

# Package ‘GAGBLUP’

May 7, 2026

**Type** Package

**Title** Genetic Algorithm Assisted Genomic Best Liner Unbiased Prediction

**Version** 1.0

**Description** Performs genetic algorithm (Scrucca, L (2013) <[doi:10.18637/jss.v053.i04](https://doi.org/10.18637/jss.v053.i04)>) assisted genomic best liner unbiased prediction for genomic selection. It also provides a binning method in natural population for genomic selection under the principle of linkage disequilibrium for dimensional reduction.

**License** GPL-3

**Encoding** UTF-8

**LazyData** true

**RoxygenNote** 7.2.3

**Depends** R (>= 4.1.0)

**Imports** GA, foreach, doParallel

**NeedsCompilation** no

**Repository** CRAN

**Date/Publication** 2023-09-05 16:50:08 UTC

**Maintainer** Yuxiang Zhang <yuxiangzhang\_99@foxmail.com>

**Author** Yuxiang Zhang [aut, cre],  
Yang Xu [aut],  
Yanru Cui [ctb],  
Chenwu Xu [ctb],  
Shizhong Xu [ctb]

## Contents

bin . . . . .	2
bins . . . . .	2
gagblup . . . . .	3
genotype . . . . .	4
phenotype . . . . .	4
<b>Index</b>	<b>5</b>

---

bin	<i>Binning Genotypes for Dimensional Reduction</i>
-----	--

---

**Description**

Binning the original genotypes into bins for dimensional reduction under the principle of linkage disequilibrium.

**Usage**

```
bin(genotype = genotype, binvar = 0.15)
```

**Arguments**

genotype	a matrix for genotypes in numeric format, coded as 1, 0 and -1, with individuals in rows and markers in cols.
binvar	a hyper-parameter between 0 and 1, the closer to 0, the fewer bins yields. Users can choose binvar based on the required number of bins, default is 0.15.

**Value**

A list with following information is returned: \$bins\_genotypes binned genotypes \$bins\_range start and stop of each bin

**Examples**

```
## load example data from GAGBLUP package
data(genotype)
## binning genotypes
bins <- bin(genotype,0.2)
```

---

bins	<i>binned genotypes</i>
------	-------------------------

---

**Description**

Binned genotypes of 410 maize hybrids.

**Usage**

```
bins
```

**Format**

A data frame with 410 rows and 810 variables.

---

gagblup	<i>Genetic algorithm assisted genomic best liner unbiased prediction for genomic selection</i>
---------	--

---

## Description

Performs genomic selection with genetic algorithm assisted genomic best liner unbiased prediction

## Usage

```
gagblup(
  genotype,
  phenotype,
  fit_fun = "HAT",
  maxiter = 2000,
  nfold = 10,
  nTimes = 1,
  seed = 123,
  n_core = 1
)
```

## Arguments

genotype	a matrix for genotypes in numeric format, with individuals in rows and markers in cols.
phenotype	a vector of phenotype, missing (NA) values are not allowed.
fit_fun	the fitness function. There are four options: fitness = "AIC"/"BIC"/"FIT"/"HAT", default is "HAT"
maxiter	max number of iterations for GAGBLUP, default is 2000
nfold	the number of folds. Default is 10.
nTimes	the number of independent replicates for the cross-validation. Default is 1.
seed	the random number. Default is 123.
n_core	the number of CPU to be used, default is 1.

## Value

A list with following information is returned: \$R2 the squared pearson correlation coefficient between the true value and the predicted value, \$predicted\_value the predicted value and the true value of the phenotype, \$marker\_selection a vector of the selected markers, with 1 indicates selected, 0 indicates not selected.

**Examples**

```
## load example data from GAGBLUP package
data(phenotype)
data(bins)
phenotype <- phenotype[1:200,3]
result <- gagblup(bins[1:200,],phenotype,fit_fun='HAT',maxiter=1,nfold=2,nTimes=1,seed=123,n_core=1)
```

---

genotype	<i>Genotype in Numeric Format</i>
----------	-----------------------------------

---

**Description**

Genotypic data of 410 maize hybrids with 4906 SNPs.

**Usage**

genotype

**Format**

A data frame with 410 rows and 4906 variables.

---

phenotype	<i>Phenotypic data of 410 hybrids</i>
-----------	---------------------------------------

---

**Description**

This dataset contains phenotypic data of 410 hybrids for grain yield in maize.

**Usage**

phenotype

**Format**

A data frame with 410 rows and 3 variables:

M The names of male parents.

F The names of female parents.

GY The grain yield of hybrids.

# Index

## \* datasets

bins, 2

genotype, 4

phenotype, 4

bin, 2

bins, 2

gagblup, 3

genotype, 4

phenotype, 4