Package 'CMplot'

May 24, 2022

Type Package

Title Circle Manhattan Plot

Version 4.1.0

Date 2022-05-23

Author LiLin-Yin

Maintainer LiLin-Yin <ylilin@163.com>

Description Manhattan plot, a type of scatter plot, was widely used to display the association re-
sults. However, it is usually time-consuming and laborious for a non-
specialist user to write scripts and adjust parameters of an elaborate plot. Moreover, the ever-
growing traits measured have necessitated the integration of results from different Genome-
wide association study researches. Circle Manhattan Plot is the first open R pack-
age that can lay out. Genome-wide association study P-value results in both traditional rectangu-
lar patterns, QQ-plot and novel circular ones. United in only one bull's eye style plot, associa-
tion results from multiple traits can be compared interactively, thereby to reveal both similari-
ties and differences between signals. Additional functions include: highlight sig-
nals, a group of SNPs, chromosome visualization and candidate genes around SNPs.

License GPL (>= 2)

Depends R(>= 2.10)

NeedsCompilation no

URL https://github.com/YinLiLin/CMplot

BugReports https://github.com/YinLiLin/CMplot/issues/new

Repository CRAN

Date/Publication 2022-05-24 02:20:06 UTC

R topics documented:

CMplot-package   *Circle Manhattan Plot*

Description

Manhattan plot, a type of scatter plot, was widely used to display the association results. However, it is usually time-consuming and laborious for a non-specialist user to write scripts and adjust parameters of an elaborate plot. Moreover, the ever-growing traits measured have necessitated the integration of results from different Genome-wide association study researches. Circle Manhattan Plot is the first open R package that can lay out Genome-wide association study P-value results in QQ-plot, both traditional rectangular patterns and novel circular ones. United in only one bull's eye style plot, association results from multiple traits can be compared interactively, thereby to reveal both similarities and differences between signals. Additional functions include: highlight signals, a group of SNPs, chromosome visualization and candidate genes around SNPs.

Usage

```
CMplot(Pmap, col=c("#4197d8", "#f8c120", "#413496", "#495226",
    "#d60b6f", "#e66519", "#d581b7", "#83d3ad", "#7c162c", "#26755d"),
    bin.size=1e6, bin.range=NULL, bin.legend.num=10, pch=19, type="p",
    band=1, H=1.5, ylim=NULL, cex.axis=1, lwd.axis=1.5, cex.lab=1.5,
    plot.type="b", multracks=FALSE, points.alpha=100L, cex=c(0.5,1,1),
    r=0.3, outward=FALSE, ylab=expression(-log[10](italic(p))),
    ylab.pos=3, xticks.pos=1, mar = c(3,6,3,3), threshold = NULL,
    threshold.col="red", threshold.lwd=1, threshold.lty=2,
    amplify= TRUE, signal.cex = 1.5, signal.pch = 19,
    signal.col=NULL, signal.line=2, highlight=NULL, highlight.cex=1,
    highlight.pch=19, highlight.type="p", highlight.col="red",
    highlight.text=NULL, highlight.text.col="black", highlight.text.cex=1,
    highlight.text.xadj=NULL, highlight.text.yadj=NULL,
    highlight.text.font=3, chr.labels=NULL, chr.border=FALSE,
    chr.labels.angle=0, chr.den.col="black", chr.pos.max=FALSE, cir.band=1,
    cir.chr=TRUE, cir.chr.h=1.5, cir.legend=TRUE, cir.legend.cex=0.6,
    cir.legend.col="black", LOG10=TRUE, box=FALSE, conf.int=TRUE,
    conf.int.col=NULL, file.output=TRUE, file=c("jpg","pdf","tiff"),
    dpi=300, height=NULL, width=NULL, memo="", main="", main.cex=1.5,
    main.font=2, trait.legend.ncol=NULL, verbose=TRUE)
```

Arguments

Pmap	a dataframe, at least four columns. The first column is the name of SNP, the second column is the chromosome of SNP, the third column is the position of SNP, and the remaining columns are the P-value of each trait(Note:each trait a column).
col	a vector or a matrix, if "col" is a vector, each circle use the same colors, it means that the same chromosome is drewed in the same color, the colors are not fixed, one, two, three or more colors can be used, if the length of the "col" is shorter

than the length the chromosome, then colors will be applied circularly. If "col" is a matrix, the row is the number of circles(traits), the columns are the colors that users want to use for different circles, each circle can be plotted in different number of colors, the missing value can be replaced by NA. For example: col=matrix(c("grey30","grey60",NA,"red","blue","green","orange",NA,NA),3,3,byrow=T).

bin.size	the size of bin for SNP_density plot.
bin.range	a vector, c(min, max). The min/max value of legend of SNP_density plot, the bin whose SNP number is smaller/bigger than 'bin.range' will use the same color.
bin.legend.num	a number, the number of breaks of legend for SNP_density plot, the default is 10.
pch	a number, the shape for the points or for traits of multi-traits Manhattan plot, is the same with "pch" in \<plot>.
type	a character, could be "p"(point), "l"(line), "h"(vertical lines) and so on, is the same with "type" in \<plot>.
band	a number, the space between chromosomes, the default is 1(if the band equals to 0, then there would be no space between chromosomes).
H	a number, the height for each circle, each circle represents a trait, the default is 1.
ylim	vector or list, CMplot will only plot the points among this interval, ylim can be vector or list, if it is a list, different traits can be assigned with different range at y-axis.
cex.axis	a number, controls the size of ticks labels of X/Y-axis and the size of chromosome labels of circle plot.
lwd.axis	a number, controls the thickness of X/Y-axis lines.
cex.lab	a number, controls the size of labels of X/Y-axis.
plot.type	a character or vector, only "d", "c", "m", "q" or "b" can be used. if plot.type="d", SNP density will be plotted; if plot.type="c", only circle-Manhattan plot will be plotted; if plot.type="m",only Manhattan plot will be plotted; if plot.type="q",only Q-Q plot will be plotted;if plot.type="b", both circle-Manhattan, Manhattan and Q-Q plots will be plotted; if plot.type=c("m","q"), Both Manhattan and Q-Q plots will be plotted.
multracks	a logical,if multracks=FALSE, all Manhattan plots will be drew in separated files, if it is TRUE, all Manhattan plots will be plotted in one file.
points.alpha	a number, the transparency of point for multiple traits plot, ranges from '0' (fully transparent) to '255' (opaque).
cex	a number or a vector, the size for the points, is the same with "size" in \<plot>, and if it is a vector, the first number controls the size of points in circle plot(the default is 0.5), the second number controls the size of points in Manhattan plot(the default is 1), the third number controls the size of points in Q-Q plot(the default is 1)
r	a number, the radius for the circle(the inside radius), the default is 1.
outward	logical, if outward=TRUE,then all points will be plotted from inside to outside for circular Manhattan plot.

ylab	a character, the labels for y axis.
ylab.pos	the distance between ylab and y axis.
xticks.pos	the distance between x ticks labels and x axis.
mar	the size of white gaps around the plot, 4 values should be provided, indicating the direction of bottom, left, up, and right.
threshold	a number or vector or list, the significant threshold. For example, Bonfferoni adjustment method: threshold=0.01/nrow(Pmap), more than one significant line can be added on the plots, if it is a list, each trait can be assigned with different thresholds, if threshold=0 or NULL, then the threshold line will not be added.
threshold.col	a character or vector, the colour for the line of threshold levels, it can also control the color of the diagonal line of QQplot.
threshold.lwd	a number or vector, the width for the line of threshold levels, it can also control the thickness of the diagonal line of QQplot.
threshold.lty	a number or vector, the type for the line of threshold levels, it can also control the type of the diagonal line of QQplot.
amplify	logical, CMplot can amplify the significant points, if amplify=T, then the points bigger than the minimal significant level will be amplified, the default: amplify=TRUE.
signal.cex	a number, if amplify=TRUE, users can set the size of significant points.
signal.pch	a number, if amplify=TRUE, users can set the shape of significant points.
signal.col	a character, if amplify=TRUE, users can set the colour of significant points, if signal.col=NULL, then the colors of significant points will not be changed.
signal.line	a number, the thickness of the lines of significant SNPs cross the circle.
highlight	a vector or list, names of SNPs which need to be highlighted. vector for single trait and list for multiple traits.
highlight.cex	a vector or list, the size of points for SNPs which need to be highlighted.
highlight.pch	a vector or list, the pch of points for SNPs which need to be highlighted.
highlight.type	a vector or list, the type of points for SNPs which need to be highlighted.
highlight.col	a vector or list, the col of points for SNPs which need to be highlighted.
highlight.text	a vector or list, the text which would be added around the highlighted SNPs.
highlight.text.col	
	a vector or list, the color for added text.
highlight.text.cex	
	a value, the size for added text.
highlight.text.xadj	
	value or vecter for single trait or list for multiple traits, -1, 0, 1 limited, control the position of text around the highlighted SNPs, -1(left), 0(center), 1(right).
highlight.text.yadj	
	value or vecter for single trait or list for multiple traits, -1, 0, 1 limited, control the position of text around the highlighted SNPs, -1(down), 0(center), 1(up).
highlight.text.font	
	text font for the highlighted SNPs.

chr.labels	a vector, the labels for the chromosomes of density plot and circle-Manhattan plot.
chr.border	logical, whether to plot the dot line between chromosomes.
chr.labels.angle	
	a value, adjust the angle of tick labels of x-axis for Manhattan plot (-90 < chr.labels.angle < 90).
chr.pos.max	logical, whether the physical positions of each chromosome contain the maximum length of the chromosome.
chr.den.col	a character or vector or NULL, the colour for the SNP density. If the length of parameter 'chr.den.col' is bigger than 1, SNP density that counts the number of SNP within given size('bin.size') will be plotted around the circle. If chr.den.col=NULL, the density bar will not be attached on the bottom of manhattan plot.
cir.band	a number, the space between circles, the default is 1.
cir.chr	logical, a boundary that represents chromosomes will be plotted on the periphery of a circle, the default is TRUE.
cir.chr.h	a number, the width for the boundary, if cir.chr=FALSE, then this parameter will be useless.
cir.legend	logical, whether to add the legend of each circle.
cir.legend.cex	a number, the size of the number of legend.
cir.legend.col	a character, the color of the axis of legend.
LOG10	logical, whether to change the p-value into log10(p-value).
box	logical, this function draws a box around the current Manhattan plot.
conf.int	logical, whether to plot confidence interval on QQ-plot.
conf.int.col	character or vector, the color of confidence interval of QQplot.
file.output	a logical, users can choose whether to output the plot results.
file	a character, users can choose the different output formats of plot, so for, "jpg", "pdf", "tiff" can be selected by users.
dpi	a number, the picture resolution for .jpg and .tiff files. The default is 300.
height	the height of output files.
width	the width of output files.
memo	add a character to the output file name.
main	character of vector, the title of the plot for manhattan plot and qqplot.
main.cex	size of title.
main.font	font of title.
trait.legend.ncol	
	number of column for the legend of multiple traits manhattan plot.
verbose	whether to print the log information.

Details

Package:	CMplot
Type:	Package
Version:	4.1.0
Date:	2022-05-23
License:	GPL(>=2)

Note

1: In the data "Pmap", autosome and heterosome can both exist simultaneously, and heterosome can also be replaced by a number.

2: only when the amplify=TRUE,then the parameter signal.cex, signal.pch, signal.col will be efficacious.

3: In the circle Manhattan plot, traits will be plotted from inside to outside,it means that the first trait occupies the innermost circle.

Author(s)

LiLin-Yin
Maintainer: LiLin-Yin **<ylilin@163.com>**

Examples

```
#plot rectangular Manhattan only one trait with 20000 markers:
data(pig60K)
CMplot(pig60K[sample(1:nrow(pig60K), 10000),c(1:4)],plot.type="m",
  threshold=c(0.01,0.05)/nrow(pig60K),threshold.col=c('red','orange'),
  multracks=FALSE, chr.den.col=NULL, file.output=FALSE)


################################################################################
#plot circular Manhattan for all traits with all markers:                     #
################################################################################
#data(pig60K)                                                                 #
#CMplot(pig60K,col=c("gray30","gray60"),r=0.5,plot.type="c",threshold=0.01,   #
# signal.col=NULL,outward=TRUE)                                               #
################################################################################
```

---

cattle50K   *Genotyped by Bovine50K chip*

Description

This dataset gives the effects of all SNPs using rrBLUP algorithm for 3 traits, individuals were genotyped by Bovine50K chip.

Usage

```
data(cattle50K)
```

Format

A dataframe containing 3 traits' SNPs effects

---

pig60K	*Genotyped by pig 60k chip*

Description

This dataset gives the results of Genome-wide association study of 3 traits, individuals were genotyped by pig 60K chip.

Usage

```
data(pig60K)
```

Format

A dataframe containing 3 traits' Pvalue

Index