Package 'delmap'

November 14, 2012

Type Package

Title DELETION MAPPING

Version 1.0

Date 2012-11-12

Author Andrew W. George

Maintainer Andrew George <andrew.george@csiro.au>

Description

Deletion mapping is a statistical technique for ordering genetic markers. Unlike linkage mapping that is reliant upon recombinations to order loci, deletion mapping depends upon genomic deletions. These genomic deletions are introduced artificially. This package visualizes and analyzes deletion data on inbreds. The data can be read in from file or simulated.

License GPL (>= 2.0)

Depends R (>= 2.15.0), TSP, seriation

R topics documented:

`delmap-package` *Deletion Mapping*

Description

A package for using deletion data to fine map genetic marker loci. It is assumed the marker loci are (almost) completely linked and unable to be ordered through linkage. It is also assumed that a plant will contain, at most, a single deletion that spans multiple marker loci. The deletions are allowed to vary in size and there may be unobserved marker data. The package contains functions to find the best ordering of the marker loci, to visualize the deletion map, and to impute missing marker genotypes.

Details

Package:	delmap
Type:	Package
Version:	1.0
Date:	2012-11-12
License:	GPL (>=2)

~~ An overview of how to use the package, including the most important ~~ ~~ functions ~~

Author(s)

Author: Andrew W. George Maintainer: Andrew W. George <andrew.george@csiro.au>

Examples

```
S <- SimDeletions(numlines = 500, plines = 0.2, nummarkers = 20,
    Enumdel = 8, p.missing = 0.02,  seed = NULL)
```

---

print.delmap.data *Print Deletion Data*

Description

'print' prints the deletion data.

Usage

```
## S3 method for class 'delmap.data'
print(x, ...)
```

Arguments

x	an object of class 'delmap.dat' that has either been read in from file or simulated.
...	further arguments passed to or from other methods.

Details

'print.delmap.data' is a customized print function for printing the contents of the 'delmap.data' object. The first five lines of data, plant names, and marker names are displayed.

Author(s)

Andrew W. George

See Also

summary

Examples

```
# generate deletion data
S <-  SimDeletions( numlines = 500, plines = 0.2, nummarkers = 20,
    Enumdel = 8, p.missing = 0.02,  seed = NULL)

# print contents of data where all genotypes are observed
print(S$nomissing)

#print contents of data where there are missing genotypes
print(S$missing)
```

`ReadData` *Read Deletion Data*

Description

Reads a file in table format and creates a delmap.data object from it, with plants corresponding to rows and genetic markers corresponding to columns in the file

Usage

```
ReadData(datafile, line.names = FALSE, na.strings = "NA", marker.names = FALSE,
```

Arguments

`datafile`	the name of the file which the data are to be read from. Each row of the table appears as one line of the file. If it does not contain an absolute path, the file name is relative to the current working directory, 'getwd()'.
`line.names`	a logical value indicating whether the file contains plant names as the first column.
`na.strings`	a character vector of strings which are to be interpreted as 'NA' values.
`marker.names`	a logical value indicating whether the file contains the names of the marker loci as its first line.
`csv`	a logical value indicating if the file is comma seperated.

Details

Deletion data are read into R with this function. The data can be integer or character where any integer/character value is valid. The integer/character value with the lowest frequency is assumed to denote a deletion. The values are converted into 0/1 data where 1 denotes a deletion. If row/column names are not specified in this file, plant names L1, L2, ... and marker names M1, M2, ... will be assigned to the rows and columns, respectively.

Value

An object of class delmap.data. It contains the data in matrix form that has 0/1 elements. It also contains the plant names and marker loci names.

Author(s)

Andrew W. George

See Also

See Also SimDeletions

Examples

```
## Not run:
# read in deletion data
deldat <- ReadData(datafile="./deldata.txt", marker.names=TRUE, line.names=TRUE)

## End(Not run)
```

SimDeletions *Simulation*

Description

'SimDeletions' returns a matrix of genotype data where the rows correspond to plants and the columns correspond to tightly linked marker loci. A column (i.e., marker locus) can contain genotypes 0/1/NA where 0 is no deletion, 1 is a deletion, and NA is missing.

Usage

```
SimDeletions(numlines = NULL, plines = 0.5, nummarkers = NULL, labels = NULL, En
```

Arguments

numlines	number of inbred lines/plants.
plines	the probability of a plant carrying a deletion.
nummarkers	the number of marker loci spanning the genomic region of interest.
labels	a character vector containing the names of the marker loci.
Enumdel	the average length (in terms of number of marker loci) for a deletion.
p.missing	the probability of a genotype being unobserved.
seed	an integer number that if set, will produce reproducable replicates.

Details

Data are generated assuming the marker loci are in the true order and that a plant can only contain a single deletion (but that deletion may span several marker loci). The length of a deletion is random and follows a poisson distribution with mean Enumdel. If p.missing is set, then the data will contain missing genotypes (coded NA).

Value

a list is returned that contains only a single element if p.missing is NULL or two elements if p.missing is set. The elements are:

nomissing	is a matrix of genotype data 0/1 where 0 is no deletion and 1 is a deletion. The number of rows corresponds to the number of plants in the simulated population and the number of columns is the number of genetic marker loci. This will be the only element in output list if p.missing is NULL.
missing	is a matrix of genotype data 0/1 that is equivalent to "nomissing" but where marker genotypes have been replaced, randomly, and set to NA.

Author(s)

Andrew W. George

See Also

See Also `plot`

Examples

```
# To generate a inbred population of 500 plants
# where genotype data are collected on 20 marker loci,
# the probability of a plant having a genomic deletion is 0.2,
# the average length of a deletion is 8 marker loci, and
# the probability of a marker genotype being missing is 0.02,
# use:
S <- SimDeletions( numlines = 500, plines = 0.2, nummarkers = 20,
    Enumdel = 8, p.missing = 0.02,  seed = NULL)

# print contents of matrix
print(S$nomissing)

# print contents of matrix
print(S$missing)
```

Index