Package 'EpiTest'

May 7, 2026

Title Test for Gene x Gene Interactions in Bi-Parental Populations

Version 1.0.0

Description Provides functions to test for gene x gene interactions in a bi-parental population of inbred lines. The data are fitted with the mixed linear model described in Rio et al. (2022) <doi:10.1101/2022.12.18.520958>, that accounts for gene x gene interactions at both the fixed effect and variance levels. The package also provides graphical tools to display the gene x gene interaction trend at the mean level and the variance component analysis.

License GPL (>= 3)

Depends R (>= 3.5.0)

Imports dplyr, ggplot2, ggrepel, magrittr, MM4LMM, purrr, stringr, utils

Encoding UTF-8

LazyData true

RoxygenNote 7.1.2

NeedsCompilation no

Author Simon Rio [aut, cre] (ORCID: <https://orcid.org/0000-0001-7014-8789>), Tristan Mary-Huard [aut] (ORCID: <https://orcid.org/0000-0002-3839-9067>)

Maintainer Simon Rio <simon.rio@cirad.fr>

Repository CRAN

Date/Publication 2022-12-20 12:30:02 UTC

Contents

`Ancestry.list` *Allele ancestries for the American maize NAM dataset*

Description

`Ancestry.list` is a list of numeric matrices corresponding to three populations of the American maize NAM experiment. Each matrix has 0/1 entries representing the ancestry of alleles (0: homozygous for B73 alleles and 1: homozygous for alternative parent alleles). Each matrix includes 191 to 196 recombinant inbred lines characterized at 1,106 markers.

Usage

```
Ancestry.list
```

Format

A list of three numeric matrices

`EpiTest.fit` *Inference and Test*

Description

This function fits the EpiTest model on a bi-parental population dataset, then outputs the estimates of the fixed and random effects, along with their associated test results.

Usage

```
EpiTest.fit(Ancestry, Pheno, Trait, ParentName, Parents = NULL, Weight = FALSE)
```

Arguments

`Ancestry`	a numeric matrix with 0/1 entries representing the allele ancestries of a bi-parental population at a set of markers (0: homozygous for parent A alleles and 1: homozygous for parent B alleles)
`Pheno`	a data.frame with a `Genotype` character column indicating the names of individuals, a numeric column for each trait to be analyzed, and possibly a `Weight` numeric column indicating the number of observation for each individual. Missing values are discarded if present for a trait

Trait	a string indicating the name of the phenotypic trait
ParentName	a character vector of length 2 providing the parent names where the first element should be named P0 (parent whose allele ancestries are coded 0) and the second element should be named P1 (parent whose allele ancestries are coded 1)
Parents	(optional) a data.frame with a Genotype character column indicating the names of parents, a Family character column indicating the names of families/populations, a numeric column for each trait to be analyzed, and possibly a Weight numeric column indicating the number of observation for each parent (see Details)
Weight	a boolean indicating whether weights indicating the number of observations should be used for the inference. If TRUE, a column Weight should be found in Pheno and Parents (see Details)

Details

Depending on cases, the data may be organized in two different ways.

In Case 1, the Ancestry matrix and the Pheno data.frame respectively provide the whole ancestry and phenotypic data, including the parental data.

In Case 2, the phenotypic data are split into two parts, the parental data being included in a separate Parents data.frame.

In both cases, one may provide a set of weights associated to each individual by including a Weight column in the Pheno and Parents data.frames. These weights correspond to the number of observations that where used to compute the parental BLUPS/mean values. These weights must be provided as a Weight column in the Parents and Pheno data.frame.

The fitted EpiTest model includes three variance components: the segregation variance, the (segregation x segregation) variance and the error variance. For each genetic variance a likelihood ratio test is performed.

The fitted EpiTest model includes three fixed parameters: the intercept, the linear regression coefficient (beta) on parent proportions and the quadratic regression coefficient (delta) on parent proportions that only involves epistatic effects. Each fixed parameter is tested using a Wald test.

Additionally, epistasis can be tested by testing that both the (segregation x segregation) variance component and the quadratic mean component (delta) are null through a likelihood ratio test. Note that in this case the LRT is based on the full (i.e. unrestricted) likelihoods.

The function outputs a list of five items: Beta is the vector of estimated fixed effects, Sigma2 is the vector of estimated variances, Test_full, Test_fixed, Text_random are 3 data.table that provide the results of the tests for no epistasis (jointly tested on fixed and variance components), and for the fixed and random effects, respectively.

Value

A list of four items: Beta, Sigma2, Test_fixed, Test_random

Examples

```
## One bi-parental population, no weighting and no parental phenotypes
data(Pheno.list)
data(Ancestry.list)
```

```
Ancestry <- Ancestry.list[[1]]
Pheno <- Pheno.list[[1]]
ParentName <- c(P0 = 'B73',P1 = 'B97')
ETest.1 <- EpiTest.fit(Ancestry = Ancestry,
                       Pheno = Pheno,
                       ParentName = ParentName,
                       Trait = "GDD_DTA")

## One bi-parental population, with weights and parental phenotypes
data(Parents)
ETest.2 <- EpiTest.fit(Ancestry = Ancestry,
                       Pheno = Pheno,
                       ParentName = ParentName,
                       Trait = "GDD_DTA",
                       Parents = Parents,
                       Weight=TRUE)

## Full NAM analysis, with weights and parental phenotypes
Parent.list <- Parents$Genotype[-1]
names(Parent.list) <- Parents$Family[-1]
ETest.nam <- purrr::imap(Parent.list, ~ EpiTest.fit(Ancestry = Ancestry.list[[.y]],
                                                    Pheno = Pheno.list[[.y]],
                                                    ParentName=c(P0 = 'B73',P1 = .x),
                                                    Parents = Parents,
                                                    Trait = 'GDD_DTA',
                                                    Weight = TRUE))
```

`EpiTest.plot` *Display variance component or directional epistasis graphs*

Description

Display variance component or directional epistasis graphs

Usage

```
EpiTest.plot(
  ETest,
  Title = NULL,
  Family_names = NULL,
  Colors = NULL,
  Type = "mean",
  Alpha = 5/100
)
```

Arguments

`ETest`	a list as obtained from the EpiTest.fit function
`Title`	(optional) a character string to be used as a title for the graph

Family_names	(optional) a character vector with family names
Colors	(optional) a character vector with three colors for the graphical display (for variance graphs only)
Type	a string with "mean" to display the directional epistasis plot or "var" to display the variance component graph
Alpha	type I error for the test, with 5% default value (for directional epistasis plot only)

Value

a ggplot graph

Examples

```
data(Pheno.list)
data(Ancestry.list)

## Full NAM analysis
Parent.list <- Parents$Genotype[-1]
names(Parent.list) <- Parents$Family[-1]
ETest.nam <- purrr::imap(Parent.list, ~ EpiTest.fit(Ancestry = Ancestry.list[[.y]],
                                                    Pheno = Pheno.list[[.y]],
                                                    ParentName = c(P0 = 'B73',P1 = .x),
                                                    Parents = Parents,
                                                    Trait = 'GDD_DTA',
                                                    Weight = TRUE))
## Variance component plot
EpiTest.plot(ETest.nam,Title = 'Days to anthesis',Type = "var")

## Directional epistasis plot
EpiTest.plot(ETest.nam,Title = 'Days to anthesis',Type = "mean",Alpha = 5/100)
```

EpiTest.plot.mean *Display directional epistasis plot*

Description

Display directional epistasis plot

Usage

```
EpiTest.plot.mean(ETest, Title = NULL, Family_names = NULL, Alpha = NULL)
```

Arguments

ETest	a list as obtained from the EpiTest.fit function
Title	(optional) a character string to be used as a title for the graph
Family_names	(optional) a character vector with family names
Alpha	type I error for the test

Value

A ggplot graph

`EpiTest.plot.var` *Display variance component graphs*

Description

Display variance component graphs

Usage

```
EpiTest.plot.var(ETest, Title = NULL, Family_names = NULL, Colors = NULL)
```

Arguments

`ETest`	a list as obtained from the EpiTest.fit function
`Title`	(optional) a character string to be used as a title for the graph
`Family_names`	(optional) a character vector with family names
`Colors`	(optional) a character vector with three colors for the graphical display

Value

a ggplot graph

`Parents` *Phenotypes for parental lines of the American maize NAM dataset*

Description

`Parents` is a data.frame with four rows. The first row corresponds to the parental inbred line common to all bi-parental populations of the American maize NAM experiment (i.e., B73), whereas the three other rows correspond to alternative inbred parental lines used to generate other bi-parental populations. For each parental inbred line, the phenotypic values for five traits are provided, along with two extra columns: Genotype (providing the line ID) and Weight (providing the weight associated with the individual, see function EpiTest.fit for details).

Usage

```
Parents
```

Format

A data.frame

pchisqmix *Chi-Squared Mixtures Distribution Function*

Description

The approximate null distribution of a likelihood ratio for 2 nested mixed models, where both fixed and random effects are tested simultaneously, is a very specific mixture of chi-square distributions. It depends on both the number of random effects and the number of fixed effects to be tested simultaneously. Note that this function is a copy of the `pchisqmix` function of the `TcGSA` package.

Usage

```
pchisqmix(quant, s, q, lower.tail = TRUE)
```

Arguments

`quant`	a quantile
`s`	number of fixed effects to be tested
`q`	number of random effects to be tested
`lower.tail`	logical. if TRUE (default), probabilities are $P[X <= x]$, otherwise $P[X > x]$.

Value

a probability.

Pheno.list *Phenotypes for the American maize NAM dataset*

Description

`Pheno.list` is a list of three data.frames corresponding to three populations of the American maize NAM experiment. Each data.frame contains the phenotypic values of 191 to 196 recombinant inbred lines (RILs) evaluated for five different traits, along with two extra columns: Genotype (providing the RIL ID) and Weight (providing the weight associated with the RIL, see function EpiTest.fit for details).

Usage

```
Pheno.list
```

Format

A list of three data.frames

VarList.comp *Compute covariance matrices*

Description

This function computes the covariance matrices associated with the variance components of the EpiTest model: the segregation variance, the (segregation x segregation) variance and the error variance

Usage

```
VarList.comp(Ancestry)
```

Arguments

Ancestry	a matrix with 0/1 entries representing the allele ancestries of a bi-parental population at a set of markers (0: homozygous for parent A alleles and 1: homozygous for parent B alleles)

Value

A list of three covariance matrices

Examples

```
## One bi-parental population, no weighting
data(Ancestry.list)
Ancestry <- Ancestry.list[[1]]
VarList <- VarList.comp(Ancestry = Ancestry)
purrr::map(VarList,~.x[1:5,1:5])
```

Index