# Package 'statgenIBD'

February 2, 2024

**Type** Package

**Title** Calculation of IBD Probabilities

**Description** For biparental, three and four-way crosses Identity by Descent
(IBD) probabilities can be calculated using Hidden Markov Models and
inheritance vectors following Lander and Green
(<https://www.jstor.org/stable/29713>) and Huang
(<doi:10.1073/pnas.1100465108>). One of a series of statistical genetic
packages for streamlining the analysis of typical plant breeding experiments
developed by Biometris.

**Version** 1.0.7

**Date** 2024-02-02

**License** GPL

**Encoding** UTF-8

**Depends** R (>= 3.6)

**Imports** data.table, ggplot2, Matrix, rlang, Rcpp, statgenGWAS (>=
1.0.9), stringi

**LinkingTo** Rcpp, RcppArmadillo

**RoxygenNote** 7.3.1

**Suggests** rmarkdown, knitr, tinytest

**VignetteBuilder** knitr

**URL** <https://biometris.github.io/statgenIBD/index.html>,
<https://github.com/Biometris/statgenIBD/>

**BugReports** <https://github.com/Biometris/statgenIBD/issues>

**NeedsCompilation** yes

**Author** Martin Boer [aut] (<https://orcid.org/0000-0002-1879-4588>),
Bart-Jan van Rossum [aut, cre]
(<https://orcid.org/0000-0002-8673-2514>),
Wenhao Li [ctb] (<https://orcid.org/0000-0001-5719-5775>),
Johannes Kruisselbrink [ctb] (<https://orcid.org/0000-0003-1673-5725>)

**Maintainer** Bart-Jan van Rossum <bart-jan.vanrossum@wur.nl>

## R topics documented:

---

c.IBDprob                           *Concatenate function for objects of class IBDprob*

---

### Description

Concatenates objects of class IBDprob. All objects that are concatenated should have the same population type and the same map. The function is mainly meant for combining information for multiple crosses with overlapping parents.

### Usage

```
## S3 method for class 'IBDprob'
c(...)
```

### Arguments

...                  Objects of class IBDprob.

### Value

An object of class IBDprob containing data for all concatenated objects.

### Examples

```
## Compute IBD probabilties for AxB.
AB <- calcIBD(popType = "F4DH",
              markerFile = system.file("extdata/multipop", "AxB.txt",
                                  package = "statgenIBD"),
              mapFile = system.file("extdata/multipop", "mapfile.txt",
                                  package = "statgenIBD"))
## Compute IBD probabilties for Axc.
```

```
AC <- calcIBD(popType = "F4DH",
              markerFile = system.file("extdata/multipop", "AxC.txt",
                                        package = "statgenIBD"),
              mapFile = system.file("extdata/multipop", "mapfile.txt",
                                     package = "statgenIBD"))

## Combine results.
ABC <- c(AB, AC)

## Check summary.
summary(ABC)
```

---

calcIBD                        *Calculate IBD probabilities*

---

### Description

Calculate IBD probabilities for different types of populations.

### Usage

```
calcIBD(
  popType,
  markerFile,
  mapFile,
  evalPos = NULL,
  evalDist = NULL,
  grid = TRUE,
  verbose = FALSE
)
```

### Arguments

| | |
|---|---|
| popType | A character string indicating the type of population. One of DH, Fx, FxDH, BCx, BCxDH, BC1Sx, BC1SxDH, C3, C3DH, C3Sx, C3SxDH, C4, C4DH, C4Sx, C4SxDH (see Details). |
| markerFile | A character string indicating the location of the file with genotypic information for the population. The file should be in tab-delimited format with a header containing marker names. |
| mapFile | A character string indicating the location of the map file for the population. The file should be in tab-delimited format. It should consist of exactly three columns, marker, chromosome and position. There should be no header. The positions in the file should be in centimorgan. |
| evalPos | A data.frame with evaluation positions to which the calculations should be limited. |

| | |
|---|---|
| evalDist | An optional numerical value indicating the maximum distance for marker positions. Extra markers will be added based on the value of `grid`. |
| grid | Should the extra markers that are added to assure the a maximum distince of `evalDist` be on a grid (TRUE) or in between marker existing marker positions (FALSE). |
| verbose | Should messages indicating the progress of the process be printed? |

## Details

IBD probabilities can be calculated for many different types of populations. In the following table all supported populations are listed. Note that the value of x in the population types is variable, with its maximum value depicted in the last column.

| Population type | Cross | Description | max. x |
|---|---|---|---|
| DH | biparental | doubled haploid population | |
| Fx | biparental | Fx population (F1, followed by x-1 generations of selfing) | 8 |
| FxDH | biparental | Fx, followed by DH generation | 8 |
| BCx | biparental | backcross, second parent is recurrent parent | 9 |
| BCxDH | biparental | BCx, followed by DH generation | 9 |
| BC1Sx | biparental | BC1, followed by x generations of selfing | 7 |
| BC1SxDH | biparental | BC1, followed by x generations of selfing and DH | 6 |
| C3 | three-way | three way cross: (AxB) x C | |
| C3DH | three-way | C3, followed by DH generation | |
| C3Sx | three-way | C3, followed by x generations of selfing | 7 |
| C3SxDH | three-way | C3, followed by x generations of selfing and DH generation | 6 |
| C4 | four-way | four-way cross: (AxB) x (CxD) | |
| C4DH | four-way | C4, followed by DH generation | |
| C4Sx | four-way | C4, followed by x generations of selfing | 6 |
| C4SxDH | four-way | C4, followed by x generations of selfing and DH generation | 6 |

## Value

An object of class IBDprob, a `list` with five elements,

**map** a `data.frame` with chromosome and position of the markers.

**markers** a 3-dimensional `array` of IBD probabilities with genotypes, markers and parents as array dimensions.

**parents** the parents.

**popType** the population type.

## Examples

```
## Compute IBD probabilities for Steptoe Morex.
SxMIBD <- calcIBD(popType = "DH",
                  markerFile = system.file("extdata/SxM", "SxM_geno.txt",
                                           package = "statgenIBD"),
                  mapFile = system.file("extdata/SxM", "SxM_map.txt",
```

```
                                           package = "statgenIBD"))

## Check summary.
summary(SxMIBD)

## Compute IBD probabilities for Steptoe Morex.
## Add extra evaluation positions in between exiting marker positions
## to assure evaluation positions are at most 5 cM apart.
SxMIBD_Ext <- calcIBD(popType = "DH",
                      markerFile = system.file("extdata/SxM", "SxM_geno.txt",
                                               package = "statgenIBD"),
                      mapFile = system.file("extdata/SxM", "SxM_map.txt",
                                            package = "statgenIBD"),
                      evalDist = 5)

## Check summary.
summary(SxMIBD_Ext)
```

---

getProbs                           *Extract Probabilities for markers*

---

### Description

Extract IBD probabilities for one or more markers from an object of class IBDprob.

### Usage

```
getProbs(IBDprob, markers, sumProbs = FALSE)
```

### Arguments

| | |
|---|---|
| IBDprob | An object of class IBDprob. |
| markers | A character vector of markers that should be extracted. |
| sumProbs | Should the probabilities by summed per parent. If TRUE the probability for e.g. parent A in a cross with parent B will be calculated as pA + 0.5 * pAB. If FALSE both pA and pAB will be output without further calculations. |

### Value

A data.frame with IBD probabilities for the extracted markers in the column and genotypes in the rows.

## Examples

```
## Compute IBD probabilities for Steptoe Morex.
SxMIBD <- calcIBD(popType = "DH",
                  markerFile = system.file("extdata/SxM", "SxM_geno.txt",
                                           package = "statgenIBD"),
                  mapFile = system.file("extdata/SxM", "SxM_map.txt",
                                        package = "statgenIBD"))

## Get probabilities for a single marker.
probOne <- getProbs(IBDprob = SxMIBD,
                    markers = "plc")
head(probOne)

## Get probabilities for a multiple markers.
probMult <- getProbs(IBDprob = SxMIBD,
                     markers = c("plc", "tuba1"))
head(probMult)
```

---

plot.IBDprob                *Plot function for objects of class IBDprob*

---

## Description

Creates a plot for an object of class IBDprob. Six types of plot can be made:

- singleGeno A plot for a single genotype showing the IBD probabilities for all parents across the genome.
- allGeno A plot showing for all genotypes the IBD probabilities of the parent with the highest probability per marker.
- pedigree A plot showing the structure of the pedigree of the population.
- map A plot of the genetic map showing the length of the chromosomes and the positions of the markers.
- meanProbs A plot showing the coverage of each parent across the population.
- totalCoverage A plot showing the total coverage of each parent.

## Usage

```
## S3 method for class 'IBDprob'
plot(
  x,
  ...,
  plotType = c("singleGeno", "allGeno", "pedigree", "map", "meanProbs", "totalCoverage"),
  genotype,
  chr = NULL,
  title = NULL,
  output = TRUE
)
```

## Arguments

| | |
|---|---|
| x | An object of class `IBDprob`. |
| ... | Further arguments. Unused. |
| plotType | A character string indicating the type of plot that should be made. |
| genotype | A character string indicating the genotype for which the plot should be made. Only for `plotType = "singleGeno"`. |
| chr | A character vector indicating the chromosomes to which the coverage should be restricted. Only for `plotType = "meanProbs"` and `plotType = "totalCoverage"`. If `NULL` all chromosomes are included. |
| title | A character string, the title of the plot. |
| output | Should the plot be output to the current device? If `FALSE`, only a ggplot object is invisibly returned. |

## Value

A ggplot object is invisibly returned.

## Examples

```
## Not run:
## Compute IBD probabilities for Steptoe Morex.
## Add extra evaluation positions in between exiting marker positions
## to assure evaluation positions are at most 2 cM apart.
SxMIBD_Ext <- calcIBD(popType = "DH",
                      markerFile = system.file("extdata/SxM", "SxM_geno.txt",
                                               package = "statgenIBD"),
                      mapFile = system.file("extdata/SxM", "SxM_map.txt",
                                            package = "statgenIBD"),
                      evalDist = 2)

## Plot results for genotype dh005.
plot(SxMIBD_Ext,
     plotType = "singleGeno",
     genotype = "dh005")

## Plot results for all genotypes.
plot(SxMIBD_Ext,
     plotType = "allGeno")

## Plot structure of the pedigree.
plot(SxMIBD_Ext,
     plotType = "pedigree")

## Plot genetic map.
plot(SxMIBD_Ext,
     plotType = "map")

## Plot coverage across population.
plot(SxMIBD_Ext,
```

```
    plotType = "meanProbs")

## Plot total coverage.
plot(SxMIBD_Ext,
     plotType = "totalCoverage")

## End(Not run)
```

---

readIBDs                      *Read IBD probabilities from file*

---

### Description

Reads IBD probabilities from a plain text, tab-delimited .txt or .ibd file. Information about the file format can be found in the vignette ( `vignette("IBDFileFormat", package = "statgenIBD")`). A data.frame with the map must be specified as well.

### Usage

```
readIBDs(infile, map)
```

### Arguments

| | |
|---|---|
| infile | A character string specifying the path of the input file. Compressed files with extension ".gz" or ".bz2" are supported as well. |
| map | A data.frame with columns chr for chromosome and pos for position. Positions should be in centimorgan (cM). They should not be cumulative over the chromosomes. Other columns are ignored. Marker names should be in the row names. These should match the marker names in the input file. |

### Value

An object of class IBDprob.

### Examples

```
## Read map for Steptoe Morex.
SxMmap <- read.delim(system.file("extdata/SxM", "SxM_map.txt",
                               package = "statgenIBD"), header = FALSE)
rownames(SxMmap) <- SxMmap$V1
SxMmap <- SxMmap[, -1]
colnames(SxMmap) <- c("chr", "pos")


## Read IBD probabilities for Steptoe Morex.
SxMIBD <- readIBDs(system.file("extdata/SxM", "SxM_IBDs.txt",
                             package = "statgenIBD"),
                 map = SxMmap)
```

```
## Print summary.
summary(SxMIBD)
```

---

| readRABBIT | *Read IBD probabilities* |
|---|---|

---

### Description

Read a file with IBD probabilities computed by the RABBIT software package. It is possible to additionally read the pedigree file that is also used by RABBIT. Reading this file allows for plotting the pedigree.

### Usage

```
readRABBIT(infile, pedFile = NULL)
```

### Arguments

infile          A character string, a link to a .csv file with IBD probabilities. Compressed .csv files with extension ".gz" or ".bz2" are supported as well.

pedFile        A character string, a link to a .csv file with pedigree information as used by RABBIT as input. Compressed .csv files with extension ".gz" or ".bz2" are supported as well.

### Value

An `IBDprob` object with map and markers corresponding to the imported information in the imported .csv file.

### References

Zheng, Chaozhi, Martin P Boer, and Fred A Van Eeuwijk. "Recursive Algorithms for Modeling Genomic Ancestral Origins in a Fixed Pedigree." G3 Genes|Genomes|Genetics 8 (10): 3231–45. https://doi.org/10.1534/G3.118.200340.

### Examples

```
## Not run:
## Read RABBIT data for barley.
genoFile <- system.file("extdata/barley", "barley_magicReconstruct.zip",
                        package = "statgenIBD")
barleyIBD <- readRABBIT(unzip(genoFile, exdir = tempdir()))

## End(Not run)
```

---

summary.IBDprob            *Summary function for objects of class IBDprob*

---

### Description

Prints a short summary for objects of class IBDprob. The summary consists of the population type, number of evaluation points, number of individuals and names of the parents in the object.

### Usage

```
## S3 method for class 'IBDprob'
summary(object, ...)
```

### Arguments

object            An object of class IBDprob.

...               Not used.

### Value

No return value, a summary is printed.

### Examples

```
## Compute IBD probabilities for Steptoe Morex.
SxMIBD <- calcIBD(popType = "DH",
                  markerFile = system.file("extdata/SxM", "SxM_geno.txt",
                                                package = "statgenIBD"),
                  mapFile = system.file("extdata/SxM", "SxM_map.txt",
                                                package = "statgenIBD"))

## Print summary
summary(SxMIBD)
```

---

writeFlapjack              *Write to Flapjack format*

---

### Description

Export the results of an IBD calculation to Flapjack format so it can be visualized there.

## Usage

```
writeFlapjack(
  IBDprob,
  outFileMap = "ibd_map.txt",
  outFileGeno = "ibd_geno.txt"
)
```

## Arguments

| | |
|---|---|
| IBDprob | An object of class IBDprob. |
| outFileMap | A character string, the full path to the output map file. |
| outFileGeno | A character string, the full path to the output genotype file. |

## Value

No output. Output files are written to a folder.

## Examples

```
## Compute IBD probabilities for Steptoe Morex.
SxMIBD <- calcIBD(popType = "DH",
                  markerFile = system.file("extdata/SxM", "SxM_geno.txt",
                                           package = "statgenIBD"),
                  mapFile = system.file("extdata/SxM", "SxM_map.txt",
                                        package = "statgenIBD"))

## Write output in Flapjack format to temporary files.
writeFlapjack(SxMIBD,
              outFileMap = tempfile(fileext = ".txt"),
              outFileGeno = tempfile(fileext = ".txt"))
```

---

| writeIBDs | *Write IBD probabilities to file.* |
|---|---|

---

## Description

Writes IBD probabilities to a plain text, tab-delimited .txt or .ibd file. Information about the file format can be found in the vignette ( vignette("IBDFileFormat", package = "statgenIBD")).

## Usage

```
writeIBDs(IBDprob, outFile, decimals = 6, minProb = 0, compress = FALSE)
```

**Arguments**

| | |
|---|---|
| IBDprob | An object of class IBDprob containing the IBD probabilities. |
| outFile | A character string specifying the path of the output file. |
| decimals | An integer value specifying the number of decimals to include in writing the output file. |
| minProb | A numerical value between zero and 1 / number of parents, specifying the minimum probability cutoff value. Probabilities below this cutoff are set to zero and other probabilities are rescaled to make sure that the probabilities sum up to one. |
| compress | Should the output be compressed to .gz format? |

**Value**

No output. The output file is created as a result of calling this function.

**Examples**

```
## Compute IBD probabilities for Steptoe Morex.
SxMIBD <- calcIBD(popType = "DH",
                  markerFile = system.file("extdata/SxM", "SxM_geno.txt",
                                            package = "statgenIBD"),
                  mapFile = system.file("extdata/SxM", "SxM_map.txt",
                                         package = "statgenIBD"))

## Write IBDs to temporary files.
writeIBDs(IBDprob = SxMIBD, outFile = tempfile(fileext = ".txt"))

## Write IBDs to file, set values <0.05 to zero and only print 3 decimals.
writeIBDs(IBDprob = SxMIBD, outFile = tempfile(fileext = ".txt"),
          decimals = 3, minProb = 0.05)
```

# Index