cgmisc

Package containing various functions useful in computational genetics,
especially in genome-wide association studies

Marcin Kierczak

August 18, 2014

Introduction

0.1 Synopsis

Package cgmisc contains miscellaneous functions, hopefully useful for extending genome-wide association study (GWAS) analyses.

0.2 Getting help

Like every other R function, the functions provided in this package are documented in the standard R-help (Rd) format and can be easily accessed by issuing **help**() or its shorter version, **?** function. For instance, if you want to get more information on how to use the clump.markers() function, type either help(clumpmarkers) or ?clump.markers and press return/enter. To see this document from within R you type vignette('cgmisc').

0.3 Purpose of this document

This document aims at presenting how to use functions provided in this package in a typical GWAS data analyses workflow. It is, however, not pretending to be a GWAS tutorial as such.

0.4 Conventions

- All R commands are written in terminal type: myfun(foo=T, bar=54)

- In the above example: *myfun* is a function and both *foo* and *bar* are its arguments

Working with `cgmisc`

0.5 Installation

In order to install `cgmisc`, you either use one of the R GUIs (native R GUI, RStudio etc.) or type the following command:

```
install.packages("cgmisc", repos="")
```

Functions in the `cgmisc` package often complement or use `GenABEL` package functions and data structures.`GenABEL` is an excellent and widely-used R package for performing genome-wide association studies and much more...

Therefore `GenABEL` will be loaded automagically when loading cgmisc. You can load `cgmisc` package by as follows:

```
require("cgmisc")

## Loading required package:  cgmisc
## Loading required package:  GenABEL
## Loading required package:  MASS
## Loading required package:  GenABEL.data
##
##  Package cgmisc contains miscellaneous functions, useful for extending
## genome-wide association study (GWAS) analyses.
##
## Package Name:  cgmisc
##  Version:  2.9.3
##  Date:  2014-08-13
##  Author:  Marcin Kierczak <marcin.kierczak@imbim.uu.se>
##  License GPL (>=2.10)
##
##  Package contains various functions useful in computational
##    genetics, especially in genome-wide association studies.
```

After having loaded the package it is time to load some data:

```
#setwd("~/Dropbox/cgmisc-devel/cgmisc_jagoda-devel")
load('data/data.Rd')

## Warning:  cannot open compressed file 'data/data.Rd', probable reason 'No
such file or directory'
## Error:  cannot open the connection
```

0.6   Association Analysis

Some of `cgmisc` functions use data which are the result of GWAS analyses. Let's perform GWAS on our data to obtain `GenABEL` `scan.gwaa-class` object :

```
an0 <- qtscore(response ~ sex, data = data)

## Error:  wrong data class:  should be gwaa.data
```

And have a look at top 5 markers

```
summary(an0, top = 5)

## Error:  object 'an0' not found
```

Once this is done, we can proceed with `cgmisc` functions.

Functions

0.7 Plot.Manhattan.LD

The `plot.Manhattan.LD` function allows you to visualize the LD pattern in a genome fragment on an enchanced Manhattan plot. You select one marker, typically the one with the strongest association to the analysed trait and all other markers in the region are coloured according to the degree of linkage disequilibrium with this index marker.

```
plot.manhattan.LD(data, an0, chr = 34, region = c(3.9e+07, 4.2e+07), index.snp = "BICF2P106334
    bonferroni = F)

## Error:  trying to get slot "gtdata" from an object of a basic class ("function")
with no slots
```

0.8 Clump.markers

`clump.markers` function implements clumping procedure described in PLINK documentation. Clumping is based on linkage disequilibrium. The function returns list of clumps which can be used for further analyses or plotted using `plot.clumps` function included in our package.

```
clumps <- clump.markers(data, gwas.result = an0, chr = 6, bp.dist = 250000,
    p1 = 1e-04, p2 = 0.01, r2 = 0.5, image = T)

## Error:  object 'an0' not found
```