Package ‘gapmap’

October 13, 2022

Title Drawing Gapped Cluster Heatmap with 'ggplot2'

Version 0.1.0

Description The gap encodes the distance between clusters and improves interpretation of cluster heatmaps. The gaps can be of the same distance based on a height threshold to cut the dendrogram. Another option is to vary the size of gaps based on the distance between clusters.

License GPL-2 | GPL-3

Encoding UTF-8

Depends ggplot2, reshape2

Imports grid

Suggests knitr, dendsort, RColorBrewer, rmarkdown

VignetteBuilder knitr

URL `https://github.com/evanbiederstedt/gapmap`

BugReports `https://github.com/evanbiederstedt/gapmap/issues`

RoxygenNote 7.1.1

NeedsCompilation no

Maintainer Evan Biederstedt <`evan.biederstedt@gmail.com`>

Author Ryo Sakai [aut],
Evan Biederstedt [cre, aut]

Repository CRAN

Date/Publication 2021-04-19 09:10:05 UTC

R topics documented:

`gapmap-package` *Draws gapped heatmap (gapmap) and gapped dendrograms using ggplot2 in [R].*

Description

Functions for drawing gapped cluster heatmap with ggplot2

Details

This is a set of tools for drawing gapmaps using `ggplot`

`gap_data` extracts data from a dendrogram object. Make sure to convert `hclust` object to `dendrogram` object by calling `as.dendrogram()`. This method generates an object class `gapdata`, consisting of a list of `data.frames`. The general workflow is as following:

1. Hierarchical clustering `hclust()`
2. Convert the `hclust` output class into `dendrogram` by calling `as.dendrogram()`
3. Generate a gapped cluster heatmap by specifying a `matrix` and `dendrogram` objects for rows and columns in `gapmap()` function

Author(s)

Ryo Sakai <`ryo.sakai@esat.kuleuven.be`>

`gapmap` *Function to draw a gapped cluster heatmap*

Description

This function draws a gapped cluster heatmap using the ggplot2 package. The input for the function is the a matrix, two dendrograms, and parameters for gaps.

Usage

```
gapmap(
  m,
  d_row,
  d_col,
  mode = c("quantitative", "threshold"),
  mapping = c("exponential", "linear"),
  ratio = 0.2,
  scale = 0.5,
  threshold = 0,
  row_threshold = NULL,
```

```
    col_threshold = NULL,
    rotate_label = TRUE,
    verbose = FALSE,
    left = "dendrogram",
    top = "dendrogram",
    right = "label",
    bottom = "label",
   col = c("#053061", "#2166AC", "#4393C3", "#92C5DE", "#D1E5F0", "#F7F7F7", "#FDDBC7",
      "#F4A582", "#D6604D", "#B2182B", "#67001F"),
    h_ratio = c(0.2, 0.7, 0.1),
    v_ratio = c(0.2, 0.7, 0.1),
    label_size = 5,
    show_legend = FALSE,
    ...
)
```

Arguments

`m`	matrix
`d_row`	a dendrogram class object for rows
`d_col`	a dendrogram class object for columns
`mode`	gap mode, either "threshold" or "quantitative"
`mapping`	in case of quantitative mode, either "linear" or "exponential" mapping
`ratio`	the percentage of width allocated for the sum of gaps.
`scale`	the sclae log base for the exponential mapping
`threshold`	the height at which the dendrogram is cut to infer clusters
`row_threshold`	the height at which the row dendrogram is cut
`col_threshold`	the height at which the column dendrogram is cut
`rotate_label`	a logical to rotate column labels or not
`verbose`	logical for whether in verbose mode or not
`left`	a character indicating "label" or "dendrogram" for composition
`top`	a character indicating "label" or "dendrogram" for composition
`right`	a character indicating "label" or "dendrogram" for composition
`bottom`	a character indicating "label" or "dendrogram" for composition
`col`	colors used for heatmap
`h_ratio`	a vector to set the horizontal ratio of the grid. It should add up to 1. top, center, bottom.
`v_ratio`	a vector to set the vertical ratio of the grid. It should add up to 1. left, center, right.
`label_size`	a numeric to set the label text size
`show_legend`	a logical to set whether to show a legend or not
`...`	ignored

Value

a ggplot object

Examples

```
set.seed(1234)
#generate sample data
x <- rnorm(10, mean=rep(1:5, each=2), sd=0.4)
y <- rnorm(10, mean=rep(c(1,2), each=5), sd=0.4)
dataFrame <- data.frame(x=x, y=y, row.names=c(1:10))
#calculate distance matrix. default is Euclidean distance
distxy <- dist(dataFrame)
#perform hierarchical clustering. default is complete linkage.
hc <- hclust(distxy)
dend <- as.dendrogram(hc)
#make a cluster heatmap plot
gapmap(m = as.matrix(distxy), d_row= rev(dend), d_col=dend)
```

`gap_data` *Generate a gapdata class object from a dendrogram object*

Description

This function takes a dendrogram class object as an input, and generate a gapdata class object as an output. By parsing the dendrogram object based on parameters for gaps, gaps between leaves in a dendrogram are introduced, and the coordinates of the leaves are adjusted. The gaps can be based on the a height (or distance) threshold to to introduce the gaps of the same width, or quantitative mapping of distance values mapped linearly or exponentially.

Usage

```
gap_data(
  d,
  mode = c("quantitative", "threshold"),
  mapping = c("exponential", "linear"),
  ratio = 0.2,
  scale = 0.5,
  threshold = 0,
  verbose = FALSE,
  ...
)
```

Arguments

`d`	dendrogram class object
`mode`	gap mode, either "threshold" or "quantitative"

`mapping`	in case of quantitative mode, either "linear" or "exponential" mapping
`ratio`	the percentage of width allocated for the sum of gaps.
`scale`	the sclae log base for the exponential mapping
`threshold`	the height at which the dendrogram is cult to infer clusters
`verbose`	logical for whether in verbose mode or not
`...`	ignored

Value

a list of data frames that contain coordinates for drawing a gapped dendrogram

---

`gap_dendrogram` *Function to draw a gapped dendrogram*

Description

This function draws a gapped dendrogram using the ggplot2 package. The input for the function is the gapdata class object, generated from gap_data() function.

Usage

```
gap_dendrogram(
  data,
  leaf_labels = TRUE,
  rotate_label = FALSE,
  orientation = c("top", "right", "bottom", "left"),
  ...
)
```

Arguments

`data`	gapdata class object
`leaf_labels`	a logical to show labels or not
`rotate_label`	a logical to rotate labels or not
`orientation`	a character to set the orientation of dendrogram. Choices are "top", "right", "bottom", "left".
`...`	ignored

Value

a ggplot object

`gap_heatmap` *Function to draw a gapped heatmap*

Description

This function draws a gapped heatmap using the ggplot2 package. The input for the function are the gapdata class objects, generated from gap_data() function, and the data matrix.

Usage

```
gap_heatmap(
  m,
  row_gap = NULL,
  col_gap = NULL,
  row_labels = TRUE,
  col_labels = TRUE,
  rotate = FALSE,
 col = c("#053061", "#2166AC", "#4393C3", "#92C5DE", "#D1E5F0", "#F7F7F7", "#FDDBC7",
    "#F4A582", "#D6604D", "#B2182B", "#67001F")
)
```

Arguments

`m`	data matrix
`row_gap`	a gapdata class object for rows
`col_gap`	a gapdata class object for columns
`row_labels`	a logical to show labels for rows
`col_labels`	a logical to show lables for columns
`rotate`	a logical to rotate row labels
`col`	colors used for heatmap

Value

a ggplot object

`gap_label` *Function to draw a gapped labels*

Description

This function draws a gapped labels using the ggplot2 package. The input for the function is the gapdata class object, generated from gap_data() function.

Usage

```
gap_label(data, orientation, label_size = 5)
```

Arguments

`data`	gapdata class object
`orientation`	orientation of the labels, "left", "top", "right", or "bottom"
`label_size`	a numeric to set the label text size

Value

a ggplot object

`sample_tcga` *Sample data matrix from the integrated pathway analysis of gastric cancer from the Cancer Genome Atlas (TCGA) study.*

Description

a multivariate table obtained from the integrated pathway analysis of gastric cancer from the Cancer Genome Atlas (TCGA) study. In this data set, each column represents a pathway consisting of a set of genes and each row represents a cohort of samples based on specific clinical or genetic features. For each pair of a pathway and a feature, a continuous value of between 1 and -1 is assigned to score positive or negative association, respectively.

Usage

```
data(sample_tcga)
```

Format

A data frame with 215 rows and 117 variables

Details

We would like to thank Sheila Reynolds and Vesteinn Thorsson from the Institute for Systems Biology for sharing this sample data set.

Index