Package 'BG2'

November 11, 2026

Type Package

Title Performs Bayesian GWAS analysis for non-Gaussian data using BG2

Version 1.12.0

Description This package is built to perform GWAS analysis for non-Gaussian data using BG2. The BG2 method uses penalized quasi-likelihood along with nonlocal priors in a two step manner to identify SNPs in GWAS analysis. The research related to this package was supported in part by National Science Foundation awards DMS 1853549 and DMS 2054173.

License GPL-3 + file LICENSE

Encoding UTF-8

LazyData false

biocViews Bayesian, AssayDomain, SNP, GenomeWideAssociation

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

Depends R (>= 4.2.0)

Suggests BiocStyle, knitr, rmarkdown, formatR, rrBLUP, testthat (>= 3.0.0)

RoxygenNote 7.2.2

VignetteBuilder knitr

Config/testthat/edition 3

git_url https://git.bioconductor.org/packages/BG2

git_branch RELEASE_3_23

git_last_commit e45b59b

git_last_commit_date 2026-04-28

Repository Bioconductor 3.23

Date/Publication 2026-11-11

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

Maintainer Jacob Williams <jwilliams@vt.edu>

Contents

BG2 2
kinship 3
SNPs 4
Y_binary 4
Y_poisson 5

Index **6**

BG2 *Performs BG2 as described in the manuscript, Xu, Williams, and Ferreira BG2: Bayesian variable selection in generalized linear mixed models with non-local priors for non-Gaussian GWAS data, Bioinformatics, Submitted.*

Description

Performs BG2 as described in the manuscript ...

Usage

```
BG2(
  Y,
  SNPs,
  FDR_Nominal = 0.05,
  Fixed = NULL,
  family = c("poisson", "bernoulli"),
  Covariance,
  Z = NULL,
  replicates = NULL,
  Tau = "uniform",
  maxiterations = 4000,
  runs_til_stop = 400
)
```

Arguments

Y	The observed phenotypes, count or binary.
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.
Fixed	A matrix of fixed covariates to control for. Do not include the intercept. The value is defaulted at NULL implying no fixed covariates.
family	Specify if the response is count ("poisson") or binary ("bernoulli").

Covariance	A list of covariance matrices that are the covariance matrices of the random effects. This matches the list of design matrices in Z.
Z	A list of matrices specifying the design matrix of each random effect of interest.
replicates	If family = "poisson", the replicates of each ecotype, can be a vector or a number if the number of replicates is the same for each ecotype. If family = "binomial", replicates = NULL.
Tau	Specifying either a fixed value for the dispersion parameter of the nonlocal prior (0.022 and 0.348 are used in the paper). Or specify a prior for tau, either uniform or Inverse Gamma centered at ...
maxiterations	The maximum iterations the genetic algorithm in the model selection step iterates for. Defaulted at 400 which is the value used in the BG2 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 BG2 paper simulation studies.

Value

The column indices of SNPs that were in the best model identified by BG2.

Examples

```
library(BG2)
data("Y_poisson");data("SNPs");data("kinship")
n <- length(Y_poisson)
covariance <- list()
covariance[[1]] <- kinship
covariance[[2]] <- diag(1, nrow = n, ncol = n)

set.seed(1330)
output_poisson <- BG2(Y=Y_poisson, SNPs=SNPs, Fixed = NULL,
                  Covariance=covariance, Z=NULL, family="poisson",
                  replicates=4, Tau="uniform",FDR_Nominal = 0.05,
                  maxiterations = 4000, runs_til_stop = 400)
```

---

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 (Julkowska et al. Genetic Components of Root Architecture Remodeling in Response to Salt Stress, The Plant Cell, Volume 29, Issue 12, December 2017, Pages 3198–3213). The kinship matrix was computed using all SNPs with minor allele frequency greater than 0.01.

Usage

```
data(kinship)
```

Format

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

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 (Julkowska et al. Genetic Components of Root Architecture Remodeling in Response to Salt Stress, The Plant Cell, Volume 29, Issue 12, December 2017, Pages 3198–3213).

Usage

```
data(SNPs)
```

Format

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

Y_binary *A. Thaliana Simulated Phenotype matrix*

Description

This is a phenotype matrix simulated from 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 a binary mixed effects model with only a kinship random effect. The data was simulated using the kinship random effect provided in the package.

Usage

```
data(Y_binary)
```

Format

'Y_binary' A vector with 328 observations corresponding to the 328 ecotypes.

Y_poisson *A. Thaliana Simulated Phenotype matrix*

Description

This is a phenotype matrix simulated from 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 a poisson mixed effects model with both a kinship random effect and an overdispersion random effect. The data was simulated using the kinship random effect provided in the package.

Usage

```
data(Y_poisson)
```

Format

'Y_poisson' A vector with 328 observations corresponding to the 328 ecotypes.

Index

∗ **datasets**
kinship, 3
SNPs, 4
Y_binary, 4
Y_poisson, 5

BG2, 2

kinship, 3

SNPs, 4

Y_binary, 4
Y_poisson, 5