# Package 'msap'

February 2, 2017

**Version** 1.1.8

**Date** 2014-05-07

**Title** Statistical analysis for Methylation-sensitive Amplification
Polymorphism data

**Author** Andres Perez-Figueroa [aut, cre]

**Maintainer** Andres Perez-Figueroa <anpefi@uvigo.es>

**Depends** R (>= 2.14.0)

**Imports** ade4, ape

**Description**
Statistical Analyses of Methylation-sensitive Amplification Polymorphism (MSAP) assays.

**License** GPL (>= 2)

**URL** http://msap.r-forge.r-project.org/

**BugReports**
https://r-forge.r-project.org/tracker/?func=browse&group_id=1422&atid=5292

**Repository** CRAN

**Repository/R-Forge/Project** msap

**Repository/R-Forge/Revision** 45

**Repository/R-Forge/DateTimeStamp** 2014-05-08 14:21:58

**Date/Publication** 2014-05-09 18:30:28

**NeedsCompilation** no

## R topics documented:

1

| | |
|---|---|
| rmsap-package | *Statistical Analyses for Methylation Sensitive Amplification Polymorphism assays* |

## Description

This package provides tools for Statistical Analyses of Methylation Sensitive Amplification Polymorphism (MSAP) assays

## Details

| | |
|---|---|
| Package: | rmsap |
| Type: | Package |
| Version: | 0.01 |
| Date: | 2012-05-08 |
| License: | GPL-2 |
| Requires: | ade4, pegas |
| LazyLoad: | yes |

## Author(s)

Andres Perez-Figueroa (<anpefi@uvigo.es>)

| | |
|---|---|
| methStatusEval | *classify a locus as either MSL or NML.* |

## Description

Classify a locus as either 'methylation-susceptible locus' (MSL) or 'non-methylated locus' (NML). See Details.

## Usage

```
methStatusEval(x, error=0.05, uninformative=TRUE)
```

## Arguments

x           A list of mixed band patterns for a given fragment across all the samples. Values of 11, 10, 1 and 0 represent patterns HPA+/MSP+,HPA+/MSP-,HPA-/MSP+,HPA-/MSP-, respectively.

| error | Repeatibility value of MSAP assays. It provides a threshold to consider methylation events as genotyping errors. |
| uninformative | A logic value determining how to deal with HPA-/MSP- pattern. FALSE indicates that HPA-/MSP- stands for hypermethylated target (ignoring genetic mutation), and TRUE (default) stands for uninformative pattern (as this could be caused by genetic mutation or hypermethylation) |

### Details

Individual fragments (loci) are classified as 'methylation-susceptible loci' (MSL) or 'non-methylated loci' (NML), depending on whether the observed proportion of discordant HPA/MSP scores suggestive of methylation(i.e. number of individuals with contrastinf HPA/MSP scores for the fragment divided by the total number of individuals assayed) exceeded a user-defined threshold (0.05 by default).

### Author(s)

Andres Perez-Figueroa (<anpefi@uvigo.es>)

---

msap                                  *Diversity analysis of MSAP data*

---

### Description

It performs all the analysis of MSAP data in order to asses both epigenetic and genetic diversity.

### Usage

```
msap(datafile, name=datafile, no.bands = "u", nDec=4, meth=TRUE,
  rm.redundant=TRUE,   rm.monomorphic=TRUE, do.pcoa=TRUE,
  do.shannon=TRUE, do.amova=TRUE, do.pairwisePhiST=FALSE,
  do.cluster=TRUE, use.groups=NULL, do.mantel=FALSE, np.mantel=1000,
  loci.per.primer=NULL, error.rate.primer=NULL, uninformative=TRUE)
```

### Arguments

| datafile | String containing the url of the csv file with the data. Required. |
| name | a name for the dataset to be included in the output files. By default, the name of the given datafile is used. |
| no.bands | String to indicate how to deal with HPA-/MSP- pattern in MSL, with two possible values: 'h' for hypermethylated state and 'u' for uninformative state (default). |
| | With 'h' researcher assumes that HPA-/MSP- (no band for both isoschizomers) pattern represents full methylation of cytosines in the target (hypermethylation) ignoring the chance of genetic change in the target. This approach could be very useful when no much genetic differences are expected between samples |

(i.e. the same individuals are compared for different treatments). In this case HPA-/MSP- will compute as 'methylated' state in the binary matrix used for subsequent analysis.

With 'u' (default value), researcher considers that pattern as uninformative as could be caused by a missing target (mutation) as well as by hypermethylation. So, in this case, HPA-/MSP- will compute as missing values in the binary matrix used for subsequent analysis.

| | |
|---|---|
| nDec | number of digits of precision for floating point output. |
| meth | Logical value switching between MSAP ('TRUE') and standard AFLP ('FALSE') analysis. The difference lies in that for AFLP (meth=FALSE) the 'enzyme' column is ignored and every row in data represent an independent sample, without combination of data. |
| rm.redundant | Not implemented yet. |
| rm.monomorphic | Logical value switching between the removal ('TRUE', by default) of monomorphic fragments (defined as those with only one state or just one ocurrence of the second state across the whole dataset) after data transformation. |
| do.pcoa | Option switcher for doing a Principal Coordinate Analysis for variation between groups. TRUE by default. |
| do.shannon | Option switcher for Shannon's Diversity Index comparison between MSL and NML. |
| do.amova | Option switcher for doing an AMOVA for differentiation between groups. TRUE by default. |
| do.pairwisePhiST | |
| | Logical value switching between the calculation of the pairwise Phi_st between pairs of groups/populations ('TRUE' ) or skip it ('FALSE' by default). |
| do.cluster | Calculates and plots a Neighbour-Joining tree ('TRUE' by default) or skip it ('FALSE'). |
| use.groups | Gives the groups/populations/treatments of the datafile to be analysed. By default all groups are considered into de the analysis. To provide a subset of the groups a vector should be passed with the names of groups to be included. For example, in a datafile with 5 groups (Control, pop1, pop2, pop3 and pop4) we are interested only in Control and pops 1 and 3. Then, msap should be called with 'use.groups=c('Control','pop1','pop3')'. |
| do.mantel | Performs a Mantel test to obtain correlation between MSL and NML ('TRUE') or skip it ('FALSE' by default). |
| np.mantel | Gives the number of permutations for the above Mantel test (1000 by default) or skip it ('FALSE'). |
| loci.per.primer | |
| | Vector providing the number of loci/fragements obtained per primer combination. Fragment classsification is performed independenty for each primer combination. These fragment should be ordered in the datafilein the sameway as specified here. If this is nor provided (by default) then all fragments should be analyzed as they come from a single combination. For example, if there are three primer combinations with 135, 234 and 210 loci each, then msap should be called with 'loci.per.primer=c(135,234,210)' |

error.rate.primer

        Gives the repeatibility value of MSAP assays for each primer combination. It provides a threshold to consider methylation events as genotyping errors.

uninformative     Deprecated. This argument is keept for back compatibility with version 1.0.x. Contains the same information as no.bands.

## Details

This function is the main interface of the msap package. The only required argument is a string with the name (uri) of the data file to be analysed.

Data file should be a .csv file with markers as columns and two rows by sample, one for each isochizomer reaction. The first row should include the markers name/references. The first column should provide the label for the group where the sample is included, with the aim to make comparisons between different gruops. Second column is reserved for an arbitrary label (i.e. to name the sample). Third column should identify the isochizomer with 'HPA' or 'MSP'.

## Value

From version 1.1.2, msap returns a list with data useful for further analysis:

groups           A factor with the name of the group of every individual analysed

patterns         A list showing the MSAP patterns (11, 10, 01 and 00 coded as u, h, i, f) in all groups

transformed.MSL

        A data frame including the binary (1: unmmethylated, 2: methylated) values for those loci classified as MSL

transformed.NML

        A data frame including the binary (1: unmmethylated, 2: methylated) values for those loci classified as NML

DM.MSL         A distance matrix object between all individuals for those polymorphic loci classified as MSL

DM.NML         A distance matrix object between all individuals for those polymorphic loci classified as NML

DM.AFLP        A distance matrix object between all individuals when analysing AFLP data

## Author(s)

Andres Perez-Figueroa (<anpefi@uvigo.es>)

## See Also

[rmsap-package](rmsap-package)

## Examples

```
#Perform all default analysis
## Not run: msap("MyDataFile", name="MyAnalysis")
#Perform all default analysis for a dataset with 2 primer combinations
  #(200 and 180 fragments each)
## Not run: msap("MyDataFile", name="MyAnalysis", loci.per.primer=c(200, 180))
#Perform analysis assumes that HPA-/MSP- represents full methylation
## Not run: msap("MyDataFile", name="MyAnalysis", uninformative=FALSE)
#The same that above but skipping the PCoA results
## Not run: msap("MyDataFile", name="MyAnalysis", uninformative=FALSE, do.pcoa=FALSE)
#Using only some of the populations
## Not run: msap("MyDataFile", name="MyAnalysis", use.groups=c('Control','pop1','pop3'))
```

---

pcoa                                    *Principal Coordinate Analyses*

---

## Description

This function uses ade4's dudi.pco and s.class to obtain a PCoA and represent its results in a 2-D figure. It also saves a .csv file with the coordinates for all individuals and another .csv file with the eigenvalues for all axis.

## Usage

```
pcoa(DM, groups, inds, name, surname)
```

## Arguments

| | |
|---|---|
| DM | A Euclidean Distance Matrix |
| groups | A factor with the label for the group for every sample (row) in 'dataM' |
| inds | Label list for individuals in 'dataM' |
| name | Prefix for filename and title |
| surname | "MNL" or "MSL" |

## Author(s)

Andres Perez-Figueroa (<anpefi@uvigo.es>)

---

polymorphic                    *Checks polymorphic state*

---

### Description

Checks if a list of binary aleles is polymorphic (with at least two ocurrences of both states)

### Usage

```
polymorphic(col)
```

### Arguments

col                 A list of binary alleles

### Author(s)

Andres Perez-Figueroa (<anpefi@uvigo.es>)

---

repMet                    *Report methylation levels*

---

### Description

Prints a report of the relative number or different HPA-MSP band patterns in all groups

### Usage

```
repMet(dataM,groups,nDec)
```

### Arguments

dataM               A matrix of mixed band patterns, where each row is a different sample. Values of
                    11, 10, 1 and 0 represent patterns HPA+/MSP+,HPA+/MSP-,HPA-/MSP+,HPA-
                    /MSP-, respectively.

groups              A list with the label for the group for every sample (row) in 'dataM'

nDec                number of digits of precision for floating point output.

### Author(s)

Andres Perez-Figueroa (<anpefi@uvigo.es>)

---

shannon                          *Shannon Index of phenotypic diversity*

---

### Description

Calculate Shannon Index of phenotypic diversity given a list of binary alleles.

### Usage

```
shannon(x)
```

### Arguments

x                          A list of binary values or alleles.

### Details

The Shannon index of phenotypic diversity, S, derived from the Shannon-Weaver index (Shannon, 1948):

$$S = -\sum_{i=1}^{n} p_i \log_2 p_i$$

where $p_i$ is the frequency of the band presence at the ith marker within the population. This index gives more weight to the presence than to the absence of bands. This has no real biological support, although it might account for the occurrence of homoplasic absences of bands (Bonin et al., 2007)

### Value

The value of the Shannon Index of phenotypic diversity

### Author(s)

Andres Perez-Figueroa (`<anpefi@uvigo.es>`)

### References

Bonin, A., Ehrich, D. and Manel S. 2007. Statistical analysis of amplified fragment length polymorphism data: a toolbox for molecular ecologists and evolucionists. *Molecular Ecology* 16, 3737-3758.

Shannon, C. E. 1948. A mathematical theory of comunication. *Bell System Technical Journal* 27, 379-423.

# Index