Package 'SHLR'

April 4, 2016

Type Package

Title Shared Haplotype Length Regression

Version 1.0

Date 2016-04-02

Author Satria Sajuthi

Maintainer Satria Sajuthi <ssajuthi@wakehealth.edu>

Description
A statistical method designed to take advantage of population genetics and microevolutionary theory, specifically by testing the association between haplotype sharing length and trait of interest.

License GPL-3

Imports stringr, geepack, Matrix, FactoMineR, kinship2, foreach, MASS, doParallel

NeedsCompilation no

Repository CRAN

Date/Publication 2016-04-04 18:39:37

R topics documented:

`read.shapeit.haps` *Helper function to read phased haplotypes produces by SHAPEIT*

Description

This function reads phased haplotypes produces by SHAPEIT and convert them into a data structure required by run.SHLR.scan()

Usage

```
read.shapeit.haps(infile)
```

Arguments

infile	The phased haplotype file produces by SHAPEIT

Value

haps	Haplotype array consists of one haplotype per row.
marker.map	Marker map contains information about marker name in the 1st column and marker position in the 2nd column

---

run.SHLR.scan *SHLR scan*

Description

This is a wrapper function that performs SHLR or SHLR-fam statistic to a genomic region one marker at a time. The scan results will be outputted to a file.

Usage

```
run.SHLR.scan(haps, phen, marker.map, outfile, outcome,
out.dist = "Gaussian", cov = "", missing.id.code = "0",
method, nSplits = 1, pca.thres = NULL, hap.freq.thres = 0.005,
window.size = 100, threads = NULL, corstr = "kinship",
std.err = "naive", gender = "", fid = "", iid = "", pid = "",
mid = "")
```

Arguments

haps	Haplotypes array stored as a string of character per row. Every two rows contains the haplotypes for each individual. (Required)
phen	Phenotype table. The phenotype table contains information about trait of interest and other covariates. (Required)
marker.map	The marker map contains information about marker name (1st column) and marker position (2nd column). (Required)
outfile	The desired name of the output file. (Required)
outcome	The column name for the phenotype of interest in the phenotype table. (Required)
out.dist	The assumed outcome distribution of the model. e.g. : gaussian, binomial, etc. (Required)
cov	The column names for the covariates in the phenotype table.

`missing.id.code`
: The code for missing id in the pedigree structure. Usually, it is coded as '0'. (Required only if method is 'SHLR-fam' and corstr is 'kinship')

`method`
: The choice for the SHLR method (i.e. : SHLR or SHLR-fam). (Required)

`nSplits`
: The number of splits. This option will determine how often the program prints out results to the output file.

`pca.thres`
: The proportion of variation retained in the design matrix of the model based on PCA.

`hap.freq.thres`
: The minimum haplotype frequency proportion threshold. Haplotypes that meet this threshold will be considered in the model.

`window.size`
: The number of markers that form the haplotype.

`threads`
: The number of CPUs for parallel computation.

`corstr`
: The type of correlation structure. The options are: 'kinship', 'exchangeable', or 'independence'. (SHLR-fam parameter)

`std.err`
: The method for estimating variance for SHLR-fam. The options are: 'naive' or 'sandwich'. (SHLR-fam parameter)

`gender`
: The column name that indicates gender in the phenotype table. (Required if method is SHLR-fam and cor.str is 'kinship')

`fid`
: The column name that indicates family id in the phenotype table. (Required if method is SHLR-fam and cor.str is 'kinship')

`iid`
: The column name that indicates individual id in the phenotype table. (Required if method is SHLR-fam and cor.str is 'kinship')

`pid`
: The column name that indicates father id in the phenotype table. (Required if method is SHLR-fam and cor.str is 'kinship')

`mid`
: The column name that indicates mother id in the phenotype table. (Required if method is SHLR-fam and cor.str is 'kinship')

Examples

```
library(SHLR)

####reading in pedigree data
fam <- read.table(system.file("extdata","plink.fam",package="SHLR"),
    header=FALSE, stringsAsFactors=FALSE)
colnames(fam) <- c("fam_id","id","dad_id","mom_id","sex","qtrait")

####add simulated covariates to create phenotype table
N <- dim(fam)[1]
cov1 <- rnorm(N, 0, 2)
cov2 <- rbinom(N, 1, 0.7)
phen <- data.frame(fam, cov1=cov1, cov2=cov2)

##Read phased haplotype
shapeit <- read.shapeit.haps(system.file("extdata","phased.haps",package="SHLR"))
haps <- shapeit$haps
marker.map <- shapeit$marker.map
```

```
##Initialize parameters for SHLR.scan
outfile <- "out.txt"
outcome <- "qtrait"
cov <- c("sex","cov1","cov2")
famid <- "fam_id"
iid <- "id"
pid <- "dad_id"
mid <- "mom_id"
gender <- "sex"
missid <- "0"
corstr <- "kinship"
out_dist <- "gaussian"
std.err <- "naive"
method <- "SHLR-fam"
window.size <- 35
pca.thres <- 0.8
nSplits <- 1

run.SHLR.scan(haps, phen, marker.map, outfile,
        outcome, out.dist=out_dist, cov=cov, missing.id.code=missid,
        method=method, nSplits=nSplits, pca.thres=pca.thres, hap.freq.thres=0.001,
        window.size=window.size, threads=2, corstr=corstr,
        std.err=std.err, gender=gender, fid=famid, iid=iid, pid=pid, mid=mid)
```

Index