

Package ‘COMBAT’

May 7, 2026

Type Package

Title A Combined Association Test for Genes using Summary Statistics

Version 0.0.4

Date 2018-01-14

Author Minghui Wang, Yiyuan Liu, Shizhong Han

Maintainer Minghui Wang <m.h.wang@live.com>

Description Genome-wide association studies (GWAS) have been widely used for identifying common variants associated with complex diseases. Due to the small effect sizes of common variants, the power to detect individual risk variants is generally low. Complementary to SNP-level analysis, a variety of gene-based association tests have been proposed. However, the power of existing gene-based tests is often dependent on the underlying genetic models, and it is not known a priori which test is optimal. Here we proposed COMBined Association Test (COMBAT) to incorporate strengths from multiple existing gene-based tests, including VEGAS, GATES and simpleM. Compared to individual tests, COMBAT shows higher overall performance and robustness across a wide range of genetic models. The algorithm behind this method is described in Wang et al (2017) <[doi:10.1534/genetics.117.300257](https://doi.org/10.1534/genetics.117.300257)>.

Depends mvtnorm, corpcor, R (>= 3.2.0)

Suggests parallel

License GPL-2

NeedsCompilation no

Repository CRAN

Date/Publication 2018-01-14 21:54:00 UTC

Contents

COMBAT	2
ext_simes	3
gates	4
ld.Rsquare	5
Index	6

COMBAT*A Combined Gene-based Association Test*

Description

This function implements a combined gene-based association test using SNP-level P values and reference genotype data.

Usage

```
COMBAT(x, snp.ref, vegas.pct = c(0.1,0.2,0.3,0.4,1), pca_cut_perc = 0.995,  
       nperm = 100, seed=12345, ncores=1)
```

Arguments

x	a vector of SNP-level P values.
snp.ref	a matrix of SNP genotypes (coded as allele counts) from reference samples, with samples in rows and SNPs in columns.
vegas.pct	a numeric vector, fraction of the top SNPs to be used in the VEGAS method.
pca_cut_perc	numeric, cutoff for percentage of sum of eigen values in the simpleM approach.
nperm	number of permutations for computing the correlation between P values of different tests.
seed	random seed to derive consistent outcome.
ncores	number of CPU cores for parallel computing.

Details

COMBAT uses simulation and the extended Simes procedure ([ext_simes](#)) to combine multiple gene-based association test statistics (currently including [gates](#), [vegas](#), and [simpleM](#)) to perform a more powerful association analysis. This method does not require raw genotype or phenotype data, but needs only SNP-level P-values and correlations between SNPs from ancestry-matched samples. The technical details about the method is described in Wang et al (2017) <doi:10.1534/genetics.117.300257>.

Value

A vector of p-values from COMAT and each individual gene-based test.

References

Minghui Wang, Jianfei Huang, Yiyuan Liu, Li Ma, James B. Potash, Shizhong Han. COMBAT: A Combined Association Test for Genes using Summary Statistics. *Genetics* 2017, 207(3): 883-891. <https://doi.org/10.1534/genetics.117.300257>.

See Also

[ext_simes](#), [gates](#), [vegas](#), [simpleM](#).

Examples

```
# read SNP P values
file1 <- paste(path.package("COMBAT"), "extdata", "SNP_info.txt.gz", sep="/")
snp.info <- read.table(file1, header = TRUE, as.is=TRUE)
snp.pvals <- as.matrix(snp.info[,2])

# read reference genotype
file2 <- paste(path.package("COMBAT"), "extdata", "SNP_ref.txt.gz", sep="/")
snp.ref <- read.table(file2, header = TRUE)
snp.ref <- as.matrix(snp.ref)
#call COMBAT

COMBAT(snp.pvals, snp.ref, nperm=100, ncores=2)
```

ext_simes

Extended Simes Procedure

Description

Combine a vector of test P values by correction for number of independent tests.

Usage

```
ext_simes(x, cor_r)
```

Arguments

x	a vector of SNP-level P values.
cor_r	correlation among P values.

Value

P value.

See Also

[COMBAT](#)

Examples

```
# see ?COMBAT
```

Description

Several gene-based association tests methods are implemented.

Usage

```
gates(x, cor_G)
vegas(x, cor_G, vegas.pct=c(0.1,0.2,0.3,0.4,1), max.simulation=1e6)
simpleM(x, cor_G, pca_cut_perc=0.995)
```

Arguments

<code>x</code>	a vector of SNP-level P values.
<code>cor_G</code>	SNP-SNP correlation matrix.
<code>vegas.pct</code>	a numeric vector, specifying the fraction of the top SNPs to be used in the VEGAS method.
<code>max.simulation</code>	maximum number of simulations to be performed. Must be at least 1e6.
<code>pca_cut_perc</code>	cutoff for percentage of sum of eigen values.

Details

Function `gates` implements the GATES method (Li et al 2011, American Journal of Human Genetics 88:283-293), `vegas` implements VEGAS with different proportion tests (Liu et al 2010, American Journal of Human Genetics 87:139-145), and `simpleM` is the simpleM method (Gao et al 2008, Genetic Epidemiology 32:361-369).

Value

P value(s).

See Also

[COMBAT](#), [ld.Rsquare](#).

Examples

```
# read SNP P values
file1 <- paste(path.package("COMBAT"), "extdata", "SNP_info.txt.gz", sep="/")
snp.info <- read.table(file1, header = TRUE, as.is=TRUE)
snp.pvals <- as.matrix(snp.info[,2])

# read reference genotype
file2 <- paste(path.package("COMBAT"), "extdata", "SNP_ref.txt.gz", sep="/")
snp.ref <- read.table(file2, header = TRUE)
```

```
snp.ref  <- as.matrix(snp.ref)

#compute correlation among SNPs
cor_G <- ld.Rsquare(snp.ref)

#call gates
(pval_gates <- gates(x=snp.pvals, cor_G=cor_G))

#call vegas
(pval_vegas <- vegas(x=snp.pvals, cor_G=cor_G))

#call simpleM
(pval_simpleM <- simpleM(x=snp.pvals, cor_G=cor_G))
```

ld.Rsquare

Linkage Disequilibrium Estimate as Correlation

Description

Compute linkage disequilibrium among SNPs using correlation coefficients.

Usage

```
ld.Rsquare(x)
```

Arguments

x a matrix of SNP genotypes with samples in the rows.

Value

A positive definite correlation matrix.

See Also

[gates](#), [vegas](#), [simpleM](#).

Index

COMBAT, [2](#), [3](#), [4](#)

ext_simes, [2](#), [3](#)

gates, [2](#), [4](#), [5](#)

ld.Rsquare, [4](#), [5](#)

simpleM, [2](#), [5](#)

simpleM(gates), [4](#)

vegas, [2](#), [5](#)

vegas(gates), [4](#)