

Package ‘MAPITR’

October 12, 2022

Title MArginal ePIstasis Test for Regions

Version 1.1.2

Date 2020-09-24

Description A genetic analysis tool and variance component model for identifying marginal epistasis between pathways and the rest of the genome. 'MAPITR' uses as input a matrix of genotypes, a vector of phenotypes, and a list of pathways. 'MAPITR' then iteratively tests each pathway for epistasis between any variants within the pathway versus any variants remaining in the rest of the genome. 'MAPITR' returns results in the form of p-values for every pathway indicating whether the null model of there being no epistatic interactions between a pathway and the rest of the genome can be rejected.

License MIT + file LICENSE

URL <https://github.com/mturchin20/MAPITR>

BugReports <https://github.com/mturchin20/MAPITR/issues>

Depends R (>= 3.3.0)

Imports stats, doParallel, Rcpp (>= 1.0.5), CompQuadForm

Suggests testthat, knitr, rmarkdown

Encoding UTF-8

LazyData true

NeedsCompilation yes

RoxygenNote 7.1.1

VignetteBuilder knitr

LinkingTo Rcpp, RcppArmadillo

Author Michael Turchin [aut, cre] (<<https://orcid.org/0000-0003-3569-1529>>),
Gregory Darnell [aut, ctb] (<<https://orcid.org/0000-0003-0425-940X>>),
Lorin Crawford [aut, ctb] (<<https://orcid.org/0000-0003-0178-8242>>),
Sohini Ramachandran [aut] (<<https://orcid.org/0000-0002-9588-7964>>),
Peter Carbonetto [ctb] (<<https://orcid.org/0000-0003-1144-6780>>)

Maintainer Michael Turchin <michael_turchin@brown.edu>

Repository CRAN

Date/Publication 2020-09-28 13:20:12 UTC

R topics documented:

MAPITR	2
MAPITR_SimData_Genotypes	4
MAPITR_SimData_Pathways	4
MAPITR_SimData_PCs	4
MAPITR_SimData_Phenotype	5
MAPITR_TestData_Genotypes	5
MAPITR_TestData_Pathways	6
MAPITR_TestData_PCs	6
MAPITR_TestData_Phenotype	6
Index	7

MAPITR	<i>MArginal ePIstasis Test for Regions (MAPITR)</i>
--------	---

Description

Run MAPITR for a group of pathways on a single phenotype and a set of genome-wide SNPs

Usage

```
MAPITR(
  Genotypes,
  Phenotype,
  Pathways,
  Covariates = NULL,
  CenterStandardize = TRUE,
  OpenMP = FALSE,
  cores = NULL,
  ...
)
```

Arguments

Genotypes	A $n \times p$ matrix containing the genotypes (0/1/2) for all p SNPs across all n individuals. No default value.
Phenotype	A vector containing phenotypic values for all individuals being analyzed. No default value.

Pathways	A $r \times 2$ matrix containing the pathway names and then a comma-separated list of the Genotypes column indices representing each SNP in the associated pathway. Note, this second column of comma-separated indices are the numeric positions for each SNP in Genotypes and not the SNP IDs or column names. No default value.
Covariates	A $n \times q$ matrix containing any q additional covariates that should be included in the M-projection matrix of the model. See Turchin et al. 2020 for details. Note that these are covariates which are applied to both sides of the model, ie the phenotype as well as the genotypes. A y-intercept term is automatically included and does not need to be part of this $n \times q$ matrix. No default value.
CenterStandardize	A logical TRUE/FALSE flag that indicates whether the genotype matrix Genotypes should be centered and standardized before analysis. This is a recommended step. Indicate FALSE if this is a preprocessing step that has already been done prior to running MAPITR. The default value is TRUE.
OpenMP	A logical TRUE/FALSE flag that indicates whether OpenMP versions of MAPITR should be implemented. OpenMP versions of the underlying Rcpp code will run more quickly, but requires the user to be operating an R version that has been installed with OpenMP access. The default value is FALSE.
cores	A numeric value providing the expected number of cores if the OpenMP version of the code is being used. <code>parallel::detectCores()</code> is used by default to assign this variable when no value is given. A value generally should only be given when needing finer control of the code or for testing purposes. The default value is NULL.
...	Additional optional arguments.

Value

A list containing two entries. First, a dataframe (`Results`) containing in the first column the list of pathways that were analyzed, in the second column the associated MAPITR p-values for each pathway, in the third column the associated MAPITR variance component estimates for each pathway, and in the fourth column the associated MAPITR PVEs for each pathway. Second, a matrix (`Eigenvalues`) containing the n associated MAPITR eigenvalues for each pathway per column.

Examples

```
data(MAPITR_TestData_Genotypes, MAPITR_TestData_Phenotype,
MAPITR_TestData_Pathways)
MAPITROutput <- MAPITR(MAPITR_TestData_Genotypes, MAPITR_TestData_Phenotype,
MAPITR_TestData_Pathways, OpenMP=FALSE)
MAPITROutput$Results
```

MAPITR_SimData_Genotypes

MAPITR SimData Genotypes

Description

A simulated dataset of genotypes for use in vignettes

Format

A data frame with 500 rows (individuals) and 750 variables (SNPs)

Source

Manually created

MAPITR_SimData_Pathways

MAPITR SimData Pathways

Description

A simulated set of pathways for use in vignettes

Format

A data frame with 5 rows (pathways) and 2 variables (pathway name, comma-separated list of pathway column indices)

Source

Manually created

MAPITR_SimData_PCs

MAPITR SimData PCs

Description

Top 10 principal components from the simulated genotypes for use in vignettes

Format

A data frame with 500 rows (individuals) and 10 variables (PCs)

Source

Manually created

MAPITR_SimData_Phenotype

MAPITR SimData Phenotype

Description

A simulated set of phenotypes for use in vignettes

Format

A data frame with 500 rows (individuals) and 1 variables (phenotype)

Source

Manually created

MAPITR_TestData_Genotypes

MAPITR TestData Genotypes

Description

A simulated dataset of genotypes for use in unit tests

Format

A data frame with 500 rows (individuals) and 750 variables (SNPs)

Source

Manually created

MAPITR_TestData_Pathways

MAPITR TestData Pathways

Description

A simulated set of pathways for use in unit tests

Format

A data frame with 5 rows (pathways) and 2 variables (pathway name, comma-separated list of pathway column indices)

Source

Manually created

MAPITR_TestData_PCs

MAPITR TestData PCs

Description

Top 10 principal components from the simulated genotypes for use in unit tests

Format

A data frame with 500 rows (individuals) and 10 variables (PCs)

Source

Manually created

MAPITR_TestData_Phenotype

MAPITR TestData Phenotype

Description

A simulated set of phenotypes for use in unit tests

Format

A data frame with 500 rows (individuals) and 1 variables (phenotype)

Source

Manually created

Index

* data

- MAPITR_SimData_Genotypes, 4
- MAPITR_SimData_Pathways, 4
- MAPITR_SimData_PCs, 4
- MAPITR_SimData_Phenotype, 5
- MAPITR_TestData_Genotypes, 5
- MAPITR_TestData_Pathways, 6
- MAPITR_TestData_PCs, 6
- MAPITR_TestData_Phenotype, 6

- MAPITR, 2
- MAPITR_SimData_Genotypes, 4
- MAPITR_SimData_Pathways, 4
- MAPITR_SimData_PCs, 4
- MAPITR_SimData_Phenotype, 5
- MAPITR_TestData_Genotypes, 5
- MAPITR_TestData_Pathways, 6
- MAPITR_TestData_PCs, 6
- MAPITR_TestData_Phenotype, 6