Package ‘emBALVI’

April 16, 2026

Title EM Bayesian Adaptive LASSO Variational Inference Based GWAS

Version 0.1.0

Description Performs Genome-Wide Association Study (GWAS) analysis using Expectation-Maximization Bayesian Adaptive LASSO with Variational Inference (emBALVI). Includes genotype preprocessing, genomic relationship matrix construction, GWAS analysis, Manhattan and QQ plotting.s.

Depends R (>= 4.0)

Imports ggplot2, RColorBrewer

Suggests rmarkdown, testthat (>= 3.0.0), roxygen2

License MIT + file LICENSE

Encoding UTF-8

LazyData true

RoxygenNote 7.3.3

NeedsCompilation no

Author Prakash Kumar [aut, cre],
Himadri Sekhar Roy [aut],
Ranjit Kumar Paul [aut],
Md. Yeasin [aut],
Neeraj Budhlakoti [aut],
Sunil Kumar Yadav [aut],
Amrit Kumar Paul [aut]

Maintainer Prakash Kumar <`prakash289111@gmail.com`>

Repository CRAN

Date/Publication 2026-04-16 08:30:08 UTC

Contents

`buildGRM` *Build VanRaden Genomic Relationship Matrix*

Description

Build VanRaden Genomic Relationship Matrix

Usage

```
buildGRM(X)
```

Arguments

`X`	Numeric genotype matrix

Value

GRM matrix

`convert_to_dosage` *Convert HMP Genotype to Dosage*

Description

Convert HMP Genotype to Dosage

Usage

```
convert_to_dosage(snp_matrix)
```

Arguments

`snp_matrix`	SNP matrix

Value

Numeric dosage matrix

`emBALVI` *EM-BALVI GWAS*

Description

Performs genome-wide association using expectation-maximization Bayesian adaptive LASSO with variational inference.

Usage

```
emBALVI(Y, X, B, G, max_iter = 20, tol = 1e-06)
```

Arguments

`Y`	Phenotype vector
`X`	Genotype matrix (numeric dosage)
`B`	Fixed effect design matrix
`G`	Genomic relationship matrix
`max_iter`	Maximum iterations
`tol`	Convergence tolerance

Value

Data frame of marker effects and p-values

`phenotypes_potatoyield` *Potato Yield Phenotype Dataset*

Description

Phenotypic data used for GWAS analysis in emBALVI.

Usage

```
phenotypes_potatoyield
```

Format

A data frame with genotype IDs and yield phenotype.

Source

Real potato yield experiment dataset

`plot_manhattan` *Manhattan Plot for emBALVI*

Description

Manhattan Plot for emBALVI

Usage

```
plot_manhattan(results)
```

Arguments

`results`	GWAS result data frame

Value

A ggplot object

`plot_qq` *QQ Plot for emBALVI GWAS Results*

Description

Generates a Quantile-Quantile (QQ) plot for GWAS p-values.

Usage

```
plot_qq(results)
```

Arguments

`results`	GWAS result data frame (must contain column 'p')

Value

A ggplot object

`snp_NN_10_hmp` *SNP Genotype Dataset (HMP Format)*

Description

Raw SNP genotype matrix in HapMap-like format.

Usage

```
snp_NN_10_hmp
```

Format

A data frame containing SNP markers and genotype calls.

Source

Real SNP dataset used for EM-BALVI GWAS

Index