Usage

\texttt{data(Interactions)}

Format

The format is: List of 4 $ piscore: \text{num}[1:282, 1:2, 1:20, 1:2, 1:2] -1.814 -2.457 -3.094 -1.448 -0.142 ... ... attr(, "dimnames")=List of 6 ...$ targetGene : chr[1:282] "TDRD6" "PRDM11" "KDM1B" "INTS12" ... ...$ targetDesign: chr[1:2] "1" "2" ... ...$ queryGene : chr[1:20] "DPF2" "SMARCA1" "SMARCC1" "SMARCD2" ... ...$ queryDesign : chr[1:2] "1" "2" ... ...$ features : chr[1:11] "count" "cell.act.m.majoraxis" "nuc.nuc.b.q001" "nuc.0.m.eccentricity" ... ...$ replicate : chr[1:2] "1" "2" $ scale : \text{num}[1:11] 0.214 0.125 0.139 0.193 0.207 ... $ padj : \text{num}[1:282, 1:2, 1:20, 1:2, 1:11] 0.6838 0.4167 0.0949 0.5786 0.7933 ... ... attr(, "dimnames")=List of 5 ...$ targetGene : chr[1:282] "TDRD6" "PRDM11" "KDM1B" "INTS12" ... ...$ targetDesign: chr[1:2] "1" "2" ... ...$ queryGene : chr[1:20] "DPF2" "SMARCA1" "SMARCC1" "SMARCD2" ... ...$ queryDesign : chr[1:2] "1" "2" ... ...$ features : chr[1:11] "count" "cell.act.m.majoraxis" "nuc.nuc.b.q001" "nuc.0.m.eccentricity" ... $ Anno : List of 3 ...$ target : \text{data.frame}: 282 obs. of 4 variables: ... ...$ ID : chr[1:282] "B1" "B2" "B3" "B4" ... ...$ Symbol: chr[1:282] "TDRD6" "PRDM11" "KDM1B" "INTS12" ... ...$ Well : chr[1:282] "B1" "B2" "B3" "B4" ... ...$ group : chr[1:282] "sample" "sample" ... ...$ phenotype: chr[1:282] "count" "cell.act.m.majoraxis" "nuc.nuc.b.q001" "nuc.0.m.eccentricity" ...

References

Laufer, Fischer et al., 2013

See Also

HD2013SGI
mainEffects

Examples

```r
data(Interactions, package="HD2013SGI")
plot(Interactions$piscore[,1,"SUV39H1","cell.act.m.majoraxis",1])
print(names(which(Interactions$padj[,1,"SUV39H1","cell.act.m.majoraxis"] <= 0.01)))
```

mainEffects

Estimated main effects

Description

Estimated main effects (single knock-down effects) for target and query genes. Additional overall effects for each phenotype are contained. The dataset contains an annotation of target genes, query genes, and phenotypes.

Usage

```r
data(mainEffects)
```

Format

The format is: List of 4

$ target: num [1:289, 1:2, 1:11, 1:2] -0.31065 -0.32253 -0.08466 -0.00367 -0.60867 ...
  - attr(*, "dimnames")=List of 5
  ..$ targetGene : chr [1:289] "TDRD6" "PRDM11" "KDM1B" "INTS12" ...
  ..$ targetDesign: chr [1:2] "1" "2"
  ..$ queryDesign : chr [1:2] "1" "2"
  ..$ features : chr [1:11] "count" "cell.act.m.majoraxis" "nuc.nuc.b.q001"
  ..$ replicate : chr [1:2] "1" "2"

$ query: num [1:2, 1:20, 1:2, 1:11, 1:2] 0.277 0.265 0.235 0.226 1.165 ...
  - attr(*, "dimnames")=List of 5
  ..$ targetDesign: chr [1:2] "1" "2"
  ..$ queryGene : chr [1:20] "DPF2" "SMARCA1" "SMARCC1" "SMARCD2" ...
  ..$ queryDesign : chr [1:2] "1" "2"
  ..$ features : chr [1:11] "count" "cell.act.m.majoraxis" "nuc.nuc.b.q001" "nuc.0.m.eccentricity"
  ..$ replicate : chr [1:2] "1" "2"

$ overall: num [1:2, 1:2, 1:11, 1:2] 0.3685 0.4638 0.3331 0.487 -0.0985 ...
  - attr(*, "dimnames")=List of 4
  ..$ targetDesign: chr [1:2] "1" "2"
  ..$ queryDesign : chr [1:2] "1" "2"
  ..$ features : chr [1:11] "count" "cell.act.m.majoraxis" "nuc.nuc.b.q001" "nuc.0.m.eccentricity"
  ..$ replicate : chr [1:2] "1" "2"

$ Anno: List of 3
  ..$ target: data.frame: 289 obs. of 4 variables:
  .. ..$ ID : chr [1:289] "B1" "B2" "B3" "B4" ...
  .. ..$ Symbol: chr [1:289] "B1" "B2" "B3" "B4" ...
  .. ..$ Well: chr [1:289] "B1" "B2" "B3" "B4" ...
  .. ..$ group: chr [1:289] "sample" "sample" "sample" "sample" ...
  ..$ query: data.frame: 20 obs. of 2 variables:
  .. ..$ ID : chr [1:20] "01" "02" "03" "04" ...
  ..$ phenotype: chr [1:20] "DPF2" "SMARCA1" "SMARCC1" "SMARCD2" ...

References

Laufer, Fischer et al., 2013

See Also

HD2013SGI
nrOfInteractionsPerTarget

Number of interactions per target gene

Description

Number of genetic interactions per target gene.

Usage

data(nrOfInteractionsPerTarget)

Format

The format is: int [1:282] 1 3 0 2 0 3 1 1 0 1 ...

References

Laufer, Fischer et al., 2013

See Also

HD2013SGI

Examples

data(nrOfInteractionsPerTarget, package="HD2013SGI")
plot(nrOfInteractionsPerTarget)

QueryAnnotation

Annotation of all query genes in the screen

Description

Annotation of all query genes in the screen.

Usage

data(QueryAnnotation)
stabilitySelection

Format
A data frame with 20 observations on the following 2 variables.

ID a character vector
Symbol a character vector

References
Laufer, Fischer et al., 2013

See Also
HD2013SGI

Examples

data(QueryAnnotation, package="HD2013SGI")
print(QueryAnnotation$Symbol)

stabilitySelection Results from the feature selection method

Description
Results from the feature selection method.

Usage
data(stabilitySelection)

Format
The format is: List of 4 $ selected: chr [1:25] "count" "cell.act.m.majoraxis" "nuc.nuc.b.q001"
"nuc.0.m.eccentricity" ... $ correlation: num [1:25] 0.917 0.972 0.938 0.928 0.896 ... $ ratioPositive:
num [1:25] 1 1 0.947 0.942 0.937 ... $ correlationAll: List of 25 ..$ : Named num [1:227] 0.917 0.884 0.93 0.897 0.882 ...
..- attr(*, "names")= chr [1:227] "count" "nuc.0.m.majoraxis" "nuc.0.m.eccentricity" "nuc.0.s.area" ...
..$ : Named num [1:225] 0.884 0.934 0.884 0.882 0.883 ... ..- attr(*, "names")= chr [1:225] "nuc.0.m.majoraxis" "nuc.0.m.eccentricity" "nuc.0.s.area" "nuc.0.s.perimeter" ...

Details
selected is a vector of the selected feature names. correlation are the Pearson correlation
coefficients of the residual features. ratioPositive is the fraction of positively correlated features
among all candidate features for selection. correlationAll contains a vector of correlations of the
residual features of all candidate features for each step in the selection process.
**TargetAnnotation**

**References**

Laufer, Fischer et al., 2013

**See Also**

HD2013SGI

**Examples**

```r
data(stabilitySelection, package="HD2013SGI")
barplot(stabilitySelection$correlation,
        names.arg=stabilitySelection$selected, las=2)
barplot(stabilitySelection$ratioPositive-0.5, offset=0.5,
        names.arg=stabilitySelection$selected, las=2)
```

---

**TargetAnnotation**  
Annotation of all target genes in the screen

**Description**

Annotation of the target genes on one target plate. It includes an ENSEMBL gene identifier, the HUGO name, the position on the plate (well), and the group of the target siRNA (sample or control).

**Usage**

```r
data(TargetAnnotation)
```

**Format**

A data frame with 345 observations on the following 4 variables.

- **ID**  a character vector
- **Symbol**  a character vector
- **Well**  a character vector
- **group**  a character vector

**References**

Laufer, Fischer et al., 2013

**See Also**

HD2013SGI

**Examples**

```r
data(TargetAnnotation, package="HD2013SGI")
print(TargetAnnotation$Symbol)
```
Index

* datasets
  datamatrix, 3
  datamatrixfull, 4
  featuresPerWell, 5
  Interactions, 10
  mainEffects, 11
  nrOfInteractionsPerTarget, 12
  QueryAnnotation, 12
  stabilitySelection, 13
  TargetAnnotation, 14

* package
  HD2013SGI-package, 2
  datamatrix, 3, 3
  datamatrixfull, 3, 4
  featuresPerWell, 3, 5
  HD2013SGI, 4–14
  HD2013SGI (HD2013SGI-package), 2
  HD2013SGI-package, 2
  HD2013SGIHeatmapHuman, 3, 6
  HD2013SGImaineffects, 3, 7
  HD2013SGIorderDim, 3, 8
  HD2013SGIselectByStability, 3, 8
  Interactions, 3, 10
  mainEffects, 3, 11
  medpolish, 7
  nrOfInteractionsPerTarget, 3, 12
  QueryAnnotation, 3, 12
  stabilitySelection, 3, 13
  TargetAnnotation, 3, 14