Package 'gpcp'

November 6, 2024

Type Package

Title Genomic Prediction of Cross Performance

Version 0.1.0

Maintainer Christine Nyaga <cmn92@cornell.edu>

Description This function performs genomic prediction of cross performance using genotype and phenotype data.
It processes data in several steps including loading necessary software, converting genotype data, processing phenotype data,
fitting mixed models, and predicting cross performance based on weighted marker effects.
For more information, see Labroo et al. (2023) <doi:10.1007/s00122-023-04377-z>.

License GPL (>= 3)

Encoding UTF-8

LazyData true

LinkingTo Rcpp, RcppArmadillo

Imports BiocManager, Rcpp, dplyr, sommer, AGHmatrix, snpStats,
VariantAnnotation, tools, magrittr, methods

RoxygenNote 7.3.2

Suggests knitr, rmarkdown, testthat (>= 3.0.0)

Config/testthat/edition 3

VignetteBuilder knitr

Depends R (>= 2.10)

NeedsCompilation yes

Author Marlee Labroo [aut],
Christine Nyaga [cre, aut],
Lukas Mueller [aut]

Repository CRAN

Date/Publication 2024-11-06 15:50:02 UTC

Contents

---

phenotypeFile   *Example Phenotype Data*

Description

This is a sample phenotype dataset used for genomic prediction.

Usage

```
phenotypeFile
```

Format

A data frame with 24 columns:

ATW Description of ATW

AUDPC_YAD Area Under Disease Progress Curve for YAD

AUDPC_YMV Area Under Disease Progress Curve for YMV

Accession Genotype IDs for each individual

Block Block information

DMC Dry Matter Content values

Design Experimental design

LOC Location of the trials

NPH Number of Plants Harvested

OXBI Oxidation Index

Oxint180Minutes Oxidation intensity after 180 minutes

PLOT Plot number

REP Replication number

Settweight Weight of the planting setts

TTNPL Total Tuber Number per Plant

TTWPL Total Tuber Weight per Plant

Trial Trial name or ID

Vigor Plant vigor score

YIELD Yield values

Year Year of the experiment

Yield.per.plot..kg. Yield per plot in kilograms

Yield_udj Unadjusted Yield

rAUDPC_YAD Relative AUDPC for YAD

rAUDPC_YMV Relative AUDPC for YMV

Source

Generated for the gpcp package example

Examples

```
data(phenotypeFile)
head(phenotypeFile)
```

| runGPCP | *Genomic Prediction of Cross Performance This function performs genomic prediction of cross performance using genotype and phenotype data.* |

Description

Genomic Prediction of Cross Performance This function performs genomic prediction of cross performance using genotype and phenotype data.

Usage

```
runGPCP(
  phenotypeFile,
  genotypeFile,
  genotypes,
  traits,
  weights = NA,
  userSexes = "",
  userFixed = NA,
  userRandom = NA,
  Ploidy = NA,
  NCrosses = NA
)
```

Arguments

phenotypeFile	A data frame containing phenotypic data, typically read from a CSV file.
genotypeFile	Path to the genotypic data, either in VCF or HapMap format.
genotypes	A character string representing the column name in the phenotype file for the genotype IDs.
traits	A string of comma-separated trait names from the phenotype file.
weights	A numeric vector specifying weights for the traits.
userSexes	A string representing the column name corresponding to the individuals' sexes.
userFixed	A string of comma-separated fixed effect variables.
userRandom	A string of comma-separated random effect variables.
Ploidy	An integer representing the ploidy level of the organism.
NCrosses	An integer specifying the number of top crosses to output.

Value

A data frame containing predicted cross performance.

Examples

```
# Load phenotype data from CSV
phenotypeFile <- read.csv(system.file("extdata", "phenotypeFile.csv", package = "gpcp"))
genotypeFile <- system.file("extdata", "genotypeFile_Chr9and11.vcf", package = "gpcp")
finalcrosses <- runGPCP(
    phenotypeFile = phenotypeFile,
    genotypeFile = genotypeFile,
    genotypes = "Accession",
    traits = "YIELD,DMC",
    weights = c(3, 1),
    userFixed = "LOC,REP",
    Ploidy = 2,
    NCrosses = 150
)
print(finalcrosses)
```

Index