Package ‘snpStatsWriter’

February 20, 2015

Type Package

Title Flexible writing of snpStats objects to flat files

Version 1.5-6

Date 2013-12-05

Author Chris Wallace

Maintainer Chris Wallace <`chris.wallace@cimr.cam.ac.uk`>

Description Write snpStats objects to disk in formats suitable for reading by snphap, phase, mach, IMPUTE, beagle, and (almost) anything else that expects a rectangular format.

Depends snpStats, colorspace

License GPL

LazyLoad yes

Collate 'snpStatsWriter-package.R' 'write.R'

OS_type unix

Suggests knitr

VignetteBuilder knitr

NeedsCompilation yes

Repository CRAN

Date/Publication 2013-12-05 17:28:32

R topics documented:

`snpStatsWriter-package` *snpStatsWriter*

Description

Write snpMatrix objects to file in other formats.

Details

Package:	snpStatsWriter
Type:	Package
Version:	1.0
Date:	2012-10-29
License:	GPL
LazyLoad:	yes

Author(s)

Maintainer: Chris Wallace <chris.wallace@cimr.cam.ac.uk>

References

David Clayton and Hin-Tak Leung (2007). An R package for analysis of whole-genome association studies. Hum Hered 64:45-51.

See Also

`snpStats`

Examples

```
data(testdata,package="snpStats")
A.small <- Autosomes[1:6,1:10]
f <- tempfile()
## write in suitable format for snphap
nsnps <- ncol(A.small)
write.simple(A.small, a1=rep("1",nsnps), a2=rep("2",nsnps), gsep=" ",
             nullallele='0', file=f,
                write.sampleid=FALSE)
unlink(f)
```

`write.beagle` *Write a snpStats object in beagle format*

Description

see `write.simple` for general information

Usage

```
write.beagle(X, a1, a2, bp, trait = NULL, gfile, mfile)
```

Arguments

`trait`	disease trait (0=missing, 1=control, 2=case)
`gfile,mfile`	`gfile`=genotype file, `pedfile`=pedigree file
`X`	SnpMatrix object
`a1`	vector of first allele at each SNP
`a2`	vector of second allele at each SNP
`bp`	vector of base pair positions for each SNP

Value

No return value, but has the side effect of writing specified output files.

Author(s)

Chris Wallace

Examples

```
data(testdata,package="snpStats")
A.small <- Autosomes[1:6,1:10]
gf <- tempfile()
mf <- tempfile()

## write in suitable format for beagle
nsnps <- ncol(A.small)
write.beagle(A.small, a1=rep("1",nsnps), a2=rep("2",nsnps), bp=1:nsnps, gfile=gf, mfile=mf)
unlink(gf)
unlink(mf)
```

`write.impute` *Write a snpStats object in IMPUTE format*

Description

see `write.simple` for general information

Usage

```
write.impute(X, a1, a2, bp, pedfile, snp.id = NULL)
```

Arguments

`pedfile`	Output file name.
`snp.id`	vector of snp ids
`X`	SnpMatrix object
`a1`	vector of first allele at each SNP
`a2`	vector of second allele at each SNP
`bp`	vector of base pair positions for each SNP

Value

No return value, but has the side effect of writing specified output files.

Author(s)

Chris Wallace

Examples

```
data(testdata,package="snpStats")
A.small <- Autosomes[1:6,1:10]
pf <- tempfile()

## write in suitable format for IMPUTE
nsnps <- ncol(A.small)
write.impute(A.small, a1=rep("1",nsnps), a2=rep("2",nsnps), bp=1:nsnps, pedfile=pf)
unlink(pf)
```

`write.mach` *Write a snpStats object in mach format*

Description

see `write.simple` for general information

Usage

```
write.mach(X, a1, a2, pedfile, mfile,
  pedigree = rownames(X), member = rep(1, nrow(X)),
  father = rep(0, nrow(X)), mother = rep(0, nrow(X)),
  sex = rep("M", nrow(X)), snp.names = colnames(X))
```

Arguments

`pedfile`	Output pedigree file name.
`mfile`	Output marker file name.
`pedigree`	Optional pedigree/member/father/mother/sex indentifier vectors, same order as rows in snpStats object. If missing, pedigree is set to rownames(X) and the others default to unrelated males
`member`	See pedigree
`father`	See pedigree
`mother`	See pedigree
`sex`	See pedigree
`snp.names`	optional SNP names to include in the marker map file. Defaults to colnames(X).
`X`	SnpMatrix object
`a1`	vector of first allele at each SNP
`a2`	vector of second allele at each SNP

Value

No return value, but has the side effect of writing specified output files.

Author(s)

Chris Wallace

Examples

```
data(testdata,package="snpStats")
A.small <- Autosomes[1:6,1:10]
pf <- tempfile()
mf <- tempfile()

## write in suitable format for MACH
nsnps <- ncol(A.small)
write.mach(A.small, a1=rep("1",nsnps), a2=rep("2",nsnps), pedfile=pf, mfile=mf)
unlink(pf)
unlink(mf)
```

`write.phase` *Write a snpStats object in PHASE/FastPHASE format*

Description

see `write.simple` for general information

Usage

```
write.phase(X, a1 = rep(1, ncol(X)),
  a2 = rep(2, ncol(X)), bp = NULL, file)
```

Arguments

`file`	Output file name.
`X`	SnpMatrix object
`a1`	vector of first allele at each SNP
`a2`	vector of second allele at each SNP
`bp`	vector of base pair positions for each SNP

Value

No return value, but has the side effect of writing specified output files.

Author(s)

Chris Wallace

Examples

```
data(testdata,package="snpStats")
A.small <- Autosomes[1:6,1:10]
f <- tempfile()

## write in suitable format for PHASE
nsnps <- ncol(A.small)
write.phase(A.small, file=f)
unlink(f)
```

`write.sbams` *write.sbams*

Description

write an sbams format file

Usage

```
write.sbams(X, response, file)
```

Arguments

`response`	vector or matrix of response variables. rows index subjects, columns index variables
`X`	SnpMatrix object
`file`	Output file name.

Details

sbams is software from Xiaoquan Wen at https://github.com/xqwen/sbams

Value

No return value, but has the side effect of writing specified output file.

Author(s)

Chris Wallace

Examples

```
data(testdata,package="snpStats")
A.small <- Autosomes[1:6,1:10]
R <- matrix(rnorm(12),ncol=2)
colnames(R) <- c("var1","var2")
f <- tempfile()
```

```
## write in suitable format for sbams
write.sbams(X=A.small, response=R, file=f)
unlink(f)
```

`write.simple` *Fast and flexible writing of snpStats objects to flat files*

Description

Different genetics phasing and analysis programs (beagle, mach, impute, snptest, phase/fastPhase, snphap, etc) have different requirements for input files. These functions aim to make creating these files from a SnpMatrix object straightfoward.

Usage

```
write.simple(X, a1, a2, file, fsep = "\t", gsep = "",
  nullallele = "N", write.header = TRUE,
  transpose = FALSE, write.sampleid = TRUE, bp = NULL,
  num.coding = FALSE)
```

Arguments

`X`	SnpMatrix object
`a1`	vector of first allele at each SNP
`a2`	vector of second allele at each SNP
`bp`	vector of base pair positions for each SNP
`fsep,gsep`	Field and genotype separators.
`nullallele`	Character to use for missing alleles
`file`	Output file name.
`write.header`	Write a header line
`transpose`	Output SNPs as rows, samples as columns if `TRUE`. The default is samples as rows, SNPs as columns, as represented internally by snpStats/SnpMatrix.
`write.sampleid`	Output sample ids
`num.coding`	Use alleles 1 and 2 instead of supplying allele vectors.

Details

It's written in C, so should be reasonably fast even for large datasets.

`write.simple` is the most flexible function. It should be able to write most rectangular based formats.

Additional functions are available tailored to software that require a bit more than a rectangular format: `write.beagle`, `write.impute`, `write.mach`, `write.phase`.

Value

No return value, but has the side effect of writing specified output files.

Warning

Any uncertain genotypes (stored by snpStats as raw codes 4 to 253) are output as missing.

The functions use "\n" as an end of line character, unless `.Platform$OS.type == "windows"`, when eol is "\r\n". I only have access to linux machines for testing.

I have tested these functions with my own data, but it is always possible that your data may contain quirks mine don't, or that input formats could change for any program mentioned here. Please do have a quick check on a small subset of data (eg, as in the example below), that the output for your exact combination of options looks sensible and matches the specified input format.

Note

This has been tested with `SnpMatrix` objects from the package `snpStats` but should also work with `snp.matrix` objects from the package `snpMatrix`.

Author(s)

Chris Wallace

References

David Clayton (2012). snpStats: SnpMatrix and XSnpMatrix classes and methods. R package version 1.6.0. http://www-gene.cimr.cam.ac.uk/clayton

phase/fastPhase: http://stephenslab.uchicago.edu/software.html

beagle: http://faculty.washington.edu/browning/beagle/beagle.html

IMPUTE: http://mathgen.stats.ox.ac.uk/impute/impute_v2.html

MACH: http://www.sph.umich.edu/csg/abecasis/MACH

snphap: https://www-gene.cimr.cam.ac.uk/staff/clayton/software/snphap.txt

Examples

```
data(testdata,package="snpStats")
A.small <- Autosomes[1:6,1:10]
f <- tempfile()
## write in suitable format for snphap
nsnps <- ncol(A.small)
write.simple(A.small, a1=rep("1",nsnps), a2=rep("2",nsnps), gsep=" ",
             nullallele='0', file=f,
                write.sampleid=FALSE)
unlink(f)
```

`write.snphap` *Write SNPHAP files*

Description

Simple wrapper to write.simple to write files in SNPHAP format

Usage

```
write.snphap(X, a1 = NULL, a2 = NULL, file)
```

Arguments

`X`	SnpMatrix object
`a1`	vector of first allele at each SNP
`a2`	vector of second allele at each SNP
`file`	Output file name.

Details

If not allele codes are given, a1 and a2 will be set to 1 and 2 for all SNPs

Value

No return value, but has the side effect of writing specified output file.

Author(s)

Chris Wallace

Examples

```
data(testdata,package="snpStats")
A.small <- Autosomes[1:6,1:10]
f <- tempfile()
## write in suitable format for snphap
write.snphap(A.small, file=f)
unlink(f)
```

Index