Package 'AFLPsim'

October 27, 2013

Type Package

Title Hybrid simulation and genome scan for dominant markers

Version 0.2-1

Encoding latin1

Date 2013-10-27

Author Francisco Balao and Juan Luis García-Castaño

Maintainer Francisco Balao <fbalao@us.es>

Depends R (>= 2.15.0)

Imports adegenet, introgress

SystemRequirements Bayescan 2.1. See README.

Description This package is developed in the Plant Reproductive Biology Lab (RNM-214) - University of Seville. It contains hybrid simulation functions for dominant genetic data. It also provides several genome scan methods.

License GPL (>= 2)

URL http://www.r-project.org,
http://personal.us.es/fbalao/software.html

LazyData true

R topics documented:

AFLPsim-package   *Hybrid simulation and genome scan for dominant markers*

Description

This package is developed in the Plant Reproductive Biology Lab (RNM-214) - University of Seville. It contains hybrid simulation functions for dominant genetic data. It also provides several genome scan methods.

More information on AFLPsim can be found at `https://github.com/fbalao/AFLPsim`.

To cite AFLPsim, please use citation("AFLPsim")

Details

Package:	AFLPsim
Type:	Package
Version:	0.2-1
Date:	2013-10-27
License:	GPL (>= 2)

Author(s)

Francisco Balao; Juan Luis Garc?a-Casta?o

---

bayescan   *Identifying candidate loci under natural selection with external application*

Description

This function calls Bayescan program from within R to identifying candidate loci under natural selection from genetic data.

Usage

```
bayescan(mat, filename, nbp = 20, pilot = 5000, burn = 50000, exec=NULL)
```

Arguments

`mat`	A matrix with genotypic data to test in hybridsim format
`filename`	a character string giving the name of the output file (without extension)
`nbp`	Number of pilot runs (default is 2)
`pilot`	Length of pilot runs (default is 50)
`burn`	Burnin length (default is 5000)
`exec`	a character string giving the path to BAYESCAN. By default it tries to guess it depending on the operating system (see details).

Details

`bayescan` tries to guess the name of the executable program depending on the operating system. Specifically, the followings are used: "bayescan_2.1" under Linux and Mac, or "C:/Program Files/BayeScan2.1/binaries/BayeScan2.1_win32bits_cmd_line.exe" under Windows.

Value

Several files with the results and a data.frame with the following variables:

`prob`	The posterior probability for the model including selection
`log10.PO.`	The logarithm of Posterior Odds to base 10
`qval`	q-values for each locus for the model including selection
`alpha`	The estimated alpha coefficient indicating the strength and direction of selection. See Bayescan 2.1 manual
`fst`	The Fst coefficient averaged over populations

Author(s)

F. Balao <fbalao@us.es>, J.L. García-Castaño

References

Foll, M. & O. Gaggiotti. 2008. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics* **180**: 977-993.

See Also

`gscan sim2bayescan`

Examples

```
## Not run
hybrids<-hybridsim(Nmarker=100, Na=30, Nb=30, Nf1=30, type="selection", Nsel=25, hybrid="
outbayes<-bayescan(hybrids, filename="bayescanresults", nbp=10, pilot=50, burn=100)
outbayes
```

| demosimhybrid | *Demographic model of introgressive hybridization* |

Description

This model simulates the proportions of parentals, F1, F2, Fx and backcross (on both sides) individuals for each generation and takes into account the initial frequencies of parentals, the assortative mating among taxa as well as fitness differences.

Usage

```
demosimhybrid(x,M,F)
```

Arguments

x	A vector indicating the initial abundances in the population.The vector should sum 1. The order of abundances is: ParentalA, ParentalB, F1, BxA, BxB and Fx
M	Matrix assortative mating.The size is 6x6 following the same order tan vector x. By default random mating (all = 1)
F	A vector indicating the diferent fecundities of the parentals and hybrids. The vector size is 6 following the same order tan vector x. By default equal fecundities (all = 1)

Details

This function simulate the model of introgresive hibridization of Epifanio and Philipp (2000)

Value

a object `demosimhybrid`. An matrix of abundances in each generation

Author(s)

Francisco Balao <fbalao@us.es>; Marcial Escudero; J.L. García-Castaño

References

Epifanio, J. & D. Philipp. 2000. Simulating the extinction of parental lineages from introgressive hybridization: the effects of fitness, initial proportions of parental taxa, and mate choice. *Reviews in Fish Biology and Fisheries* **10**: 339-354.

See Also

```
hybridsim
```

```
plot.demosimhybrid
```

Examples

```
## Example 1. Simulation under parental proportions,
## similar fecundities and random mating
inivalues<-c(0.5,0.5,0,0,0,0)
epi0.5<-demosimhybrid(inivalues)
epi0.5


## Example 2. Simulation under higher frecuency of parental B,
## and higher fecundy of parental A and random mating
inivalues2<-c(0.25,0.75,0,0,0,0)
fecundities<-c(1,0.5,0.5,0.5,0.5,0.5)
epi0.75<-demosimhybrid(x=inivalues, F=fecundities)
epi0.75
```

---

gscan   *Genome scan for hybrids*

Description

This function fits genomic scan to dominant genotypic data using the method described by *Gagnaire et al (2009)* and the new method by Balao et al (2013; *in preparation*). Significance testing for outlier loci is included.

Usage

```
gscan(mat, type=c("F1","BxA","BxB"), method=c("bal&gar-ca","gagnaire"))
```

Arguments

mat	an object of class 'hybridsim' produced by 'hybridsim' or 'hybridize' functions
type	the type of hybrid classes; either "F1", "BxA" or "BxB"
method	a character string specifying the method to test significance of outlier loci; either "gagnaire" or "bal&gar-ca". See Details.

Details

These genome scan methods calculate the null distribution of frequencies under a neutral model.

Gagnaire's method uses a binomial test to outlier significance. For more conservative and unbiased method, "Bal&gar-car" method calculates the 95% confidence expected hybrid frequencies by the Clopper-Pearson 'exact' procedure (Clopper & Pearson 1934; Brown et al. 2001).

In both methods, the False Discovery Rate (FDR) correction (Benjamini & Hochberg 1995) is used to counteract for multiple comparisons and control the expected proportion of incorrectly rejected null hypotheses.

Value

A list with the following components:

`P-values`   a matrix with P values after False Discovery Rate correction for each loci

`Outlier`   a vector with outliers

Author(s)

F. Balao `<fbalao@us.es>`, J.L. García-Castaño

References

Balao, F., Casimiro-Soriguer, R., García-Castaño, J.L., Terrab, A., Talavera, S. 2013. Big thistle eats the little thistle: Non-neutral unidirectional introgression endangers the conservation of *Onopordum hinojense*. *Molecular Ecology*, *in preparation*.

Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B* **57**: 289-300.

Brown LD, Cai TT, Anirban D (2001) Interval estimation for a binomial proportion. *Statistical Science* **16**: 101-117.

Clopper CJ, Pearson ES (1934) The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika* **26**: 404-413

Gagnaire, P.A., V. Albert, B. Jonsson, L. Bernatchez. 2009. Natural selection influences AFLP intraspecific genetic variability and introgression patterns in Atlantic eels. *Molecular Ecology* **18**: 1678-1691.

See Also

`hybridsim`

Examples

```
hybrids<-hybridsim(Nmarker=100, Na=30, Nb=30, Nf1=30, type="selection", S=5,Nsel=25, hybr

outliers<-gscan(hybrids, type="F1", method="bal&gar-ca")
```

  `hybridindex`   *Estimate hybrid index por* `hybridsim` *objects*

Description

This function finds maximum likelihood estimates of hybrid index as described by *Buerkle (2005)* using the packages **introgress**

Usage

```
hybridindex(data)
```

Arguments

data   an `hybridsim` object with the profiles of parentals and hybrids.

Details

`hybridindex` returns a hybrid index estimate with its 95% confidence interval. See `est.h` and *Buerkle (2005)* for additional details.

Value

A data frame with estimates of hybrid index and upper and lower limits of its 95% confidence intervals, which falls within two support units of the maximum-likelihood estimate:

lower   95% confidence interval lower limit.

h   Maximum-likelihood estimate of the hybrid index.

upper   95% confidence interval upper limit.

Author(s)

F. Balao `<fbalao@us.es>`; J.L. Garc?a-Casta?o

References

Buerkle C. A. (2005) Maximum-likelihood estimation of a hybrid index based on molecular markers. *Molecular Ecology Notes*, **5**: 684-687.

Gompert Z. and Buerkle C. A. (2010) introgress: a software package for mapping components of isolation in hybrids. *Molecular Ecology Resources*,**10**, 378-384

See Also

`est.h`

Examples

```
## simulate parentals and F1 hybrids
hybrids<-hybridsim(Nmarker=50, Na=10, Nb=10, Nf1=10, type="neutral", hybrid="F1")

## estimate hybrid index
hest<-hybridindex(hybrids)
```

---

hybridize   *AFLP simulator with selection from parental data*

Description

This function simulates AFLP profiles (or other dominant markers) of several hybrid classes (F1 and backcrosses on both parentals) from two parental populations. In addition, selection on several markers can be simulated for the hybrid progeny.

Usage

```
hybridize(pa,pb,Nf1, Nbxa = Nf1, Nbxb = Nf1, Nf2 = Nf1, type = "selection",
 hybrid = "all", Nsel = Nmarker * 0.1, S = 0)
```

Arguments

pa	AFLP profile of Parental A. A data.frame or matrix.
pb	AFLP profile of Parental B. A data.frame or matrix.
Nf1	number of AFLP profiles for F1.
Nbxa	number of AFLP profiles for BxA.
Nbxb	number of AFLP profiles for BxA.
Nf2	number of AFLP profiles for F2.
type	type of simulation: 'neutral' or 'selection'.
hybrid	hybrid classes to simulate. By default 'all'. Also 'F1', 'BxA', 'BxB' or 'F2'.
Nsel	number of loci under selection.
S	Selection coefficient (see Details)

Details

simulates F1, F2, Backcross to Parental A (BxA) and Backcross to Parental B (BxB) with and without selection. Parental profiles must be included (individuals in rows and markers in columns). The selection coefficient S is a integer value from -10 to 10 (i.e. it can include negative, neutral or positive selection).

Value

Object hybridsim with:

PA	Parental A AFLP profile
PB	Parental B AFLP profile
F1	F1 hybrid AFLP profile
F2	F2 hybrid AFLP profile
BxA	Backcross to Parental A AFLP profile

BxB	Backcross to Parental B AFLP profile
Nsel	Loci under selection (default, under neutral selection)
S	Selection coefficient

Author(s)

F. Balao <fbalao@us.es>, J.L. García-Castaño

References

Wright, S. 1931. Evolution in Mendelian populations. *Genetics* **16**: 97-159.

See Also

hybridsim

Examples

```
## AFLP profile Species A
SpeciesA<-read.table(system.file("/files/SpeciesA.txt",package="AFLPsim"),header=TRUE, ro


## AFLP profile Species B
SpeciesB<-read.table(system.file("/files/SpeciesB.txt",package="AFLPsim"),header=TRUE, ro

## simulate F1 hybrids
F1hybrids<-hybridize(pa=SpeciesA,pb=SpeciesB, Nf1=30, type="neutral", hybrid="F1")
```

---

hybridsim	*AFLP simulator with selection*

Description

This function simulates AFLP profiles (or other dominant markers) of parentals and several hybrid classes (F1 and backcrosses on both parentals). In addition, selection on several markers can be simulated for the hybrid progeny.

Usage

```
hybridsim(Nmarker, Na, Nb, Nf1, Nbxa = Nf1, Nbxb = Nf1, Nf2 = Nf1,
 type = "selection", hybrid = "all", Nsel = Nmarker * 0.1, S = 0, apa = 0.5, apb
```

Arguments

`Nmarker`	The number of AFLP loci to simulate.
`Na`	number of AFLP profiles for Parental A.
`Nb`	number of AFLP profiles for Parental B.
`Nf1`	number of AFLP profiles for F1.
`Nbxa`	number of AFLP profiles for BxA.
`Nbxb`	number of AFLP profiles for BxB.
`Nf2`	number of AFLP profiles for F2.
`type`	type of simulation: `'neutral'` or `'selection'`.
`hybrid`	hybrid classes to simulate. By default `'all'`. Also `'F1'`, `'BxA'`, `'BxB'` or `'F2'`.
`Nsel`	number of loci under selection.
`S`	Selection coefficient (see Details)
`apa`	value for parameter 1 of the beta distribution
`apb`	value for parameter 2 of the beta distribution

Details

Simulate dominant markers (AFLP, RFLP...). Parental allele frequencies are calculated following a beta distribution (Wright 1931). F1, F2, Backcrosses to parental A (BxA) and Backcrosses to parental b (BxB) can be simulated with and without selection. The selection coefficient S is a integer value from -1 to 1 (negative and positive selection).

Value

Object `hybridsim` with:

`PA`	AFLP profile parental population A
`PB`	AFLP profile parental population A
`F1`	AFLP profile F1 hybrid population
`F2`	AFLP profile F2 hybrid population
`BxA`	AFLP profile Backcrosses to parental population A
`BxB`	AFLP profile Backcrosses to parental population B
`Nsel`	Loci under selection. NA under neutral selection
`S`	Selection Coefficient

Author(s)

F. Balao <fbalao@us.es>, J.L. García-Castaño

References

Wright, S. 1931. Evolution in Mendelian Populations. *Genetics* **16**: 97-159.

See Also

```
hybridize
```

Examples

```
hybrids<-hybridsim(Nmarker=100, Na=30, Nb=30, Nf1=30, type="selection", Nsel=25, hybrid="
```

---

plot.demosimhybrid *Plotting* demosimhybrid *objects*

Description

A demosimhybrid object can be plotted using the function plot.demosimhybrid, which is also used as the dedicated plot method. This function plots the frequency of parentals and hybrid classes on each generation.

Usage

```
## S3 method for class 'demosimhybrid'
plot(x, col = c(2, 3, 4, "orange", "orchid", 7), ...)
```

Arguments

x	a demosimhybrid object.
col	the colors for the hybrid classes.
...	Arguments to be passed to methods, such as graphical parameters (see par).

Author(s)

F. Balao <fbalao@us.es>, J.L. García-Castaño

See Also

```
demosimhybrid
```

Examples

```
## Example 1. Simulation under parental proportions,
## similar fecundities and random mating
inivalues<-c(0.5,0.5,0,0,0,0)
epi0.5<-demosimhybrid(inivalues)
epi0.5
plot.demosimhybrid(epi0.5)

## Example 2. Simulation under higher frecuency of Parental B,
## and higher fecundy of Parental A and random mating
inivalues2<-c(0.25,0.75,0,0,0,0)
fecundities<-c(1,0.5,0.5,0.5,0.5,0.5)
```

```
epi0.75<-demosimhybrid(x=inivalues, F=fecundities)
epi0.75
plot.demosimhybrid(epi0.75)
```

---

plot.hybridsim   *Plotting hybridsim objects*

Description

A `hybridsim` object can be plotted using the function `plot.hybridsim`, which is also used as the dedicated plot method. This function represents expected hybrid markers frequencies on a neutral sheet.

Usage

```
## S3 method for class 'hybridsim'
plot(x,hybrid = c("F1", "BxA", "BxB"), col = "lightgreen",
 shade = 0.8, markers = x$SelMarkers, ...)
```

Arguments

x	A `hybridsim` object.
hybrid	The hybrid classes to simulate: `"F1"`, `"BxA"` or `"BxB"`.
col	A specification for the default plotting color.
shade	A specification for the default alpha value.
markers	a numeric vector with markers to plot. By default markers under selection by `hybridsim` function.
...	Arguments to be passed to methods, such as graphical parameters (see `par`).

Author(s)

F. Balao <fbalao@us.es>, J.L. García-Castaño

See Also

`hybridsim` `hybridize`

Examples

```
## simulate parentals and F1 hybrids
hybrids<-hybridsim(Nmarker=100, Na=30, Nb=30, Nf1=30, type="selection", S=1, Nsel=25, hyt

plot.hybridsim(hybrids, hybrid="F1")
```

---

sim2arlequin   *Converting hybridsim object to a Arlequin input file*

Description

The function sim2arlequin converts a hybridsim object into a Arlequin input file.

Usage

```
sim2arlequin(x,filename)
```

Arguments

x   a `hybridsim` object

filename   a character string indicating the name of the output file

Value

Arlequin input file

Author(s)

F. Balao <fbalao@us.es>, J.L. García-Castaño

References

Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47-50.

See Also

sim2bayescan sim2introgress sim2newhybrids sim2popgene sim2structure

Examples

```
## simulate F1 hybrids
F1hybrids<-hybridsim(Nmarker=100,Na=100,Nb=100,Nf1=30, type="neutral", hybrid="F1")

## convert to Arlequin input file
sim2arlequin(F1hybrids,filename="F1hybrids_Arlequin.txt")
```

---

sim2bayescan   *Converting hybridsim object to a Bayescan input file*

Description

The function sim2bayescan converts a hybridsim object into a `Bayescan` *(Foll & Gaggiotti 2008)* input file.

Usage

```
sim2bayescan(x,filename)
```

Arguments

x   a `hybridsim` object. Only with F1 hybrids

filename   a character string indicating the name of the output file

Value

`Bayescan` input file

Author(s)

F. Balao <fbalao@us.es>, J.L. García-Castaño

References

Foll, M. & O. Gaggiotti. 2008. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics* **180**: 977-993.

See Also

`sim2arlequin` `sim2introgress` `sim2newhybrids` `sim2popgene` `sim2structure`

Examples

```
## simulate F1 hybrids
F1hybrids<-hybridsim(Nmarker=100,Na=100,Nb=100,Nf1=30, type="selection", S=5, Nsel=25, hy

## convert to Bayescan input file
sim2bayescan(F1hybrids,filename="F1hybrids_Bayescan.txt")
```

---

sim2genind   *Converting a hybridsim object to a genind object*

Description

The function `sim2genind` converts a `hybridsim` object into a `genind` object of the package
adegenet. It is wrapper of the function `df2genind`

Usage

```
sim2genind(x)
```

Arguments

x   a `hybridsim` object

Value

A `genind` object

Author(s)

F. Balao `<fbalao@us.es>`, J.L. García-Castaño

See Also

`genind df2genind`

Examples

```
## simulate F1 hybrids
F1hybrids<-hybridsim(Nmarker=100,Na=100,Nb=100,Nf1=30, type="neutral", hybrid="F1")

## convert to genind object
F1gen<-sim2genind(F1hybrids)
```

---

sim2introgress   *Converting hybridsim object to introgress*

Description

The function sim2introgress converts a hybridsim object into an `introgress` input file. It is a
wrapper to the function `prepare.data` of the package **introgress**

Usage

```
sim2introgress(x)
```

Arguments

x   a `hybridsim` object

Value

a list returned by the function `prepare.data` of the package **introgress**

Author(s)

F. Balao `<fbalao@us.es>`, J.L. García-Castaño

References

Gompert, Z. & C.A. Buerkle. 2010. introgress: a software package for mapping components of isolation in hybrids. *Molecular Ecology Resources* **10**: 378-384.

See Also

`prepare.data sim2arlequin sim2bayescan sim2newhybrids sim2popgene sim2structure`

Examples

```
## simulate hybrids
hybrids<-hybridsim(Nmarker=100,Na=100,Nb=100,Nf1=30, type="selection", hybrid="all")

## convert to introgress input file
hybrids2<-sim2introgress(hybrids)
```

---

 sim2newhybrids   *Converting hybridsim object to a NewHybrids input file*

Description

The function `sim2newhybrids` converts a AFLPsim object into a NewHybrids input file.

Usage

```
sim2newhybrids(x,filename)
```

Arguments

x   a `hybridsim` object
filename   a character string indicating the name of the output file

Value

a NewHybrids input file

Author(s)

F. Balao `<fbalao@us.es>`, J.L. Garc?a-Casta?o

References

Anderson, E.C. 2008. Bayesian inference of species hybrids using multilocus dominant genetic markers. *Philosophical transactions of the Royal Society of London. Series B, Biological Sciences* **363**: 2841-2850.

Anderson, E.C., & E.A. Thompson. 2002. A model-based method for identifying species hybrids using multilocus genetic data. *Genetics* **160**: 1217-1229.

See Also

`sim2arlequin sim2bayescan sim2introgress sim2popgene sim2structure`

Examples

```
## simulate hybrids
hybrids<-hybridsim(Nmarker=100,Na=30,Nb=30,Nf1=30, type="neutral", hybrid="all")

## convert to NewHybrids input file
sim2newhybrids(hybrids,filename="newhybridsinput.txt")
```

`sim2popgene`	*Converting hybridsim object to a PopGene input file*

Description

The function `sim2popgene` converts a hybridsim object into a PopGene input file.

Usage

```
sim2popgene(x,filename)
```

Arguments

x	a `hybridsim` object
filename	a character string indicating the name of the output file

Value

a PopGene input file

Author(s)

F. Balao `<fbalao@us.es>`, J.L. García-Castaño

References

Yeh, F.C., R.C. Yang, T.B.J. Boyle, Z.H. Ye & J.X. Mao (1997) Popgene, the User-Friendly Share-warefor Population Genetic Analysis. Molecular Biology and Biotechnology Centre, University of Alberta, Canada (program available from: `http://www.ualberta.ca/~fyeh/`).

See Also

`sim2arlequin sim2bayescan sim2introgress sim2newhybrids sim2structure`

Examples

```
## simulate F1 hybrids
F1hybrids<-hybridsim(Nmarker=100,Na=100,Nb=100,Nf1=30, type="neutral", hybrid="F1")

## convert to genepop input file
sim2popgene(F1hybrids,filename="F1hybrids_Popgene.txt")
```

`sim2structure`	*Convert a hybridsim object to a STRUCTURE input file*

Description

The function `sim2structure` converts a hybridsim object into a Structure 2.3 input file.

Usage

```
sim2structure(x,filename)
```

Arguments

`x`	a `hybridsim` object
`filename`	a character string indicating the name of the output file

Value

a Structure input file

Author(s)

F. Balao <fbalao@us.es>, J.L. García-Castaño

References

Falush, D., M. Stephens, J.K. Pritchard. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles.*Molecular Ecology Notes* **7**: 574-578.

Pritchard, J.K., M. Stephens, P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155**: 945-959.

See Also

`sim2arlequin sim2bayescan sim2introgress sim2newhybrids sim2popgene`

Examples

```
## simulate F1 hybrids
F1hybrids<-hybridsim(Nmarker=30,Na=30,Nb=100,Nf1=30, type="neutral", hybrid="F1")

## convert to STRUCTURE input file
sim2structure(F1hybrids,filename="F1hybrids_Structure.txt")
```