

The snp.plotter Package

February 16, 2008

Version 0.2

Date 2007-01-17

Title snp.plotter

Author Augustin Luna <augustin@mail.nih.gov>, Kristin K. Nicodemus <knicodem@jhsph.edu>, <nicodemusk@mail.nih.gov>

Maintainer Augustin Luna <augustin@mail.nih.gov>

Depends R (>= 2.0.0), genetics, grid

Description Creates plots of p-values using single SNP and/or haplotype data. Main features of the package include options to display a linkage disequilibrium (LD) plot and the ability to plot multiple datasets simultaneously. Plots can be created using global and/or individual haplotype p-values along with single SNP p-values. Images are created as either PDF/EPS files.

License GPL version 2 or newer

URL <http://cbdb.nimh.nih.gov/~kristin/snp.plotter.html>

R topics documented:

snp.plotter 1

Index 7

snp.plotter	<i>SNP/haplotype association p-value and linkage disequilibrium plotter</i>
-------------	---

Description

Creates plots of p-values using single SNP and/or haplotype data. Main features of the package include options to display a linkage disequilibrium (LD) plot and the ability to plot multiple set of results simultaneously. Plots can be created using global and/or individual haplotype p-values along with single SNP p-values. Images are created as either PDF/EPS files.

Usage

```
snp.plotter(EVEN.SPACED = FALSE,
            PVAL.THRESHOLD = 1,
            USE.GBL.PVAL = TRUE,
            SYMBOLS = NA,
            SAMPLE.LABELS = NULL,
            LAB.Y = "log",
            DISP.HAP = FALSE,
            DISP.SNP = TRUE,
            DISP.COLOR.BAR = TRUE,
            DISP.PHYS.DIST = TRUE,
            DISP.LEGEND = TRUE,
            DISP.MARKER.LINES = TRUE,
            DISP.LDMAP = FALSE,
            DISP.TYPE = "symbol",
            DISP.MULT.LAB.X = FALSE,
            LD.TYPE = "dprime",
            LD.COLOR.SCHEME = "heat",
            USE.COLORS = TRUE,
            COLOR.LIST = NULL,
            PALETTE.FILE = NULL,
            IMAGE.TITLE = NULL,
            IMAGE.NAME = "snp.plotter",
            IMAGE.TYPE = "pdf",
            IMAGE.SIZE = 3.5,
            CONNECTING.LINES.FACTOR = 1,
            CONNECTING.LINES.ADJ = 0,
            CONNECTING.LINES.FLEX = 0,
            SNP.FILE = NULL,
            HAP.FILE = NULL,
            GENOTYPE.FILE = NULL,
            config.file = NULL
            )
```

Arguments

<code>config.file</code>	Name of a configuration file for snp.plotter parameters in the form ATTRIBUTE=value. This option can be used in place of specifying options at the R command line.
<code>SNP.FILE</code>	Tab-delimited input file containing p-values for single SNPs (see note below). The contents of each SNP.FILE includes four necessary columns ASSOC, SNP.NAME, LOC, and SS.PVAL corresponding to positive or negative association (indicating susceptibility or protective alleles), a SNP label, the location, and a p-value for each SNP. SNP labels cannot start with numbers. Example: "s1.txt,s2.txt,s3.txt" MANDATORY
<code>HAP.FILE</code>	Tab-delimited input file containing p-values for haplotypes (see note below). The contents of each HAP.FILE includes three necessary columns ASSOC,

G.PVAL, and I.PVAL followed by a set of columns of SNPs with corresponding haplotypes. Haplotypes are presented in a step-wise fashion with the major allele given as 1 and the minor allele as 2; haplotype variants for a set of SNPs should be grouped. SNP labels in HAP.FILE must be the same as in SNP.FILE, and only SNPs with corresponding haplotypes need to be included. Example: "h1.txt,h2.txt,h3.txt" OPTIONAL

GENOTYPE.FILE	Tab-delimited input file containing genotypes as a PED file with 6 columns preceding the genotype data: family ID, individual ID, father ID, mother ID, sex, and affection status; these columns are not used in the creation of the LD plot. This file is used for calculating D' or r^2 values for the LD heatmap plot. Only one LD plot can be shown (see note below). OPTIONAL
EVEN.SPACED	Logical. Should the p-values be displayed at even spacing or at genetic map distances?
USE.GBL.PVAL	Logical. Use global haplotype p-values (as opposed to individual p-values)? Unfilled symbols connected by solid lines are used to indicate global haplotype p-values, default symbol: circle. Unfilled and filled symbols are used to indicate alleles 1 and 2, respectively connected by solid lines and dashed lines for positive and negative association (indicating susceptibility or protective haplotypes) when using individual haplotype p-values.
DISP.HAP	Logical. Display haplotype p-values?
DISP.SNP	Logical. Display single SNP p-values?
DISP.LDMAP	Logical. Display the LD heatmap?
DISP.PHYS.DIST	Logical. Display the range of the X-scale?
DISP.LEGEND	Logical. Display a legend with sample labels and corresponding symbols?
DISP.COLOR.BAR	Logical. Display bar showing colors and corresponding values of LD plot?
DISP.TYPE	Options: "symbol"
DISP.MULT.LAB.X	Logical. Display evenly spaced X-axis tick-labels; up to 5 labels are shown.
DISP.MARKER.LINES	Logical. Display lines at p-value thresholds of 0.05, 0.01, 0.001, etc.
USE.COLORS	Logical. Restrict LD heatmap colors and default symbol colors to gray-scale
COLOR.LIST	List of colors (one for each sample) known to GraphApp (see note below) for displaying p-value symbols. Example: "red,blue,green,black,orange"
SYMBOLS	Options: circle, square, diamond, triangle; Symbols can either be filled or not filled by appending "-fill" e.s., square-fill. NA may be specified. In this case, the SNP.FILE ASSOC column is read and an up-triangle and down-triangle are used to indicate positive and negative association (indicating susceptibility or protective alleles), respectively. Example: "circle,NA,diamond-fill,triangle"
PALETTE.FILE	Colors are hexadecimal HTML color codes; one color per line. OPTIONAL
SAMPLE.LABELS	Labels for each sample. Example: "d-cc,d2-cc,d1-fam"

LAB.Y	Options: ln (natural log) or log (log10)
IMAGE.TYPE	Options: "pdf" or "eps"
IMAGE.TITLE	Title of the image in quotes. Note: Title text may not wrap.
IMAGE.SIZE	Options: 3.5 or 7. Sizes are in inches and correspond to 1 or 2 columns per printed page.
IMAGE.NAME	Name of the output file. The correct extension will be appended depending on the value of IMAGE.TYPE
PVAL.THRESHOLD	The minimum value of the the Y-scale will be set to this value. Default: 1 (to ignore option).
LD.TYPE	LD metric. Options: "dprime" or "rsquare"
LD.COLOR.SCHEME	LD heatmap color scheme. Options: heat (red-yellow-white), cm (cyan-magenta), topo (topographical map colors), gray (gray-scale), or custom; custom requires palette file (PALETTE.FILE) to be defined
CONNECTING.LINES.FACTOR	Adjusts the length of the connecting lines. Range: 0-2
CONNECTING.LINES.ADJ	Can be used adjust the position of connecting lines in relation to SNP names. Negative values shift the connecting lines to the left and positive values shift the lines to the right. Range: 0-1
CONNECTING.LINES.FLEX	Adjusts the spread of the connecting lines. Range: 0-2

Details

snp.plotter produces publishable-quality plots of p-values using single SNP and/or haplotype data. Main features of the package include options to display a linkage disequilibrium (LD) plot below the p-value plot using either the r^2 or D' LD metric with a user-specified LD heatmap color scheme, setting the X-axis to equal spacing or to use the physical SNP map, and specification of plot labels, colors and symbols for displaying p-values. A major strength of the package is that it can plot multiple set of results simultaneously. Plots can be created using global and/or individual haplotype p-values along with single SNP p-values. The package provides a simple way to convey both association and LD information in a single appealing graphic and requires virtually no knowledge of the R programming language. Code to create the LD map was modified from the LDHeatmap package by Ji-Hyung Shin, et al. (2006, version 0.2)

Value

A list containing two items: `config.var` and `gbl.var`, which includes the values of all significant variables used by `snp.plotter`

Note

Configuration Files Due to the large number of parameters implemented for flexibility, it is suggested that `snp.plotter` be run using the `config.file` argument.

- Example Datasets** Examples of SNP.FILE, HAP.FILE, GENOTYPE.FILE, and configuration files are provided at <http://cbdb.nimh.nih.gov/~kristin/snp.plotter.html> with further explanation on the file formats.
- Lists** Comma delimited lists (SNP.FILE, HAP.FILE, COLOR.LIST, SYMBOLS, etc) should not have spaces between entries. If using the config.file argument, these lists should not have quotations in the configuration file. Example: "red,blue,green,black,orange"
- Colors** COLOR.LIST colors are limited to those known to GraphApp. A short list can be found at http://en.wikipedia.org/wiki/X11_color_names; the complete list is located in the R source code file
- Palettes** PALETTE.FILE colors are hexadecimal HTML color codes http://en.wikipedia.org/wiki/X11_color_names. The first and last colors correspond to the lowest and highest value of the chosen LD metric, respectively. One color per line.
- PDFs** The error "unable to start device pdf" may occur when attempting to overwrite an open PDF document.
- P-values** A p-value of 1 or NA can be used in SNP.FILE to prevent displaying information about a single SNP
- Number of Datasets** snp.plotter handles 10 set of results, but provides default values for only 5 set of results
- File Input** All input files should be placed in the same directory

Author(s)

Augustin Luna (augustin.luna@mail.nih.gov), Kristin K. Nicodemus (knicodem@jhsph.edu), (nicodemusk@mail.nih.gov).
 Website: <http://cbdb.nimh.nih.gov/~kristin/snp.plotter.html>

Examples

```
## Not run:
  snp.plotter(config.file = "config.txt")
## End(Not run)
## Not run:
  snp.plotter(EVEN.SPACED = FALSE,
             PVAL.THRESHOLD = 1,
             USE.GBL.PVAL = TRUE,
             SYMBOLS = NA,
             SAMPLE.LABELS = NULL,
             LAB.Y = log,
             DISP.HAP = FALSE,
             DISP.SNP = TRUE,
             DISP.COLOR.BAR = TRUE,
             DISP.PHYS.DIST = TRUE,
             DISP.LEGEND = TRUE,
             DISP.MARKER.LINES = TRUE,
             DISP.LDMAP = FALSE,
             DISP.TYPE = "symbol",
             DISP.MULT.LAB.X = FALSE,
             LD.TYPE = "dprime",
             LD.COLOR.SCHEME = "heat",
             USE.COLORS = TRUE,
             COLOR.LIST = NULL,
```

```
    PALETTE.FILE = NULL,  
    FONT.FACTOR = NA,  
    IMAGE.TITLE = NULL,  
    IMAGE.NAME = "snp.plotter",  
    IMAGE.TYPE = "pdf",  
    IMAGE.SIZE = 3.5,  
    CONNECTING.LINES.FACTOR = 1,  
    CONNECTING.LINES.ADJ = 0,  
    CONNECTING.LINES.FLEX = 0,  
    SNP.FILE = NULL,  
    HAP.FILE = NULL,  
    GENOTYPE.FILE = NULL,  
    config.file = NULL  
  )  
## End(Not run)
```

Index

*Topic **aplot**

`snp.plotter`, 1

*Topic **hplot**

`snp.plotter`, 1

`snp.plotter`, 1