Package 'EnhancedVolcano'

April 15, 2020

Type Package

Title Publication-ready volcano plots with enhanced colouring and labeling

Version 1.4.0

Author Kevin Blighe, Sharmila Rana, Myles Lewis

Maintainer Kevin Blighe <kevin@clinicalbioinformatics.co.uk>

Description Volcano plots represent a useful way to visualise the results of differential expression analyses. Here, we present a highly-configurable function that produces publication-ready volcano plots. EnhancedVolcano will attempt to fit as many point labels in the plot window as possible, thus avoiding 'clogging' up the plot with labels that could not otherwise have been read. Other functionality allows the user to identify up to 4 different types of attributes in the same plot space via colour, shape, size, and shade parameter configurations.

License GPL-3

Depends ggplot2, ggrepel

Imports

Suggests RUnit, BiocGenerics, knitr, DESeq2, pasilla, airway, gridExtra, magrittr

URL https://github.com/kevinblighe/EnhancedVolcano

biocViews RNASeq, GeneExpression, Transcription, DifferentialExpression, ImmunoOncology

VignetteBuilder knitr

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

git_branch RELEASE_3_10

git_last_commit 6086485

git_last_commit_date 2019-10-29

Date/Publication 2020-04-14

R topics documented:

EnhancedVolcano-package

EnhancedVolcano: Publication-ready volcano plots with enhanced colouring and labeling.

Description

Volcano plots represent a useful way to visualise the results of differential expression analyses. Here, we present a highly-configurable function that produces publication-ready volcano plots. EnhancedVolcano will attempt to fit as many point labels in the plot window as possible, thus avoiding 'clogging' up the plot with labels that could not otherwise have been read. Other functionality allows the user to identify up to 3 different types of attributes in the same plot space via colour, shape, and shade parameter configurations.

EnhancedVolcano   *Publication-ready volcano plots with enhanced colouring and labeling.*

Description

Volcano plots represent a useful way to visualise the results of differential expression analyses. Here, we present a highly-configurable function that produces publication-ready volcano plots. EnhancedVolcano [@EnhancedVolcano] will attempt to fit as many labels in the plot window as possible, thus avoiding 'clogging' up the plot with labels that could not otherwise have been read. Other functionality allows the user to identify up to 3 different types of attributes in the same plot space via colour, shape, size, and shade parameter configurations.

Usage

```
EnhancedVolcano(
  toptable,
  lab,
  x,
  y,
  selectLab = NULL,
  xlim = c(min(toptable[[x]], na.rm=TRUE),
    max(toptable[[x]], na.rm=TRUE)),
  ylim = c(0, max(-log10(toptable[[y]]), na.rm=TRUE) + 5),
  xlab = bquote(~Log[2]~ "fold change"),
  ylab = bquote(~-Log[10]~italic(P)),
  axisLabSize = 18,
  title = 'Volcano plot',
  subtitle = 'EnhancedVolcano',
  caption = paste0('Total = ', nrow(toptable), ' variables'),
  titleLabSize = 18,
  subtitleLabSize = 14,
  captionLabSize = 14,
  pCutoff = 10e-6,
  pLabellingCutoff = pCutoff,
```

```
FCcutoff = 1.0,
cutoffLineType = 'longdash',
cutoffLineCol = 'black',
cutoffLineWidth = 0.4,
transcriptPointSize = 0.8,
transcriptLabSize = 3.0,
transcriptLabCol = 'black',
transcriptLabFace = 'plain',
transcriptLabhjust = 0,
transcriptLabvjust = 1.5,
pointSize = 2.0,
labSize = 3.0,
labCol = 'black',
labFace = 'plain',
labhjust = 0,
labvjust = 1.5,
boxedlabels = FALSE,
boxedLabels = FALSE,
shape = 19,
shapeCustom = NULL,
col = c("grey30", "forestgreen", "royalblue", "red2"),
colCustom = NULL,
colAlpha = 1/2,
legend = c("NS","Log2 FC","P","P & Log2 FC"),
legendLabels = c('NS', expression(Log[2]~FC),
  "p-value", expression(p-value~and~log[2]~FC)),
legendPosition = "top",
legendLabSize = 14,
legendIconSize = 4.0,
legendVisible = TRUE,
shade = NULL,
shadeLabel = NULL,
shadeAlpha = 1/2,
shadeFill = "grey",
shadeSize = 0.01,
shadeBins = 2,
drawconnectors = FALSE,
drawConnectors = FALSE,
widthConnectors = 0.5,
typeConnectors = 'closed',
endsConnectors = 'first',
lengthConnectors = unit(0.01, 'npc'),
colConnectors = 'grey10',
hline = NULL,
hlineType = 'longdash',
hlineCol = 'black',
hlineWidth = 0.4,
vline = NULL,
vlineType = 'longdash',
vlineCol = 'black',
vlineWidth = 0.4,
gridlines.major = TRUE,
```

```
  gridlines.minor = TRUE,
  border = "partial",
  borderWidth = 0.8,
  borderColour = "black")
```

Arguments

toptable	A data-frame of test statistics (if not, a data frame, an attempt will be made to convert it to one). Requires at least the following: column for transcript names (can be rownames); a column for log2 fold changes; a column for nominal or adjusted p-value. REQUIRED.
lab	A column name in toptable containing transcript names. Can be rownames(toptable). REQUIRED.
x	A column name in toptable containing log2 fold changes. REQUIRED.
y	A column name in toptable containing nominal or adjusted p-values. REQUIRED.
selectLab	A vector containing a subset of lab. DEFAULT = NULL. OPTIONAL.
xlim	Limits of the x-axis. DEFAULT = c(min(toptable[,x], na.rm=TRUE), max(toptable[,x], na.rm=TRUE)). OPTIONAL.
ylim	Limits of the y-axis. DEFAULT = c(0, max(-log10(toptable[,y]), na.rm=TRUE) + 5). OPTIONAL.
xlab	Label for x-axis. DEFAULT = bquote(~Log[2]~ "fold change"). OPTIONAL.
ylab	Label for y-axis. DEFAULT = bquote(~-Log[10]~italic(P)). OPTIONAL.
axisLabSize	Size of x- and y-axis labels. DEFAULT = 18. OPTIONAL.
title	Plot title. DEFAULT = 'Volcano plot'. OPTIONAL.
subtitle	Plot subtitle. DEFAULT = 'EnhancedVolcano'. OPTIONAL.
caption	Plot caption. DEFAULT = paste0('Total = ', nrow(toptable), ' variables'). OPTIONAL.
titleLabSize	Size of plot title. DEFAULT = 18. OPTIONAL.
subtitleLabSize	
	Size of plot subtitle. DEFAULT = 14. OPTIONAL.
captionLabSize	Size of plot caption. DEFAULT = 14. OPTIONAL.
pCutoff	Cut-off for statistical significance. A horizontal line will be drawn at -log10(pCutoff). DEFAULT = 10e-6. OPTIONAL.
pLabellingCutoff	
	Labelling cut-off for statistical significance. DEFAULT = pCutoff. OPTIONAL
FCcutoff	Cut-off for absolute log2 fold-change. Vertical lines will be drawn at the negative and positive values of log2FCcutoff. DEFAULT = 1.0. OPTIONAL.
cutoffLineType	Line type for FCcutoff and pCutoff ("blank", "solid", "dashed", "dotted", "dotdash", "longdash", "twodash"). DEFAULT = "longdash". OPTIONAL.
cutoffLineCol	Line colour for FCcutoff and pCutoff. DEFAULT = "black". OPTIONAL.
cutoffLineWidth	
	Line width for FCcutoff and pCutoff. DEFAULT = 0.4. OPTIONAL.
transcriptPointSize	
	transcriptPointSize argument deprecated in v1.4 - please use pointSize. DEFAULT = 0.8. OPTIONAL.

transcriptLabSize	transcriptLabSize argument deprecated in v1.4 - please use labSize. DEFAULT = 3.0. OPTIONAL.
transcriptLabCol	transcriptLabCol argument deprecated in v1.4 - please use labCol. DEFAULT = 'black'. OPTIONAL.
transcriptLabFace	transcriptLabFace argument deprecated in v1.4 - please use labFace. DEFAULT = 'plain'. OPTIONAL.
transcriptLabhjust	transcriptLabhjust argument deprecated in v1.4 - please use labhjust. DEFAULT = 0. OPTIONAL.
transcriptLabvjust	transcriptLabvjust argument deprecated in v1.4 - please use labvjust. DEFAULT = 1.5. OPTIONAL.
pointSize	Size of plotted points for each transcript. Can be a single value or a vector of sizes. DEFAULT = 2.0. OPTIONAL.
labSize	Size of labels for each transcript. DEFAULT = 3.0. OPTIONAL.
labCol	Colour of labels for each transcript. DEFAULT = 'black'. OPTIONAL.
labFace	Font face of labels for each transcript. DEFAULT = 'plain'. OPTIONAL.
labhjust	Horizontal adjustment of label for each transcript. DEFAULT = 0. OPTIONAL.
labvjust	Vertical adjustment of label for each transcript. DEFAULT = 1.5. OPTIONAL.
boxedlabels	boxedlabels argument deprecated in v1.4 - please use boxedLabels. DEFAULT = FALSE. OPTIONAL.
boxedLabels	Logical, indicating whether or not to draw labels in boxes. DEFAULT = FALSE. OPTIONAL.
shape	Shape of the plotted points. Either a single value for all points, or 4 values corresponding to $< abs(FCcutoff) \&\& > pCutoff$, $> abs(FCcutoff)$, $< pCutoff$, $> abs(FCcutoff) \&\& < pCutoff$. DEFAULT = 19. OPTIONAL.
shapeCustom	Named vector / key-value pairs that will over-ride the default shape scheme. The order must match that of toptable. Names / keys relate to groups / categories; values relate to shape encodings. DEFAULT = NULL. OPTIONAL.
col	Colour shading for plotted points, corresponding to $< abs(FCcutoff) \&\& > pCutoff$, $> abs(FCcutoff)$, $< pCutoff$, $> abs(FCcutoff) \&\& < pCutoff$. DEFAULT = c("grey30", "forestgreen", "royalblue", "red2"). OPTIONAL.
colCustom	Named vector / key-value pairs that will over-ride the default colour scheme. The order must match that of toptable. Names / keys relate to groups / categories; values relate to colour. DEFAULT = NULL. OPTIONAL.
colAlpha	Alpha for purposes of controlling colour transparency of transcript points. DEFAULT = 1/2. OPTIONAL.
legend	Plot legend key. DEFAULT = c("NS", "Log2 FC", "P", "P & Log2 FC"). OPTIONAL.
legendLabels	Plot legend text labels. DEFAULT = c('NS', expression(Log[2]~FC), "p-value", expression(p-value~and~log[2]~FC). OPTIONAL
legendPosition	Position of legend ("top", "bottom", "left", "right"). DEFAULT = "top". OPTIONAL.
legendLabSize	Size of plot legend text. DEFAULT = 14. OPTIONAL.

legendIconSize	Size of plot legend icons / symbols. DEFAULT = 4.0. OPTIONAL.
legendVisible	Logical, indicating whether or not to show the legend. DEFAULT = TRUE. OPTIONAL.
shade	A vector of transcript names to shade. DEFAULT = NULL. OPTIONAL.
shadeLabel	Label for the transcrips to shade. DEFAULT = NULL. OPTIONAL.
shadeAlpha	Alpha for purposes of controlling colour transparency of shaded regions. DE-FAULT = 1/2. OPTIONAL.
shadeFill	Colour of shaded regions. DEFAULT = "grey". OPTIONAL.
shadeSize	Size of the shade contour lines. DEFAULT = 0.01. OPTIONAL.
shadeBins	Number of bins for the density of the shade. DEFAULT = 2. OPTIONAL.
drawconnectors	drawconnectors argument deprecated since v1.2 - please use drawConnectors. DEFAULT = FALSE. OPTIONAL.
drawConnectors	Logical, indicating whether or not to connect plot labels to their corresponding points by line connectors. DEFAULT = FALSE. OPTIONAL.
widthConnectors	
	Line width of connectors. DEFAULT = 0.5. OPTIONAL.
typeConnectors	Have the arrow head open or filled ('closed')? ('open', 'closed'). DEFAULT = 'closed'. OPTIONAL.
endsConnectors	Which end of connectors to draw arrow head? ('last', 'first', 'both'). DEFAULT = 'first'. OPTIONAL.
lengthConnectors	
	Length of the connectors. DEFAULT = unit(0.01, 'npc'). OPTIONAL
colConnectors	Line colour of connectors. DEFAULT = 'grey10'. OPTIONAL.
hline	Draw one or more horizontal lines passing through this/these values on y-axis. For single values, only a single numerical value is necessary. For multiple lines, pass these as a vector, e.g., c(60,90). DEFAULT = NULL. OPTIONAL.
hlineType	Line type for hline ('blank', 'solid', 'dashed', 'dotted', 'dotdash', 'longdash', 'twodash'). DEFAULT = 'longdash'. OPTIONAL.
hlineCol	Colour of hline. DEFAULT = 'black'. OPTIONAL.
hlineWidth	Width of hline. DEFAULT = 0.4. OPTIONAL.
vline	Draw one or more vertical lines passing through this/these values on x-axis. For single values, only a single numerical value is necessary. For multiple lines, pass these as a vector, e.g., c(60,90). DEFAULT = NULL. OPTIONAL.
vlineType	Line type for vline ('blank', 'solid', 'dashed', 'dotted', 'dotdash', 'longdash', 'twodash'). DEFAULT = 'longdash'. OPTIONAL.
vlineCol	Colour of vline. DEFAULT = 'black'. OPTIONAL.
vlineWidth	Width of vline. DEFAULT = 0.4. OPTIONAL.
gridlines.major	
	Logical, indicating whether or not to draw major gridlines. DEFAULT = TRUE. OPTIONAL
gridlines.minor	
	Logical, indicating whether or not to draw minor gridlines. DEFAULT = TRUE. OPTIONAL
border	Add a border for just the x and y axes ('partial') or the entire plot grid ('full')? DEFAULT = 'partial'. OPTIONAL.
borderWidth	Width of the border on the x and y axes. DEFAULT = 0.8. OPTIONAL.
borderColour	Colour of the border on the x and y axes. DEFAULT = "black". OPTIONAL.

Details

Volcano plots represent a useful way to visualise the results of differential expression analyses. Here, we present a highly-configurable function that produces publication-ready volcano plots [@EnhancedVolcano]. EnhancedVolcano will attempt to fit as many transcript names in the plot window as possible, thus avoiding 'clogging' up the plot with labels that could not otherwise have been read.

Value

A [ggplot2](#) object.

Author(s)

Kevin Blighe <kevin@clinicalbioinformatics.co.uk>

Examples

```
library("pasilla")
pasCts <- system.file("extdata", "pasilla_gene_counts.tsv",
  package="pasilla", mustWork=TRUE)
pasAnno <- system.file("extdata", "pasilla_sample_annotation.csv",
  package="pasilla", mustWork=TRUE)
cts <- as.matrix(read.csv(pasCts,sep="\t",row.names="gene_id"))
coldata <- read.csv(pasAnno, row.names=1)
coldata <- coldata[,c("condition","type")]
rownames(coldata) <- sub("fb", "", rownames(coldata))
cts <- cts[, rownames(coldata)]
library("DESeq2")
dds <- DESeqDataSetFromMatrix(countData = cts,
  colData = coldata,
  design = ~ condition)

featureData <- data.frame(gene=rownames(cts))
mcols(dds) <- DataFrame(mcols(dds), featureData)
dds <- DESeq(dds)
res <- results(dds)

EnhancedVolcano(res,
  lab = rownames(res),
  x = "log2FoldChange",
  y = "pvalue",
  pCutoff = 10e-4,
  FCcutoff = 1.333,
  xlim = c(-5.5, 5.5),
  ylim = c(0, -log10(10e-12)),
  pointSize = 1.5,
  labSize = 2.5,
  shape = c(6, 6, 19, 16),
  title = "DESeq2 results",
  subtitle = "Differential expression",
  caption = "FC cutoff, 1.333; p-value cutoff, 10e-4",
  legendPosition = "right",
  legendLabSize = 14,
  col = c("grey30", "forestgreen", "royalblue", "red2"),
  colAlpha = 0.9,
  drawConnectors = TRUE,
```

```
    hline = c(10e-8),
    widthConnectors = 0.5)
```

Index