Package ‘synbreedData’

April 29, 2015

Type    Package
Title   Data for the Synbreed Package
Version 1.5
Date 2015-04-29
Author Valentin Wimmer, Theresa Albrecht, Hans-Juergen Auinger, Chris-Carolin Schoen with contributions by Malena Erbe, Ulrike Ober and Christian Reimer
Depends R (>= 2.10)
Maintainer Valentin Wimmer <Valentin.Wimmer@kws.com>
Description Data sets for the 'synbreed' package with three data sets from cattle, maize and mice to illustrate the functions in the 'synbreed' R package. All data sets are stored in the gpData format introduced in the 'synbreed' package. This research was funded by the German Federal Ministry of Education and Research (BMBF) within the AgroClustEr Synbreed - Synergistic plant and animal breeding (FKZ 0315528A).
URL http://synbreed.r-forge.r-project.org/
License GPL-2
LazyLoad yes
LazyData no
ZipData no
NeedsCompilation no
Repository CRAN
Date/Publication 2015-04-29 12:21:46

R topics documented:

cattle ................................................................. 2
maize ............................................................... 2
mice ................................................................. 3

Index 5
**cattle**  
_Dairy cattle data_

**Description**

Data set contains genotypic, phenotypic, map and pedigree data of 500 bulls. All individuals are labeled with an unique ID, starting with ID1430 and ending with ID1929. Genotypic and pedigree data is based on a real cattle data set while phenotypes were built artificially. Pedigree information is available at least on parents and grandparents of the phenotyped individuals.

There are two quantitative phenotypes available. The heritabilities of these traits are 0.41 and 0.66, estimated with a pedigree-based animal model using the data set on hand.

Genotypic data consists of 7250 biallelic SNP markers for every phenotyped individual with missing data included. SNPs are mapped across all 29 autosomes. Distances in the SNP map are given in mega bases (Mb).

**Usage**

```r
data(cattle)
```

**Format**

Object of class gpData

**Examples**

```r
## Not run:
library(synbreed)
data(cattle)
summary(cattle)

## End(Not run)
```

---

**maize**  
_Simulated maize data_

**Description**

This is a simulated dataset of a maize breeding program. Data comprises 1250 doubled haploid (DH) lines that were genotyped with 1117 polymorphic SNP markers and phenotyped in a testcross with a single tester for one quantitative trait. All individuals are labeled with a unique ID, starting from 11360 to 12609. Markers are distributed along all 10 chromosomes of maize. Pedigree information starts with basis population and is available up to 15 generations. The 1250 lines belong to 25 full sib families with 50 individuals in each family. In the simulation of true breeding
values (TBV), 1000 biallelic quantitative trait loci (QTL) with equal and additive (no dominance or epistasis) effects were generated. True breeding values for individuals were calculated according to

\[ t_{bv} = \sum_{k=1}^{1000} QTL_k \]

where \( QTL_k \) is the effect of the \( k \)-th QTL. Phenotypic values were simulated according to

\[ y_i = t_{bv_i} + \epsilon_i \]

where \( \epsilon_i \sim N(0, \sigma^2) \). The value for \( \sigma^2 \) was chosen in a way that a given plot heritability of \( h^2 = 0.197 \) is realized. Note that true breeding values for 1250 phenotyped lines are stored as \( t_{bv} \) in covar of gpData object. Reported phenotypic values of lines are adjusted values testcross means for yield [dt/ha] evaluated in 3 locations.

**Usage**

```r
data(maize)
```

**Format**

Object of class gpData

**Examples**

```r
## Not run:
library(synbreed)
data(maize)
summary(maize)

## End(Not run)
```

---

**mice**

*Heterogenous stock mice population*

---

**Description**

Data set comprises public available data of 2527 (1293 males and 1234 females) heterogenous stock mice derived from eight inbred strains (A/J, AKR/J, BALBc/J, CBA/J, C3H/HeJ, C57BL/6J, DBA/2J and LP/J) followed by 50 generations of pseudorandom mating. All individuals are labeled with a unique ID, starting with A048005080. For all individuals, family, sex (females=0, males=1), month of birth (1-12), birthyear, coat color, cage density and litter is available and stored in covar.

The measured traits are described in Solberg et al. (2006). Here, the body weight at age of 6 weeks [g] and growth slope between 6 and 10 weeks age [g/day] are available. The heritabilities of these traits are reported as 0.74 and 0.30, respectively (Valdar et al, 2006b). Phenotypic data was taken from [http://mus.well.ox.ac.uk/GSCAN/HS_PHENOTYPES/Weight.txt](http://mus.well.ox.ac.uk/GSCAN/HS_PHENOTYPES/Weight.txt).

Genotypic data consists of 12545 biallelic SNP markers and is available for 1940 individuals. Raw genotypic data from [http://mus.well.ox.ac.uk/GSCAN/HS_GENOTYPES/](http://mus.well.ox.ac.uk/GSCAN/HS_GENOTYPES/) is given in the
Ped-File Format with two columns for each marker. Both alleles were combined to a single genotype for each marker in mice data. The SNPs are mapped in a sex-averaged genetic map with distances given in centimorgan (Shifman et al. (2006)). SNPs are mapped across all 19 autosomes and X-chromosome where distances between adjacent markers vary from 0 to 3 cM.

Usage

data(mice)

Format

Object of class gpData

Source

Welcome Trust Centre for Human Genetics, Oxford University, data available from http://gsan. well.ox.ac.uk

References


Examples

```r
## Not run:
library(synbreed)
data(mice)
summary(mice)

## End(Not run)
```
Index

*Topic datasets
  cattle, 2
  maize, 2
  mice, 3

  cattle, 2
  maize, 2
  mice, 3