

Package ‘ofGEM’

July 13, 2018

Title A Meta-Analysis Approach with Filtering for Identifying Gene-Level Gene-Environment Interactions with Genetic Association Data

Version 1.0

Date 2018-07-06

Author Jiebiao Wang, Qianying Liu and Lin S. Chen

Maintainer Jiebiao Wang <randel.wang@gmail.com>

Description Offers a gene-based meta-analysis test with filtering to detect gene-environment interactions (GxE) with association data, proposed by Wang et al. (2018) <doi:10.1002/gepi.22115>. It first conducts a meta-filtering test to filter out unpromising SNPs by combining all samples in the consortia data. It then runs a test of omnibus-filtering-based GxE meta-analysis (ofGEM) that combines the strengths of the fixed- and random-effects meta-analysis with meta-filtering. It can also analyze data from multiple ethnic groups.

License GPL

Depends CompQuadForm, MASS, forestplot

URL <https://github.com/randel/ofGEM>

BugReports <https://github.com/randel/ofGEM/issues>

RoxygenNote 6.0.1

NeedsCompilation no

Repository CRAN

Date/Publication 2018-07-13 14:30:03 UTC

R topics documented:

forest_plot	2
ofGEM	3
sim_dat	4

Index	6
--------------	----------

`forest_plot`*Forest plot*

Description

It makes a forest plot for gene-environment interactions (GxE) across multiple sub-studies for all the SNPs in a set (e.g., a gene).

Usage

```
forest_plot(coef, se, sort = TRUE, exp = FALSE)
```

Arguments

<code>coef</code>	a matrix contains the GxE effect sizes. Each row corresponds to a SNP in a set, and each column represents a study. If the matrix has colnames and/or rownames for SNPs and studies respectively, they will be used as labels in the forest plot.
<code>se</code>	a matrix contains the standard errors for the GxE effect estimates. Each row corresponds to a SNP in a set, and each column represents a study.
<code>sort</code>	logical. If TRUE, the SNPs are sorted by averaged effect sizes across different studies. The default is TRUE.
<code>exp</code>	logical. If TRUE, coef will be exponentiated. This option is mainly used for log odds ratios obtained from logistic regressions. The default is FALSE.

Value

A forest plot across multiple sub-studies for all the SNPs in a set (e.g., a gene).

Examples

```
coef = matrix(rnorm(6 * 6), 6, 6)
se = matrix(abs(runif(6 * 6, 0.1, 0.15)), 6, 6)

forest_plot(coef, se)
```

ofGEM	<i>A meta-analysis approach with filtering for identifying gene-level gene-environment interactions with genetic association data</i>
-------	---

Description

This function first conducts a meta-filtering test to filter out unpromising SNPs. It then runs a test of omnibus-filtering-based GxE meta-analysis (ofGEM) that combines the strengths of the fixed- and random-effects meta-analysis with meta-filtering. It can also analyze data from multiple ethnic groups. The p-values are calculated using a sequential sampling approach.

Usage

```
ofGEM(Z, X, R, weight, threshold = 0.1, maxSim = 1e+06, tol = 10)
```

Arguments

Z	a matrix of test statistics for gene-environment interactions (GxE) from consortium data. Each row corresponds to a SNP in a set (e.g., a gene), and each column represents a study. For multi-ethnic groups, Z is a list with each element being the matrix for each ethnic group.
X	a matrix of filtering statistics for GxE. Each row corresponds to a SNP in a set, and each column represents a study. For multi-ethnic groups, X is a list with each element being the matrix for each ethnic group.
R	the correlation matrix of test statistics for SNPs in a set. One may use the genotype LD matrix for the set of SNPs to approximate it. This matrix is used when sampling correlated testing and filtering statistics under the null hypothesis and to obtain the null meta-analysis statistics. For multi-ethnic groups, R is a list with each element being the correlation matrix for each ethnic group.
weight	the weight vector for each study, or the weight matrix for each SNP and each study. If the weight is the same across SNPs, it is a vector with length equaling to the number of studies. If the weight is different for different SNPs, it is a matrix with each row corresponding to each SNP and each column representing each study.
threshold	a fixed p-value threshold for filtering test. The default is 0.1.
maxSim	the maximum number of samplings performed in obtaining the sets of correlated testing and filtering statistics under the null. The default is 1e6. This number determines the precision of the p-value calculation.
tol	the tolerance number to stop the sequential sampling procedure. The default is 10. We count the number of sampling-based null meta-statistics that is more extreme than the observed meta-statistics. We sequentially increase the number of sampling with an increment of 100. The sequential sampling will stop if the cumulative count reaches tol. The idea is to stop pursuing a higher precision with more sampling of null if the p-value appears to be not significant. If tol = 0, the number of samplings equals to maxSim.

Value

A list containing

- pval_random_mf the p-value based on the random-effects meta-analysis test with its corresponding meta-filtering.
- pval_fixed_mf the p-value based on the fixed-effects meta-analysis test with its corresponding meta-filtering.
- pval_ofGEM the p-value based on using Fisher's method to aggregating the p-values of fixed- and random-effects meta-analysis tests with meta-filtering
- nsim the number of samplings being performed.

References

Wang, Jiebiao, Qianying Liu, Brandon L. Pierce, Dezheng Huo, Olufunmilayo I. Olopade, Habibul Ahsan, and Lin S. Chen. "A meta-analysis approach with filtering for identifying gene-level gene-environment interactions." *Genetic epidemiology* (2018). <https://doi.org/10.1002/gepi.22115>

Examples

```
data(sim_dat)

pval = ofGEM(Z = sim_dat$Z, X = sim_dat$X, R = sim_dat$R, weight = rep(1/6, 6))
```

<code>sim_dat</code>	<i>A simulated data example</i>
----------------------	---------------------------------

Description

This simulated data list is for demonstration.

Value

A list containing

- Z a simulated matrix of testing statistics for GxE. Each row corresponds to a SNP in a set, and each column represents a study.
- X a simulated matrix of filtering statistics for GxE. Each row corresponds to a SNP, and each column represents a study.
- R a simulated correlation matrix for testing (or filtering) statistics in a set under the null. The testing statistics for SNPs in LD will be also correlated. The correlations can be approximated by the LD matrix. We resample SNP-level testing and filtering statistics from multivariate normal distributions to calculate the null meta-analysis statistics.

sim_dat

5

Examples

```
data(sim_dat)
```

Index

forest_plot, 2

ofGEM, 3

sim_dat, 4