Package 'GWAS.BAYES'

June 6, 2023

Type Package

Title Bayesian analysis of Gaussian GWAS data

Version 1.10.0

Description This package is built to perform GWAS analysis using Bayesian techniques. Currently, GWAS.BAYES has functionality for the implementation of BICOSS for Gaussian phenotypes (Williams, J., Ferreira, M. A., and Ji, T. (2022). BICOSS: Bayesian iterative conditional stochastic search for GWAS. BMC Bioinformatics 23, 475). The research related to this package was supported in part by National Science Foundation awards DMS 1853549, DMS 1853556, and DMS 2054173.

License GPL-3

Encoding UTF-8

LazyData true

biocViews Bayesian, AssayDomain, SNP, GenomeWideAssociation

Imports GA (>= 3.2), caret (>= 6.0-86), memoise (>= 1.1.0), Matrix (>= 1.2-18), limma (>= 3.54.0), stats (>= 4.2.2), MASS (>= 7.3-58.1)

Depends R (>= 4.2.0)

Suggests BiocStyle, knitr, rmarkdown, formatR, rrBLUP

VignetteBuilder knitr

RoxygenNote 7.2.2

git_url https://git.bioconductor.org/packages/GWAS.BAYES

git_branch RELEASE_3_17

git_last_commit 090ccc5

git_last_commit_date 2023-04-25

Date/Publication 2023-06-06

Author Jacob Williams [aut, cre] (<https://orcid.org/0000-0002-6425-1365>), Marco Ferreira [aut] (<https://orcid.org/0000-0002-4705-5661>), Tieming Ji [aut]

Maintainer Jacob Williams <jwilliams@vt.edu>

R topics documented:

BICOSS *BICOSS for Gaussian Phenotypes*

Description

Performs BICOSS analysis as described in Williams, J., Ferreira, M.A.R. & Ji, T. BICOSS: Bayesian iterative conditional stochastic search for GWAS. BMC Bioinformatics 23, 475 (2022). https://doi.org/10.1186/s12859-022-05030-0.

Usage

```
BICOSS(
  Y,
  SNPs,
  FDR_Nominal = 0.05,
  kinship = diag(nrow(SNPs)),
  maxiterations = 400,
  runs_til_stop = 40,
  P3D = TRUE
)
```

Arguments

Y	The observed numeric phenotypes
SNPs	The SNP matrix, where each column represents a single SNP encoded as the numeric coding 0, 1, 2. This is entered as a matrix object.
FDR_Nominal	The nominal false discovery rate for which SNPs are selected from in the screening step.
kinship	The observed kinship matrix, has to be a square positive semidefinite matrix. Defaulted as the identity matrix. The function used to create the kinship matrix used in the BICOSS paper is A.mat() from package rrBLUP.
maxiterations	The maximum iterations the genetic algorithm in the model selection step iterates for. Defaulted at 400 which is the value used in the BICOSS paper simulation studies.
runs_til_stop	The number of iterations at the same best model before the genetic algorithm in the model selection step converges. Defaulted at 40 which is the value used in the BICOSS paper simulation studies.

P3D — Population previous determined, if TRUE BICOSS uses approximated variance parameters estimated from the baseline model when conducting both the screening and the model selection steps. Setting P3D = TRUE is significantly faster. If FALSE, uses exact estimates of the variance parameters all models in both the screening and model selection step.

Value

The column indices of SNPs that were in the best model identified by BICOSS.

Examples

```
library(GWAS.BAYES)
BICOSS(Y = Y, SNPs = SNPs, kinship = kinship,
    FDR_Nominal = 0.05,P3D = TRUE,
    maxiterations = 400,runs_til_stop = 40)
```

---

kinship — *A. Thaliana Kinship matrix*

Description

This is a kinship matrix from the TAIR9 genotype information for 328 A. Thaliana Ecotypes from the paper Components of Root Architecture Remodeling in Response to Salt Stress. The kinship matrix was computed using all SNPs with minor allele frequency greater than 0.01.

Usage

```
kinship
```

Format

'kinship' A matrix with 328 rows and 328 columns corresponding to the 328 ecotypes.

SMA — *Performs Single Marker Association tests for both Linear Mixed Models and Linear models.*

Description

Performs Single Marker Association tests for both Linear Mixed Models and Linear models.

Usage

```
SMA(Y, SNPs, kinship = FALSE, P3D = FALSE)
```

Arguments

Y	The observed numeric phenotypes
SNPs	The SNP matrix, where each column represents a single SNP encoded as the numeric coding 0, 1, 2. This is entered as a matrix object.
kinship	The observed kinship matrix, has to be a square positive semidefinite matrix. Defaulted as the identity matrix. The function used to create the kinship matrix used in the BICOSS paper is A.mat() from package rrBLUP.
P3D	Population previous determined, if TRUE BICOSS uses approximated variance parameters estimated from the baseline model when conducting both the screening and the model selection steps. Setting P3D = TRUE is significantly faster. If FALSE, uses exact estimates of the variance parameters all models in both the screening and model selection step.

Value

The p-values corresponding to every column provided in SNPs. These p-values can be used with any threshold of your choosing or with p.adjust().

SNPs *A. Thaliana Genotype matrix*

Description

This is a matrix with 328 observations and 9,000 SNPs. Each row is contains 9,000 SNPs from a single A. Thaliana ecotype in the paper Components of Root Architecture Remodeling in Response to Salt Stress.

Usage

```
SNPs
```

Format

'SNPs' A matrix with 328 observations and 9,000 SNPs.

Y *A. Thaliana Simulated Phenotype matrix*

Description

This is a phenotype matrix simulated from the 9,000 SNPs. SNPs at positions 450, 1350, 2250, 3150, 4050, 4950, 5850, 6750, 7650, and 8550 have nonzero coefficients. Further, the data was simulated under the linear mixed model specified in the vignette and the BICOSS manuscript using the kinship matrix (kinship).

Usage

```
Y
```

Format

'Y' A matrix with 328 rows corresponding to the 328 ecotypes.

Index