Package 'EnhancedVolcano'

April 15, 2019

Type Package

Title Publication-ready volcano plots with enhanced colouring and labeling

Version 1.0.1

Author Kevin Blighe

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 transcript names in the plot window as possible, thus avoiding 'clogging' up the plot with labels that could not otherwise have been read.

License GPL-3

Depends R (>= 3.5.0), ggplot2, ggrepel

Imports

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

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

biocViews ImmunoOncology, RNASeq, GeneExpression, Transcription, DifferentialExpression

VignetteBuilder knitr

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

git_branch RELEASE_3_8

git_last_commit 45c76cc

git_last_commit_date 2019-01-04

Date/Publication 2019-04-15

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 transcript names in the plot window as possible, thus avoiding 'clogging' up the plot with labels that could not otherwise have been read.

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 transcript names in the plot window as possible, thus avoiding 'clogging' up the plot with labels that could not otherwise have been read.

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 = 16,
    pCutoff = 0.05,
    pLabellingCutoff = pCutoff,
    FCcutoff = 2.0,
    title = "",
    titleLabSize = 16,
    transcriptPointSize = 0.8,
    transcriptLabSize = 3.0,
    col = c("grey30", "forestgreen", "royalblue", "red2"),
    colOverride = NULL,
    colAlpha = 1/2,
    legend = c("NS", "Log2 FC", "P", "P & Log2 FC"),
    legendPosition = "top",
    legendLabSize = 10,
```

```
        legendIconSize = 3.0,
        DrawConnectors = FALSE,
        widthConnectors = 0.5,
        colConnectors = "black",
        cutoffLineType = "longdash",
        cutoffLineCol = "black",
        cutoffLineWidth = 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. Only values in selectLab that pass FCcutoff and pCutoff thresholds will be labelled in the plot. 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 = 16. OPTIONAL.
pCutoff	Cut-off for statistical significance. A horizontal line will be drawn at -log10(pCutoff). DEFAULT = 0.05. 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 FCCutoff. DEFAULT = 2.0. OPTIONAL.
title	Plot title. DEFAULT = "". OPTIONAL.
titleLabSize	Size of plot title. DEFAULT = 16. OPTIONAL.
transcriptPointSize	
	Size of plotted points for each transcript. DEFAULT = 0.8. OPTIONAL.
transcriptLabSize	
	Size of labels for each transcript. DEFAULT = 3.0. 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.

colOverride	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 = 0.5. OPTIONAL.
legend	Plot legend text. DEFAULT = c("NS", "Log2 FC", "P", "P & Log2 FC"). OPTIONAL.
legendPosition	Position of legend ("top", "bottom", "left", "right"). DEFAULT = "top". OPTIONAL.
legendLabSize	Size of plot legend text. DEFAULT = 10. OPTIONAL.
legendIconSize	Size of plot legend icons / symbols. DEFAULT = 3.0. OPTIONAL.
DrawConnectors	Fit labels onto plot and connect to their respective points by lines (TRUE/FALSE). DEFAULT = FALSE. OPTIONAL.
widthConnectors	Line width of connectors to plot points. DEFAULT = 0.5. OPTIONAL.
colConnectors	Line colour of connectors to plot points. DEFAULT = "black". 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.
gridlines.major	Draw major gridlines? (TRUE/FALSE). DEFAULT = TRUE. OPTIONAL
gridlines.minor	Draw minor gridlines? (TRUE/FALSE). 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 = "padj",
    ylab = bquote(~-Log[10]~adjusted~italic(P)),
    pCutoff = 10e-4,
    FCcutoff = 1.333,
    xlim=c(-5.5, 5.5),
    ylim=c(0, -log10(10e-12)),
    transcriptLabSize = 3.5,
    title = "DESeq2 results",
    legendPosition = "right",
    legendLabSize = 14,
    col = c("grey30", "forestgreen", "royalblue", "red2"),
    colAlpha=0.9,
    DrawConnectors = TRUE,
    widthConnectors=0.2)
```

Index