

Package ‘pedantics’ documentation

of

December 8, 2009

Type Package

Title Functions to facilitate power and sensitivity analyses for genetic studies of natural populations

Version 1.01

Date 2009-07-16

Depends R (>= 2.4.0), MCMCglmm, kinship, grid, MasterBayes

Author Michael Morrissey

Maintainer Michael Morrissey <michael.morrissey@ed.ac.uk>

Description More to come

License GPL-2 | GPL-3

LazyLoad yes

R topics documented:

drawPedigree	1
fixPedigree	4
fpederr	5
genomesim	7
gryphons	9
makePedigreeFactor	10
makePedigreeNumeric	11
microsim	13
pedantics-package	15
pedigreeStats	16
pedStatSummary	19
phensim	19
rpederrBird	21
rpederr	24

drawPedigree	<i>Produce a graphical representation of a pedigree</i>
---------------------	---

Description

Plots a pedigree, with options specific to considerations for pedigrees used for quantitative genetic inference in natural populations. Pedigrees containing only those individuals that are informative with respect to (genetic) variation in an arbitrary trait can be plotted, potentially overlain on a complete pedigree. Functions also exist to plot various types of pedigree links associated with focal individuals.

Usage

```
drawPedigree(Ped, cohorts = NULL, sex = NULL, dat = NULL, dots = "n",
             plotfull = "y", writeCohortLabels = "n", links = "all",
             sexInd = c(0, 1), dotSize = 0.001, dataDots = "n",
             dataDots.cex = 2, cohortLabs.cex = 1, retain="informative",
             focal=NULL, sexColours=c('red','blue'), ...)
```

Arguments

Ped	An ordered pedigree with 3 columns: id, dam, sire
cohorts	An optional numeric vector of the same length as the pedigree designating, for example cohort affinities or birth years
sex	An optional numeric vector of the same length as the pedigree containing the sexes (may be unknown) of all individuals with entries in the pedigree. Defaults (modifiable with sexInd) are 0=male and 1=female
dat	An optional vector or data frame containing indicators of data availability. If dat contains only ones and zeros, then any individual with any entry of one will be considered as having data records. If dat contains values other than ones and zeros, individuals in the pedigree with rows in dat that contain at least one available record, i.e., one data record is not NA, will be treated as having data.
dots	If 'y', then a dot will be printed representing each individual in the pedigree. If sexes are available, dots will be colour coded by sex.
plotfull	To be used when dat is supplied. If 'y' (the default), individuals in the pedigree that are uninformative with respect to the available data have their pedigree links plotted in gray.
writeCohortLabels	To be used when cohorts is used. Will plot the cohort values on the left hand side of the pedigree image.
links	Default is 'all', other values are 'mums' to print only maternal pedigree links and 'dads' to print only paternal pedigree links.
sexInd	To be used with if sex is supplied and if the vector of sex specifiers differs from the default.

<code>dotSize</code>	Set the dot size bigger or smaller
<code>dataDots</code>	Will print dots over the dots denoting individuals, but denoting individuals with available data as indicated by <code>dat</code> .
<code>dataDots.cex</code>	controls the size of <code>dataDots</code> relative to <code>dots</code> .
<code>cohortLabs.cex</code>	controls the size of cohort labels.
<code>retain</code>	When those pedigree links only informative relative to phenotypic data availability are to be plotted, this controls whether or not a pruned pedigree based on phenotypic data is plotted (if set to "pruned"), or whether strictly only those informative pedigree links are plotted (the default)
<code>focal</code>	An optional list containing the id of an individual and the kinds of relatives of the focal individual to which to plot pedigree links. Available types are 'offspring', 'descendants', 'parents', 'ancestors', and 'kin'.
<code>sexColours</code>	The colours that will be used to draw points and or lines associated with males and females.
<code>...</code>	Additional graphical parameters.

Details

If necessary, more details than the description above

Author(s)

Michael Morrissey (michael.morrissey@ed.ac.uk)

References

Morrissey, M.B, and A.J. Wilson, 2009. *pedantics*, an R package for pedigree-based genetic simulation, and pedigree manipulation, characterisation, and viewing. *Molecular Ecology Resources*.

See Also

`fixPedigree` to prepare pedigrees that may not explicitly contain records for all individuals (i.e., where founding individuals may only appear in the dam or sire column.)

Examples

```
data(gryphons)
pedigree<-fixPedigree(gryphons[,1:3])

## draw the gryphon pedigree by pedigree depth
drawPedigree(pedigree)

## draw the gryphon pedigree by cohort
drawPedigree(pedigree,cohorts=gryphons$cohort,writeCohortLabels='y',
             cohortLabs.cex=1)

## draw the gryphon pedigree by cohort with only maternal links
```

```
drawPedigree(pedigree,cohorts=gryphons$cohort,links='mums')

## draw the gryphon pedigree by cohort with colour only for those
## individuals that are informative relative to the quantitative
## genetics of a hypothetical trait only measured for individuals
## in the last two cohorts, emphasize the phenotyped individuals
## with large black dots, and all other individuals with dots
## colour coded by sex:
dataAvailability<-(gryphons$cohort>=(max(gryphons$cohort)-1))+0
drawPedigree(pedigree,cohorts=gryphons$cohort,sex=gryphons$sex,
  dots='y',dat=dataAvailability,writeCohortLabels='y',dataDots='y')
```

fixPedigree	<i>Manipulating pedigrees to prepare them for requirements of subsequent analyses</i>
-------------	---

Description

Prepares a pedigree to conform with requirements of many softwares used in quantitative genetic analysis, as well as for many of the functions in pedantics.

Usage

```
fixPedigree(Ped, dat = NULL)
```

Arguments

Ped	An ordered pedigree with 3 columns: id, dam, sire
dat	An optional data frame, the same length as the pedigree

Value

Returns a pedigree in which all individuals that exist in the dam and sire columns are represented by their own record lines, occurring before the records of their first offspring. If data are supplied, then fixPedigree will return a dataframe, the first two columns are the 'fixed' pedigree, and the following columns of which contain appropriately reordered data.

Author(s)

Michael Morrissey (michael.morrissey@ed.ac.uk)

References

Morrissey, M.B, and A.J. Wilson, 2009. pedantics, an R package for pedigree-based genetic simulation, and pedigree manipulation, characterisation, and viewing. Molecular Ecology Resources.

Examples

```
## a valid pedigree, i.e., no loops, no bisexuality, etc.,
## but where not all parents have a record line, and where
## parents do not necessarily occur before their offspring:
pedigree<-as.data.frame(matrix(c(
  10,1,2,
  11,1,2,
  12,1,3,
  13,1,3,
  14,4,5,
  15,6,7,
  4,NA,NA,
  5,NA,NA,
  6,NA,NA,
  7,NA,NA),10,3,byrow=TRUE))
names(pedigree)<-c("id","dam","sire")
pedigree
fixedPedigree<-fixPedigree(Ped=pedigree)
fixedPedigree
```

fpederr	<i>Simulates a pedigree with errors and missing data from a complete pedigree.</i>
---------	--

Description

Implements the 'forward' approach to producing pairs of pedigrees for power and sensitivity analyses.

Usage

```
fpederr(truePedigree, founders = NULL, sex = NULL, samp = NULL,
        sireE = NULL, damE = NULL, sireA = NULL, damA = NULL,
        cohort = NULL, first = NULL, last = NULL, monocey = 0,
        modifyAssumedPedigree = 0)
```

Arguments

truePedigree	A complete pedigree with records for all individuals and parental ID's for all non-founders
founders	A vector the same length as the pedigree containing indicator variables 1 = founder, 0 = non-founder
sex	A vector the same length as the pedigree indicating sex, 0=male, 1=fe-male, any other value = unknown sex
samp	A vector the same length as the pedigree indicating whether or not each individual is sampled (1), or an unsampled dummy individual (0).

<code>sireE</code>	Value(s) indicating the paternal error rate. If it is a single number (between 0 and 1), it is applied to the entire pedigree; if it is a vector the length of the pedigree, then probabilities can vary among individuals.
<code>damE</code>	Value(s) indicating the maternal error rate. If it is a single number (between 0 and 1), it is applied to the entire pedigree; if it is a vector the length of the pedigree, then probabilities can vary among individuals.
<code>sireA</code>	Value(s) indicating the paternal pedigree link assignment rate. If it is a single number (between 0 and 1), it is applied to the entire pedigree; if it is a vector the length of the pedigree, then probabilities can vary among individuals.
<code>damA</code>	Value(s) indicating the maternal pedigree link assignment rate. If it is a single number (between 0 and 1), it is applied to the entire pedigree; if it is a vector the length of the pedigree, then probabilities can vary among individuals.
<code>cohort</code>	A numeric vector the same length as the pedigree containing cohorts
<code>first</code>	A numeric vector the same length as the pedigree indicating the first cohort for which an individual is to be considered a potential parent
<code>last</code>	A numeric vector the same length as the pedigree indicating the last cohort for which an individual is to be considered a potential parent
<code>monoecy</code>	An indicator specifying whether or not bisexuality is allowed (0=no (default), 1=yes)
<code>modifyAssumedPedigree</code>	An indicator variable specifying whether or not an assumed pedigree with errors but no missing links should be supplied.

Value

<code>assumedPedigree</code>	A pedigree differing from the supplied pedigree so as to mimic patterns of pedigree errors and missing data that might occur in a real study.
<code>truePedigree</code>	Echos the pedigree supplied.
<code>supplementalPedigree</code>	(optional) a 'assumed' pedigree containing errorsbut no missing links.

Author(s)

Michael Morrissey (michael.morrissey@ed.ac.uk)

References

Morrissey et al. 2007. Journal of Evolutionary Biology 20:2309-2321., Morrissey, M.B, and A.J. Wilson, 2009. pedantics, an R package for pedigree-based genetic simulation, and pedigree manipulation, characterisation, and viewing. Molecular Ecology Resources.

See Also

rpederr,fpederr

Examples

```
testData<-as.data.frame(matrix(c(
1,  NA,  NA,  1,  1,  1,  2,  2,
2,  NA,  NA,  1,  1,  1,  2,  2,
3,  NA,  NA,  1,  1,  1,  2,  2,
4,  NA,  NA,  1,  0,  1,  2,  2,
5,  NA,  NA,  1,  0,  1,  2,  2,
6,  1,  4,  0,  -1,  2,  3,  3,
7,  1,  4,  0,  -1,  2,  3,  3,
8,  1,  4,  0,  -1,  2,  3,  3,
9,  1,  4,  0,  -1,  2,  3,  3,
10, 2,  5,  0,  -1,  2,  3,  3,
11, 2,  5,  0,  -1,  2,  3,  3,
12, 2,  5,  0,  -1,  2,  3,  3,
13, 2,  5,  0,  -1,  2,  3,  3,
14, 3,  5,  0,  -1,  2,  3,  3,
15, 3,  5,  0,  -1,  2,  3,  3,
16, 3,  5,  0,  -1,  2,  3,  3,
17, 3,  5,  0,  -1,  2,  3,  3),
17,8,byrow=TRUE))

names(testData)<-c("id","dam","sire","founder","sex",
                 "cohort","first","last")
pedigree<-as.data.frame(cbind(testData$id,testData$dam,
                              testData$sire))
for(x in 1:3) pedigree[,x]<-as.factor(pedigree[,x])
names(pedigree)<-c("id","dam","sire")
pedigree

## some missing sire links:
fpederr(truePedigree=pedigree,founders=testData$founder,
        sex=testData$sex,sireA=0.5,cohort=testData$cohort,
        first=testData$first,last=testData$last)$assumedPedigree

## some erroneous sire links:
fpederr(truePedigree=pedigree,founders=testData$founder,
        sex=testData$sex,sireE=0.5,cohort=testData$cohort,
        first=testData$first,last=testData$last)$assumedPedigree
```

genomesim

A function to simulate QTL and/or SNP data.

Description

Simulates a chromosome of arbitrary length with arbitrary numbers, types, and spacings of genetic loci over arbitrary pedigrees.

Usage

```
genomesim(pedigree, founders=NULL, positions=NULL, initHe=NULL,
          mutationType=NULL, mutationRate=NULL, phenotyped=NULL,
          founderHaplotypes=NULL, genotyped=NULL, returnG='n',initFreqs=NULL)
```

Arguments

pedigree	A pedigree
founders	A vector of indicator variables denoting founder status (1=founder, 0=non-founder)
positions	Genome locations in cM for markers
initHe	Initial levels of expected heterozygosity
mutationType	A vector of locus types - see details
mutationRate	A vector of mutation rates
founderHaplotypes	A matrix or dataframe containing founder haplotypes
phenotyped	A vector of IDs of those individuals for which to return phenotypic data
genotyped	A vector of IDs of those individuals for which to return genotypic data
returnG	If 'y' then genotypic data for all loci (including cIAM loci) will be returned.
initFreqs	A list of allele frequencies for all loci. If initFreqs is specified, it will override information from initHe . extractA from package MasterBayes can be used to obtain initFreqs form a sample of genotypes. For cIAM loci, allele names in initFreqs should be allelic substitution effects.

Details

Valid mutation types are 'Micro', 'Dom', 'dIAM' and 'cIAM', for microsatellite, dominant (AFLP), discrete infinite alleles mutation model loci (SNPs), and continuous infinite alleles mutation model loci (polymorphisms effecting phenotypic variation). cIAM loci have mutational allelic substitution effects taken drawn from a normal distribution with mean 0 and variance 1.

Value

Phenotypes	A vector of phenotypes. Calculated as the sum of all allelic effects. Scaling is currently left to be done post-hoc.
MarkerData	A vector of marker genotypes, i.e. alleles at all loci except those designated 'cIAM'

Author(s)

Michael Morrissey (michael.morrissey@ed.ac.uk)

References

Morrissey, M.B, and A.J. Wilson, 2009. pedantics, an R package for pedigree-based genetic simulation, and pedigree manipulation, characterisation, and viewing. *Molecular Ecology Resources*.

See Also

phensim

Examples

```
testData<-as.data.frame(matrix(c(
  1,    NA,    NA,    1,    1,    1,    2,    2,
  2,    NA,    NA,    1,    1,    1,    2,    2,
  3,    NA,    NA,    1,    1,    1,    2,    2,
  4,    NA,    NA,    1,    0,    1,    2,    2,
  5,    NA,    NA,    1,    0,    1,    2,    2,
  6,    1,    4,    0,    -1,   2,    3,    3,
  7,    1,    4,    0,    -1,   2,    3,    3,
  8,    1,    4,    0,    -1,   2,    3,    3,
  9,    1,    4,    0,    -1,   2,    3,    3,
  10,   2,    5,    0,    -1,   2,    3,    3,
  11,   2,    5,    0,    -1,   2,    3,    3,
  12,   2,    5,    0,    -1,   2,    3,    3,
  13,   2,    5,    0,    -1,   2,    3,    3,
  14,   3,    5,    0,    -1,   2,    3,    3,
  15,   3,    5,    0,    -1,   2,    3,    3,
  16,   3,    5,    0,    -1,   2,    3,    3,
  17,   3,    5,    0,    -1,   2,    3,    3),
  17,8,byrow=TRUE))

names(testData)<-c("id","dam","sire","founder","sex",
                  "cohort","first","last")
pedigree<-as.data.frame(cbind(testData$id,testData$dam,
                              testData$sire))
for(x in 1:3) pedigree[,x]<-as.factor(pedigree[,x])
names(pedigree)<-c("id","dam","sire")
pedigree

##make up some microsatellite and gene allele frquencies:
sampleGenotypes<-as.data.frame(matrix(c(
  1,2,-1.32,0.21,2,1,0.21,0.21),2,4,byrow=TRUE))
testFreqs<-extractA(sampleGenotypes)

## note that alleles at the gene locus are given as their
## allelic substitution effects:
testFreqs

## simulate data for these individuals based on a single QTL
## with two equally alleles with balanced frequencies in the
## founders, linked (2 cM) to a highly polymorphic microsatellite:
genomesim(pedigree=pedigree,founders=testData$founder,positions=c(0,2),
          mutationType=c('Micro','cIAM'),mutationRate=c(0,0),initFreqs=testFreqs,returnG='y')
## since we specified returnG='y', we can check that the phenotypes add up to the
## allelic substitution effects for the second locus.
```

gryphons
Example dataset for pedantics examples and tutorial

Description

This contains pedigree and life history data of a fictional population. The data are relevant to power and sensitivity analyses for quantitative genetic studies of natural populations.

Usage

```
gryphons
```

Format

A table.

makePedigreeFactor
Converts a numeric pedigree to a pedigree with factors

Description

Some internal `pedantics` modules require that pedigrees be specified only by numerical values, including numerical values for missing data, this converts them back to factors

Usage

```
makePedigreeFactor(id, sire, dam, key)
```

Arguments

<code>id</code>	Numeric individual identifiers
<code>sire</code>	Numeric sire codes
<code>dam</code>	Numeric dam codes
<code>key</code>	A dataframe, as produced by <code>makePedigreeNumeric</code> , specifying factor codes for numeric values in <code>is</code> , <code>sire</code> , and <code>dam</code>

Value

returns the pedigree with all ids specified as factors according to `key`

Author(s)

Michael Morrissey (michael.morrissey@ed.ac.uk)

References

Morrissey, M.B, and A.J. Wilson, 2009. pedantics, an R package for pedigree-based genetic simulation, and pedigree manipulation, characterisation, and viewing. *Molecular Ecology Resources*.

See Also

makePedigreeNumeric

Examples

```
## first we'll implement the example from makePedigreeNumeric(),
## and use makePedigreeFactor() to turn it back again:

pedigree<-as.data.frame(matrix(c(
  "m1",  NA,  NA,
  "m2",  NA,  NA,
  "m3",  NA,  NA,
  "d4",  NA,  NA,
  "d5",  NA,  NA,
  "o6",  "m1", "d4",
  "o7",  "m1", "d4",
  "o8",  "m1", "d4",
  "o9",  "m1", "d4",
  "o10", "m2", "d5",
  "o11", "m2", "d5",
  "o12", "m2", "d5",
  "o13", "m2", "d5",
  "o14", "m3", "d5",
  "o15", "m3", "d5",
  "o16", "m3", "d5",
  "o17", "m3", "d5"),17,3,byrow=TRUE))

names(pedigree)<-c("id","dam","sire")
for(x in 1:3) pedigree[,x]<-as.factor(pedigree[,x])

## make the test pedigree numeric with NAs denoted by -1
test<-makePedigreeNumeric(id=as.character(pedigree[,1]),
                          dam=as.character(pedigree[,2]),
                          sire=as.character(pedigree[,3]),
                          missingVal=-1)

test$numericPedigree

test$idKey

## and turn it back again
makePedigreeFactor(id=test$numericPedigree$id,
                  dam=test$numericPedigree$dam,
                  sire=test$numericPedigree$sire,
                  key=test$idKey)
```

`makePedigreeNumeric` *Converts a pedigree with individuals specified as factors to a numeric pedigree*

Description

Some internal `pedantics` modules require that pedigrees be specified only by numerical values, including numerical values for missing data, this provides that conversion

Usage

```
makePedigreeNumeric(id, sire, dam, missingVal = NULL)
```

Arguments

<code>id</code>	Individual identifiers - pass using <code>as.character()</code>
<code>sire</code>	Sire codes - pass using <code>as.character()</code>
<code>dam</code>	Dam codes - pass using <code>as.character()</code>
<code>missingVal</code>	the indicator that should be substituted for missing values

Value

<code>numericPedigree</code>	The factor pedigree in numeric form
<code>idKey</code>	A key to facilitate conversion back to the original identifiers

Author(s)

Michael Morrissey (michael.morrissey@ed.ac.uk)

References

Morrissey, M.B, and A.J. Wilson, 2009. `pedantics`, an R package for pedigree-based genetic simulation, and pedigree manipulation, characterisation, and viewing. *Molecular Ecology Resources*.

See Also

`makePedigreeFactor`

Examples

```
pedigree<-as.data.frame(matrix(c(
  "m1",  NA,  NA,
  "m2",  NA,  NA,
  "m3",  NA,  NA,
  "d4",  NA,  NA,
  "d5",  NA,  NA,
  "o6",  "m1", "d4",
  "o7",  "m1", "d4",
  "o8",  "m1", "d4",
  "o9",  "m1", "d4",
  "o10", "m2", "d5",
  "o11", "m2", "d5",
  "o12", "m2", "d5",
  "o13", "m2", "d5",
  "o14", "m3", "d5",
  "o15", "m3", "d5",
  "o16", "m3", "d5",
  "o17", "m3", "d5"),17,3,byrow=TRUE))
names(pedigree)<-c("id","dam","sire")
for(x in 1:3) pedigree[,x]<-as.factor(pedigree[,x])

## make the test pedigree numeric with NAs denoted by -1
makePedigreeNumeric(id=as.character(pedigree[,1]),
                    dam=as.character(pedigree[,2]),
                    sire=as.character(pedigree[,3]),
                    missingVal=-1)
```

microsim

Simulates microsatellite data across a pedigree.

Description

Uses a pedigree with parents identified for all non-founding individuals and simulates microsatellite genotypes

Usage

```
microsim(pedigree, genFreqs = NULL, genotypesSample = NULL,
         knownGenotypes = NULL, records = NULL, eRate1 = 0,
         eRate2 = 0, eRate3 = 0)
```

Arguments

pedigree A pedigree

genFreqs (optional) A list of allele frequencies, can be produced with `extractA` in `MasterBayes`

genotypesSample	(required if genFreqs is not supplied) a sample of genotypes from which to estimate population allele frequencies
knownGenotypes	(not yet implemented) a data frame of genotypes for (potentially a subset) of founder individuals
records	Record availability, see details.
eRate1	The rate of genotypic substitution errors, i.e., when a true genotype at a given locus is replaced by a pair of alleles selected at random based on the population allele frequencies
eRate2	The rate of allelic substitution errors, i.e. when an allele is erroneously replaced at a given locus by an allele chosen at random based on the population allele frequencies
eRate3	The rate of large allele dropouts, simulated by setting the value of the larger allele at a locus to the value of the smaller allele

Details

Error rates and data availability rates can be specified as either (1) single values to be applied to all individuals and all loci, (2) as a vector the same length as the number of loci, representing locus-specific rates to be applied uniformly to all individuals, or (3) as data frames with rows for each individual and columns for each locus. In the third option, observed patterns of data availability can be simulated by supplying 0s and 1s for missing and available individual genotypes, respectively.

Value

trueGenotypes	A data frame of true genotypes
observedGenotypes	A data frame of plausible observed genotypes, given specified patterns of missingness and errors.

Author(s)

Michael Morrissey <michael.morrissey@ed.ac.uk>

References

Morrissey et al. 2007. *Journal of Evolutionary Biology* 20:2309-2321., Morrissey, M.B, and A.J. Wilson, 2009. *pedantics*, an R package for pedigree-based genetic simulation, and pedigree manipulation, characterisation, and viewing. *Molecular Ecology Resources*.

Examples

```
pedigree<-as.data.frame(matrix(c(
  "m1",  NA,  NA,
  "m2",  NA,  NA,
  "m3",  NA,  NA,
  "d4",  NA,  NA,
```

```

"d5", NA, NA,
"o6", "m1", "d4",
"o7", "m1", "d4",
"o8", "m1", "d4",
"o9", "m1", "d4",
"o10", "m2", "d5",
"o11", "m2", "d5",
"o12", "m2", "d5",
"o13", "m2", "d5",
"o14", "m3", "d5",
"o15", "m3", "d5",
"o16", "m3", "d5",
"o17", "m3", "d5"),17,3,byrow=TRUE))
names(pedigree)<-c("id","dam","sire")
for(x in 1:3) pedigree[,x]<-as.factor(pedigree[,x])

## some sample genotypes, very simple, two markers with He = 0.5
sampleGenotypes<-as.data.frame(matrix(c(
1,2,1,2,2,1,2,1),2,4,byrow=TRUE))

## locus names
names(sampleGenotypes)<-c("loc1a","loc1b","loc2a","loc2b")

## simulate some genotypes
microsim(pedigree=pedigree,genotypesSample=sampleGenotypes)

```

pedantics-package	<i>Tools to facilitate quantitative genetic studies of natural populations, especially with respect to the use of pedigrees in such problems.</i>
-------------------	---

Description

`pedantix` contains three types of functions. The first are functions specifically designed to aid power and sensitivity analyses for quantitative genetic studies, particularly with thought to accommodating the problems and data structures that arise in data from natural populations. There are basic utility functions for manipulating pedigrees. Finally there are functions for visualizing and statistically characterizing pedigrees.

Details

Package:	pedantics
Type:	Package
Version:	1.01
Date:	2009-07-21
License:	GPL-2 GPL-3
LazyLoad:	yes

See the tutorial, `pedantics-Tutorial.pdf` for detailed example analyses using `pedantics`

Author(s)

Michael Morrissey (michael.morrissey@ed.ac.uk)

Maintainer: Michael Morrissey (michael.morrissey@ed.ac.uk)

References

Morrissey et al. 2007. *Journal of Evolutionary Biology* 20:2309-2321., Morrissey, M.B, and A.J. Wilson, 2009. `pedantics`, an R package for pedigree-based genetic simulation, and pedigree manipulation, characterisation, and viewing. *Molecular Ecology Resources*.

<code>pedigreeStats</code>	<i>Calculates a range of statistics of pedigrees.</i>
----------------------------	---

Description

Statistics are those that will hopefully be useful for describing pedigrees to be used in quantitative genetic analyses of natural populations. This module will be most useful when cohort affinities for all individuals can be provided. All outputs are produced in a numerical form as well as in graphical summaries.

Usage

```
pedigreeStats(Ped, cohorts = NULL, dat = NULL,
              retain='informative', graphicalReport = "y",
              includeA=TRUE, lowMem=FALSE, grContrast=FALSE)
```

Arguments

<code>Ped</code>	A pedigree
<code>cohorts</code>	(Optional) Cohort affinities for members of the pedigree
<code>dat</code>	(Optional) Available data based upon which the pedigree can be pruned for just informative individuals
<code>retain</code>	The default value ('informative') results in pedigree being pruned to only those individuals who's records contribute to estimation of quantitative genetic parameters with respect to the available data specified in <code>dat</code> . Otherwise, specifying a value of 'ancestors' will result in the inclusion of all ancestors of phenotyped individuals.
<code>graphicalReport</code>	Controls whether or not graphical output is produced.
<code>includeA</code>	If TRUE, additive genetic relatedness matrix is returned.
<code>lowMem</code>	If TRUE, then stats based on calculation of A are not performed.
<code>grContrast</code>	If TRUE, then uglier shades of red and blue are used to denote male and female statistics in graphical reports, but these colours provide better contrast in greyscale.

Value

totalMaternities	Total number of maternities defined by the pedigree.
totalPaternities	Total number of paternities defined by the pedigree.
totalFullSibs	Total number of pair-wise full sib relationships defined by the pedigree.
totalMaternalSibs	Total number of pair-wise maternal sib relationships defined by the pedigree. To get the number of maternal half sibs, subtract totalFullSibs.
totalPaternalSibs	Total number of pair-wise paternal sib relationships defined by the pedigree. To get the number of paternal half sibs, subtract totalFullSibs.
totalMaternalGrandmothers	Total number of maternal grandmothers defined by the pedigree.
totalMaternalGrandfathers	Total number of maternal grandfathers defined by the pedigree.
totalPaternalGrandmothers	Total number of paternal grandmothers defined by the pedigree.
totalPaternalGrandfathers	Total number of paternal grandfathers defined by the pedigree.
pedigreeDepth	The pedigree depth, i.e. maximum number of ancestral generations, for each individual.
inbreedingCoefficients	Individual inbreeding coefficients
maternalSibships	Sibship size of each individual appearing in the dam column of the pedigree.
paternalSibships	Sibship size of each individual appearing in the sire column of the pedigree.
cumulativeRelatedness	Proportion of pair-wise relatedness values less than values ranging from 0 to 1.
relatednessCategories	Discretized distribution of relatedness.
analyzedPedigree	Returns the pedigree.
sampleSizesByCohort	(Optional) Number of individuals belonging to each cohort.
maternitiesByCohort	(Optional) Number of assigned maternities by offspring cohort.
paternitiesByCohort	(Optional) Number of assigned paternities by offspring cohort.
fullSibsByCohort	(Optional) Number of pair-wise full sib relationships by cohort - note the sum of these need not be equal to totalFullSibs in pedigrees of long-lived organisms.

maternalSibsByCohort
 (Optional) Number of pair-wise maternal sib relationships by cohort - note the sum of these need not be equal to totalMaternalSibs in pedigrees of long-lived organisms.

paternalSibsByCohort
 (Optional) Number of pair-wise paternal sib relationships by cohort - note the sum of these need not be equal to totalPaternalSibs in pedigrees of long-lived organisms.

maternalGrandmothersByCohort
 (Optional) Numbers of maternal grandmother assignments by offspring cohort.

maternalGrandfathersByCohort
 (Optional) Numbers of maternal grandfather assignments by offspring cohort.

paternalGrandmothersByCohort
 (Optional) Numbers of paternal grandmother assignments by offspring cohort.

paternalGrandfathersByCohort
 (Optional) Numbers of paternal grandfather assignments by offspring cohort.

cumulativePedigreeDepth
 (Optional) Distributions of pedigree depth by cohort.

meanRelatednessAmongCohorts
 (Optional) Mean relatedness among cohorts.

cohorts (Optional) Returns cohort designations.

Graphical summaries of a number of these summary statistics are printed to the console when `codegraphicalReports==y`.

Author(s)

Michael Morrissey (michael.morrissey@ed.ac.uk)

References

Morrissey, M.B, and A.J. Wilson, 2009. *pedantics*, an R package for pedigree-based genetic simulation, and pedigree manipulation, characterisation, and viewing. *Molecular Ecology Resources*.

See Also

`fixPedigree`

Examples

```
data(gryphons)
pedigree<-gryphons[,1:3]

gryphonsPedigreeSummary<-pedigreeStats(pedigree,
                                       cohorts=gryphons$cohort,graphicalReport='n')
```

```
gryphonsPedigreeSummary$totalMaternities
gryphonsPedigreeSummary$totalPaternities

gryphonsPedigreeSummary$maternitiesByCohort
gryphonsPedigreeSummary$paternitiesByCohort
```

pedStatSummary *Post-processes output from pedigreeStats*

Description

Generates a manageable summary of pedigree-wide statistics reported by pedigreeStats, either for a single pedigree or for a comparison between pedigrees

Usage

```
pedStatSummary(pedStats, pedStats2=NULL)
```

Arguments

`pedStats` An output data list from pedigreeStats
`pedStats2` An optional output data list from pedigreeStats

Value

Returns a table of numbers of records, maternities, paternities, pairwise sibship relationships, numbers of different classes of grand-parental relationships, pedigree depth, number of founders, mean subship sizes, simple statistics of numbers of inbred and non-inbred individuals, and proportions of pairwise relationship coefficients equal to or greater than several thresholds.

Author(s)

Michael Morrissey <michael.morrissey@ed.ac.uk>

References

Morrissey, M.B, and A.J. Wilson, 2009. pedantics, an R package for pedigree-based genetic simulation, and pedigree manipulation, characterisation, and viewing. Molecular Ecology Resources.

phensim *A function to simulated phenotypic data*

Description

Simulates phenotypic data across arbitrary pedigrees. `phensim` simulate direct, maternal and paternal genetica and environmental effects for an arbitrary number of traits with arbitrary patterns of missing data.

Usage

```
phensim(pedigree, traits = 1, randomA = NULL, randomE = NULL,
        parentalA = NULL, parentale = NULL, sampled = NULL,
        records = NULL, returnAlleffects = FALSE)
```

Arguments

<code>pedigree</code>	A pedigree
<code>traits</code>	The number of traits for which data should be simulated.
<code>randomA</code>	An additive genetic covariance matrix, with dimensions a multiple of traits - see details
<code>randomE</code>	An additive environmental covariance matrix, with dimensions a multiple of traits - see details
<code>parentalA</code>	A vector indicating which effects in <code>randomA</code> (if any) to treat as parental effects
<code>parentale</code>	A vector indicating which effects in <code>randomE</code> (if any) to treat as parental effects
<code>sampled</code>	A vector indicating which individuals are sampled
<code>records</code>	A single value, array of matrix specifying data record availability - see details
<code>returnAlleffects</code>	If <code>TRUE</code> then all individual breeding values and environmental effects are returned

Details

`randomA` and `randomE` are square matrices with dimension equal to the sum of the number direct and indirect effects. This must be a multiple of the number of traits, i.e. if an indirect effect is to be simulated for only one of multiple traits, those traits with no indirect effect should be included with (co)variances of zero.

`parentalA` and `parentale` are optional vectors of characters indicating which trait positions in `randomA` and `randomE` are to be treated as indirect effects, and which effects to treat as maternal or paternal. Valid values are 'd', 'm', and 'p', for direct, maternal indirect and paternal indirect effects, respectively.

records can be specified either (1) as a single value to be applied to all individuals and traits, (2) as a vector the same length as the number of traits, representing trait-specific rates to be applied uniformly to all individuals, or (3) as data frames with rows for each individual and columns for each trait. In the third option, observed patterns of data availability can be simulated by supplying 0s and 1s for missing and available individual genotypes, respectively.

Value

phenotypes A dataframe containing phenotypes for all individuals specified to have records.

allEffects (optional) A dataframe with all direct and indirect genetic and environmental effects.

Author(s)

Michael Morrissey (michael.morrissey@ed.ac.uk)

References

Morrissey et al. 2007. *Journal of Evolutionary Biology* 20:2309-2321., Morrissey, M.B, and A.J. Wilson, 2009. pedantics, an R package for pedigree-based genetic simulation, and pedigree manipulation, characterisation, and viewing. *Molecular Ecology Resources*.

Examples

```
## make up a pedigree
id<- c("a1","a2","a3","a4","a5","a6","a7","a8","a9")
dam<- c(NA,NA,NA,"a1","a1","a1","a4","a4","a4")
sire<- c(NA,NA,NA,"a2","a2","a2","a5","a6","a6")
pedigree<-as.data.frame(cbind(id,sire,dam))

traits<-2
## no correlations
randomA<-diag(4)
randomE<-diag(4)
parentalA<-c("d","d","m","m")
parentalE<-c("d","d","m","m")

## generate phenotypic data based on this architecture
phensim(pedigree=pedigree,traits=2,randomA=randomA,randomE=randomE,
        parentalA=parentalA,parentalE=parentalE)

## let's do it again but see how the phenotypes were composed
phensim(pedigree=pedigree,traits=2,randomA=randomA,randomE=randomE,
        parentalA=parentalA,parentalE=parentalE,returnAllEffects=TRUE)
```

rpederrBird
Permutes a bird pedigree to create a plausible complete pedigree

Description

Permutes a pedigree to create a plausible complete pedigree, with errors accounted for in terms of extra-pair paternity.

Usage

```
rpederrBird(assumedPedigree, founders = NULL, sex = NULL, samp = NULL,
            sireS = NULL, damS = NULL, cohort = NULL, first = NULL,
            last = NULL, broods = NULL, monocecy = 0, broodData = NULL,
            EPPsireData = NULL, propEPPbroods = 1,
            propEPPchicksGivenEPPbrood = 0.5, propEPPbroodsTwoFathers = 0,
            modifyAssumedPedigree = 0, EPPlambda=0.001,
            malePhenotypes=NULL, EPPbeta=0, EPPgamma=0)
```

Arguments

assumedPedigree	A pedigree
founders	A vector of indicator variables denoting founder status (1=founder, 0=non-founder)
sex	A vector of indicator variables denoting sex (0=male,1=female,anything else=unknown)
samp	A vector denoting whether or not individuals are sampled (1), or dummy individuals (0) added to the pedigree for the purpose of simulating potential "true" pedigree links that go outside the sampled population
sireS	Proportion of "true" simulated sires that are to be taken from the unsampled portion of the pedigree.
damS	Proportion of "true" simulated dams that are to be taken from the unsampled portion of the pedigree.
cohort	A numeric vector the same length as the pedigree containing cohorts
first	A numeric vector the same length as the pedigree indicating the first cohort for which an individual is to be considered a potential parent
last	A numeric vector the same length as the pedigree indicating the last cohort for which an individual is to be considered a potential parent
broods	A numeric vector the same length as the pedigree indicating the brood to which each individual belongs. Can be any integer value for individuals that do not belong to broods as specified in broodData (see below).
monocecy	An indicator specifying whether or not bisexuality is allowed (0=no (default), 1=yes)

<code>broodData</code>	A five column data frame containing (1) brood IDs corresponding to broods, (2) years, (3) brood latitudes, (4) brood longitudes, and (5) an indicator (0=no or 1=yes) specifying whether or not the particular brood potentially contains extra-pair chicks.
<code>EPPsireData</code>	A four column data frame with (1) ids of all sires that may be extra-pair fathers, (2) years (cohorts) in which an individual might have been an extra pair father (one separate line for each year), and (3) latitude and (4) longitude in each year.
<code>propEPPbroods</code>	Proportion of broods with extra-pair chicks.
<code>propEPPchicksGivenEPPbrood</code>	Mean proportion of chicks that have extra-pair fathers in extra-pair broods.
<code>propEPPbroodsTwoFathers</code>	Proportion of extra-pair broods that have two extra-pair fathers
<code>modifyAssumedPedigree</code>	An indicator variable specifying whether or not an assumed pedigree with errors but no missing links should be supplied.
<code>EPPlambda</code>	The rate of decline of the probability of extra pair paternity with distance between males and nests.
<code>malePhenotypes</code>	(optional) if specified, <code>zp</code> will take these values in the calculation of male relative fertilities, see details
<code>EPPbeta</code>	(optional) coefficient of linear selection acting through undetected extra-pair breeding success, see details
<code>EPPgamma</code>	(optional) coefficient of quadratic selection acting through undetected extra-pair breeding success, see details

Value

<code>assumedPedigree</code>	echos the supplied pedigree
<code>truePedigree</code>	A plausible pedigree with no errors and no missing links
<code>supplementalPedigree</code>	A plausible pedigree with errors but no missing links

Author(s)

Michael Morrissey <michael.morrissey@ed.ac.uk>

References

Morrissey, M.B, and A.J. Wilson, 2009. `pedantics`, an R package for pedigree-based genetic simulation, and pedigree manipulation, characterisation, and viewing. *Molecular Ecology Resources*.

See Also

`rpederr` for a simpler module that may be adequate when pedigree errors are (assumably) independent between offspring.

Examples

```
## test data for a single brood
id<- c("a1","a2","a3","a4","a5","a6","a7","a8","a9")
dam<- c(NA,NA,NA,NA,"a1","a1","a1","a1","a1")
sire<- c(NA,NA,NA,NA,"a2","a2","a2","a2","a2")
broods<- c(NA,NA,NA,NA,"r1","r1","r1","r1","r1")
founders<-c(1,1,1,1,0,0,0,0,0)
samp<- c(1,1,1,1,1,1,1,1,1)
sex<- c(1,0,0,0,0,0,1,0,0)
sireS<- rep(1,9)
damS<- rep(1,9)
cohort<-c(1,1,1,1,2,2,2,2,2)
first<-c(2,2,2,3,3,3,3,3,3)
last<-c(2,2,2,3,3,3,3,3,3)

broodIDs<-c("r1")
broodYears<-2
broodN<-c(10)
broodE<-c(10)
broodEPPpotential<-c(1)
EPPsireIDs<-c("a2","a3","a4")
EPPsireYears<-c(2,2,2)
sireN<-c(10,11,5)
sireE<-c(10,11,5)
EPPlambda<- 0.4

## put the test data in the required forms
assumedPedigree<-as.data.frame(cbind(id,sire,dam))
broodData<-as.data.frame(cbind(broodIDs,broodYears,broodN,
                               broodE,broodEPPpotential))
EPPsireData<-as.data.frame(cbind(EPPsireIDs,EPPsireYears,sireN,sireE))

## simulate extra-pair paternity (try a few times and see how often
## the paterity goes to the nearer candidate EPP father):
rpederrBird(assumedPedigree=assumedPedigree,founders=founders,sex=sex,
            samp=samp,sireS=sireS,damS=damS,cohort=cohort,first=first,
            last=last,broods=broods,broodData=broodData,EPPsireData=EPPsireData,
            propEPPbroods=1,propEPPchicksGivenEPPbrood=0.5,
            propEPPbroodsTwoFathers=0,modifyAssumedPedigree=0)
```

rpederr

Permutes a pedigree to create a plausible complete pedigree

Description

Given estimates of individual life histories and rates and patterns of errors in pedigree links, `rpederr` probabilistically assigns "true" parents given an incomplete and potentially erroneous pedigree.

Usage

```
rpederr(assumedPedigree, founders = NULL, sex = NULL, samp = NULL,
        sireE = NULL, damE = NULL, sireS = NULL, damS = NULL,
        cohort = NULL, first = NULL, last = NULL, monocecy = 0,
        modifyAssumedPedigree = 0)
```

Arguments

assumedPedigree A pedigree

founders A vector of indicator variables denoting founder status (1=founder, 0=non-founder)

sex A vector of indicator variables denoting sex (0=male,1=female,anything else=unknown)

samp A vector denoting whether or not individuals are sampled (1), or dummy individuals (0) added to the pedigree for the purpose of simulating potential "true" pedigree links that go outside the sampled population

sireE Sire assignment error rates, see details

damE Dam assignment error rates, see details

sireS Proportion of "true" simulated sires that are to be taken from the unsampled portion of the pedigree.

damS Proportion of "true" simulated dams that are to be taken from the unsampled portion of the pedigree.

cohort A numeric vector the same length as the pedigree containing cohorts

first A numeric vector the same length as the pedigree indicating the first cohort for which an individual is to be considered a potential parent

last A numeric vector the same length as the pedigree indicating the last cohort for which an individual is to be considered a potential parent

monocecy An indicator specifying whether or not bisexuality is allowed (0=no (default), 1=yes)

modifyAssumedPedigree An indicator variable specifying whether or not an assumed pedigree with errors but no missing links should be supplied.

Value

assumedPedigree echos the supplied pedigree

truePedigree A plausible pedigree with no errors and no missing links

supplementalPedigree A plausible pedigree with errors but no missing links

Author(s)

Michael Morrissey (michael.morrissey@ed.ac.uk)

References

Morrissey et al. 2007. Journal of Evolutionary Biology 20:2309-2321., Morrissey, M.B, and A.J. Wilson, 2009. pedantics, an R package for pedigree-based genetic simulation, and pedigree manipulation, characterisation, and viewing. Molecular Ecology Resources.

See Also

rpederrBird

Examples

```
id<- c("a1","a2","a3","a4","a5","a6","a7","a8","a9")
dam<- c(NA,NA,NA,"a1","a1","a1","a4","a4","a4")
sire<- c(NA,NA,NA,NA,NA,NA,NA,"a5","a5","a5")
found<-c(1,1,1,0,0,0,0,0,0)
samp<- c(1,1,1,1,1,1,1,1,1)
sex<- c(1,0,0,1,0,0,1,0,0)
dade<- rep(0,9)
mume<- rep(0,9)
dads<- rep(1,9)
mums<- rep(1,9)
cohort<-c(1,1,1,2,2,2,3,3,3)
first<-c(2,2,2,3,3,3,4,4,4)
last<-c(2,2,2,3,3,3,4,4,4)
pedigree<-as.data.frame(cbind(id,sire,dam))

### don't simulate any errors, just fill in the missing sires
rpederr(assumedPedigree=pedigree,founders=found,sex=sex,
        samp=samp,cohort=cohort,first=first,last=last)

## fill in the missing sires, and additionally simulate a problem
## with the second maternal sibship note that it is probabilistic,
## so this example may need to be run a couple of times before the
## error comes up, given the very small example dataset
fatherErrors<-c(0,0,0,0,0,0,0.8,0.8,0.8)
rpederr(assumedPedigree=pedigree,founders=found,sex=sex,samp=samp,
        sireE=fatherErrors,cohort=cohort,first=first,last=last)
```

Index

- *Topic **datagen**
 - fpederr, 5
 - genomesim, 7
 - microsim, 13
 - phensim, 19
 - rpederr, 24
 - rpederrBird, 21
- *Topic **datasets**
 - gryphons, 9
- *Topic **hplot**
 - drawPedigree, 1
- *Topic **manip**
 - fixPedigree, 4
 - makePedigreeFactor, 10
 - makePedigreeNumeric, 11
 - pedStatSummary, 19
- *Topic **package**
 - pedantics-package, 15

drawPedigree, 1

fixPedigree, 3, 4, 18

fpederr, 5, 6

genomesim, 7

gryphons, 9

makePedigreeFactor, 10, 12

makePedigreeNumeric, 10, 11

microsim, 13

pedantics (*pedantics-package*), 15

pedantics-package, 15

pedigreeStats, 16

pedStatSummary, 19

phensim, 8, 19

rpederr, 6, 23, 24

rpederrBird, 21, 25