Package: gpcp (via r-universe)

November 7, 2024

Type Package

Title Genomic Prediction of Cross Performance

Version 0.1.0

Author Marlee Labroo, Christine Nyaga, Lukas Mueller

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.

License MIT + file LICENSE

Encoding UTF-8

LazyData true

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

RoxygenNote 7.3.2

Config/pak/sysreqs make libpng-dev libxml2-dev libssl-dev

Repository https://cmn92.r-universe.dev

RemoteUrl https://github.com/cmn92/gpcp

RemoteRef HEAD

RemoteSha 1d834e19d37a696876231382b97adf2e1fdd000d

Contents

---

runGPCP	*Genomic Prediction of Cross Performance*

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.

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	A file path to the genotypic data, either in VCF format or as a HapMap.
genotypes	A character string representing the column name in the phenotype file that corresponds to the genotype IDs.
traits	A string of comma-separated trait names from the phenotype file, which will be used for genomic prediction.
weights	A numeric vector specifying the weights for the traits. The order of weights should correspond to the order of traits.
userSexes	Optional. A string representing the column name in the phenotype file corresponding to the individuals' sexes.
userFixed	A string of comma-separated fixed effect variables from the phenotype file. If no fixed effects are required, set to NA.
userRandom	A string of comma-separated random effect variables from the phenotype file. If no random effects are required, set to NA.
Ploidy	An integer representing the ploidy level of the organism (e.g., 2, 4, 6).
NCrosses	An integer specifying the number of top crosses to output. Maximum is a full diallel.

Details

This function is designed for genomic prediction of cross performance and can handle both diploid and polyploid species. It processes genotype data, calculates genetic relationships, and fits mixed models using the 'sommer' package. It outputs the best predicted crosses based on user-defined traits and weights.

Value

A data frame containing predicted crosses with the following columns:

Parent1	First parent genotype ID.
Parent2	Second parent genotype ID.
CrossPredictedMerit	
	Predicted merit of the cross.
P1Sex	Optional. Sex of the first parent if userSexes is provided.
P2Sex	Optional. Sex of the second parent if userSexes is provided.

Note

This function relies on the 'sommer', 'dplyr', and 'AGHmatrix' packages for processing mixed models and genomic data.

Author(s)

Marlee Labroo, Christine Nyaga, Lukas Mueller

References

Xiang, J., et al. (2016). "Mixed Model Methods for Genomic Prediction." *Nature Genetics*. Batista, L., et al. (2021). "Genetic Prediction and Relationship Matrices." *Theoretical and Applied Genetics*.

See Also

sommer, dplyr, Gmatrix

Examples

```
# Load phenotype data from CSV
phenotypeFile <- read.csv("~/Documents/GCPC_input_files/2020_TDr_PHENO (1).csv")

# Genotype file path
genotypeFile <- "~/Documents/GCPC_input_files/genotypeFile.vcf"


# Define inputs
genotypes <- "Accession"
traits <- c("rAUDPC_YMV", "YIELD", "DMC")
weights <- c(0.2, 3, 1)
userFixed <- c("LOC", "REP")
Ploidy <- 2
NCrosses <- 150

# Run genomic prediction of cross performance
finalcrosses <- runGPCP(
    phenotypeFile = phenotypeFile,
    genotypeFile = genotypeFile,
    genotypes = genotypes,
```

```
    traits = paste(traits, collapse = ","),
    weights = weights,
    userFixed = paste(userFixed, collapse = ","),
    Ploidy = Ploidy,
    NCrosses = NCrosses
)

# View the predicted crosses
print(finalcrosses)
```

Index