Package ‘rtk’

June 13, 2020

Type Package
Title Rarefaction Tool Kit
Version 0.2.6.1
Date 2020-06-13
Author Paul Saary, Falk Hildebrand
Maintainer Paul Saary <rtk@paulsaary.de>
Description Rarefy data, calculate diversity and plot the results.
License GPL (>= 2)
Imports Rcpp (>= 0.12.3),methods
LinkingTo Rcpp
SystemRequirements C++11
Suggests testthat
NeedsCompilation yes
Repository CRAN
Date/Publication 2020-06-13 09:00:02 UTC

R topics documented:

  rtk-package .................................................. 2
  collectors.curve ............................................ 3
  get.diversity ................................................. 4
  plot ............................................................ 5
  rtk ............................................................ 7

Index 10
Description

Rarefy data, calculate diversity and plot the results.

Details

The DESCRIPTION file:

<table>
<thead>
<tr>
<th>Package</th>
<th>rtk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td>Package</td>
</tr>
<tr>
<td>Title</td>
<td>Rarefaction Tool Kit</td>
</tr>
<tr>
<td>Version</td>
<td>0.2.6.1</td>
</tr>
<tr>
<td>Date</td>
<td>2020-06-13</td>
</tr>
<tr>
<td>Author</td>
<td>Paul Saary, Falk Hildebrand</td>
</tr>
<tr>
<td>Maintainer</td>
<td>Paul Saary <a href="mailto:rtk@paulsaary.de">rtk@paulsaary.de</a></td>
</tr>
<tr>
<td>Description</td>
<td>Rarefy data, calculate diversity and plot the results.</td>
</tr>
<tr>
<td>License</td>
<td>GPL (&gt;= 2)</td>
</tr>
<tr>
<td>Imports</td>
<td>Rcpp (&gt;= 0.12.3),methods</td>
</tr>
<tr>
<td>LinkingTo</td>
<td>Rcpp</td>
</tr>
<tr>
<td>SystemRequirements</td>
<td>C++11</td>
</tr>
<tr>
<td>Suggests</td>
<td>testthat</td>
</tr>
</tbody>
</table>

Index of help topics:

collectors.curve    collectors.curve
get.diversity       get.diversity
plot                Plot rarefaction results
rtk                 Rarefy tables
rtk-package         Rarefaction Tool Kit

This package might be used to rarefy data and compute diversity measures. Rarefied tables can be returned to R and be further processed.

Author(s)

Paul Saary, Falk Hildebrand

Maintainer: Paul Saary <rtk@paulsaary.de>

References

collectors.curve

See Also

rtk, plot.rtk, collectors.curve

collectors.curve  collectors.curve

Description

Collectors curves visualize the richness gained by picking more samples.

Usage

collectors.curve(x, y = NULL, col = 1, times = 10, bin = 3, add = FALSE, ylim = NULL, xlim = NULL, doPlot = TRUE, rareD = NULL, cls = NULL, pch = 20, col2 = NULL, accumOrder = NULL, ...)

Arguments

x  
Input a rarefaction object with one matrix and one depth or dataframe/matrix or the output of collectors.curve itself

y  
secondary input matrix for comparative plots

col  
fill color of the boxplots (set to c(0) for no color)

times  
Number of times the sampeling of samples should be perfomed

bin  
Number of samples to be added each step. Usefull to adjust for a quick glance.

add  
add the plot to an existing plot?

ylim  
Limits for Y-scale

xlim  
Limits for X-scale

doPlot  
should this function plot the collectors curve, or just return an object that can be plotted later with this function?

rareD  
Depth to which rarefy the dataset using rtk

cls  
vector describing the class of each input sample

pch  
Plotting symbols

col2  
Color for the border of the boxplot, defaults to col

accumOrder  
accumulate successively within each class, given by cls in the order given in this vector. All classes in cls must be represented in this vector.

...  
Options passed to plot or boxplot

Details

The function collectors.curve can visualize the richness a dataset has, if sampels are picked at random. It can handle rareafction results as well as normal dataframes.
Author(s)
Falk Hildebrand, Paul Saary

References

See Also
Use `plot.rtk` for how to plot your results.

Examples
```r
require("rtk")
data <- matrix(sample(x = c(rep(0, 15000), rep(1:10, 100)), size = 10000, replace = TRUE), ncol = 80)
data.r <- rtk(data, ReturnMatrix = 1, depth = min(colSums(data)))

# Collectors curve on dataframe/matrix
collectors.curve(data, xlab = "No. of samples", ylab = "richness")
collectors.curve(data.r, xlab = "No. of samples (rarefied data)", ylab = "richness")

# if you want to have an accumulated order, t compare various studies to one another:
cls <- rep_len(c("a","b","c","d"), ncol(data)) # study origin of each sample
accumOrder <- c("b","a","d","c") # define the order, for the plot
colors <- c(1,2,3,4)

# Collectorscurves visualize the richness gained by picking more samples.

get.diversity Usage
get.diversity(obj, div = "richness", multi = FALSE)
get.mean.diversity(obj, div = "richness")
get.median.diversity(obj, div = "richness")
```
**Arguments**

- **obj**: Object of type rtk
- **div**: diversity measure as string e.g "richness"
- **multi**: Argument set to true if called recursively and class should not be checked. Should not be set in normal use case.

**Details**

This set of functions allows fast and easy access to calculated diversity measures by rtk. It returns a matrix, when rarefaction was only performed to one depth and a list of matrices or vectors if rarefaction was done for multiple depths.

**Author(s)**

Falk Hildebrand, Paul Saary

**References**


**See Also**

Use rtk before calling this function.

**Examples**

```r
require("rtk")
# Collectors Curve dataset should be broad and contain many samples (columns)
data     <- matrix(sample(x = c(rep(0, 15000), rep(1:10, 100)),
                   size = 10000, replace = TRUE), ncol = 80)
data.r   <- rtk(data, depth = min(colSums(data)))
get.diversity(data.r)
get.median.diversity(data.r)
get.mean.diversity(data.r)
```

**Description**

Plot rarefaction results

**Usage**

```r
## S3 method for class 'rtk'
plot(x, div = c("richness"), groups = NA, col = NULL, lty = 1,
     pch = NA, fit = "arrenerius", legend = TRUE, legend.pos = "topleft",
     boxplot = FALSE, ...)
```
Arguments

x  a rare result object

div  Diversity measure to plot. Can be any of c('richness','shannon','simpson','invsimpson','chao1','eve')
groups  If grouping is desired a vector of factors corresponding to the input samples
col  Colors used for plotting. Can be a vector of any length which will be recycled if it is too small. By default a rainbow is used.
lty  Linetypes used for plotting. Can be a vector of any length which will be recycled if it is too small.
pch  Symbols used for plotting. Can be a vector of any length which will be recycled if it is too small.
fit  Fit the rarefaction curve. Possible values: c("arrhenius","michaelis-menten","logis")
legend  Logical indicating if a legend should be created or not
legend.pos  Position of the said legend
log.dim  Character vector indicating which scale log log transform for plotting rarefaction curves.
boxplot  If a boxplot should be added to the lineplot of the rarefaction curve.
...  Other plotting input will be passed to plot or boxplot respectively

Details

To create plots from the rarefaction results you can easily just call a plot on the resulting elements. This will either produce a rarefaction curve, if more than one depth was rarefied to, or a boxplot for a single depth. Grouping of samples is possible by simply passing a vector of the length of the samples to the option groups.

Rarefaction curves can be fitted to either the arrhenius-equation, the michaelis-menten (SSmicmen) equation or the logis function SSlogis. To disable fitting fit must be set to FALSE.

Author(s)

Falk Hildebrand, Paul Saary

References


See Also

rtk, collectors.curve
Examples

```r
require("rtk")
data <- matrix(sample(x = c(rep(0, 1500), rep(1:10, 500), 1:1000), size = 120, replace = TRUE), 40)
samplesize <- min(colSums(data))
d1 <- rtk(input = data, depth = samplesize)
d2 <- rtk(input = data, depth = round(seq(1, samplesize, length.out = 10)))

# just the richness of all three samples as boxplot
plot(d1, div = "richness")
# rarefaction curve for each sample with fit
plot(d2, div = "evenness", fit = "arrhenius", pch = c(1, 2, 3))
# Rarefaction curve with boxplot, samples pooled together (grouped)
plot(d2, div = "richness", fit = FALSE, boxplot = TRUE, col = 1, groups = rep(1, ncol(data)))
```

r
t
k

**rtk**

**Rarefy tables**

**Description**

Rarefy datasets in R or from a path.

**Usage**

```r
rtk(input, repeats = 10, depth = 1000, ReturnMatrix = 0, margin = 2,
     verbose = FALSE, threads = 1, tmpdir = NULL, seed = 0)
```

**Arguments**

- **input**: This can be either a numeric matrix or a path to a text file in tab-delimited format on the locally available storage. The later option is for very big matrices, to avoid unnecessary memory consumption in R.
- **repeats**: Number of times to compute diversity measures. (default: 10)
- **depth**: Number of elements per row/column to rarefy to. The so called rarefaction depth or samplesize. Can also be a vector of ints. (default: 1000)
- **ReturnMatrix**: Number of rarefied matrices which are returned to R. Set to zero to only measure diversity. (default: 1)
- **margin**: Indicates which margin in the matrix represents the Samples and Species. Default is to rarefy assuming columns represent single samples (margin=2). If margin=1, rows are assumed to be samples. (default: 2 (columns))
- **verbose**: If extra output should be printed to std::out or not to see progress of rarefaction. (default: TRUE)
threads Number of threads to use during rarefaction
tmpdir Location to store temporary files
seed Set seed to integer > 0 to get reproducible results. default: 0

Details

Function rare takes a dataset and calculates the diversity measures, namely the shannon diversity, richness, simpson index, the inverse simpson index, chaol and evenness.

If wished for the function can also return one or multiple rarefied matrices rarefied to one or multiple depths. Those can then also be used to create collectorcurves (see collectors.curve).

Value

The function rare returns an object of class 'rarefaction', containing the objects divvs, raremat, skipped, div.median and depths. If more than one depth was computed the elements 1-4 are inside a list themself and can be accessed by the index of the desired depth.

The object divvs contains a list of diversity measures for each sample provided.

raremat is one or multiple rarefied matrices. Samples with not enough counts are removed, thus not all raremat-matrices for different depths might be of the same size. If and which samples where excluded is denoted in the element skipped using the names of the respective samples.

depths just contains the input variable and might be useful for further analysis of the results.

It is possible to plot the results of the rarefaction, depending on the parameters passed to rare. See plot.rtk for examples.

Author(s)

Paul Saary, Falk Hildebrand

References


See Also

plot.rtk, collectors.curve

Examples

require("rtk")
# generate semi sparse example data
data <- matrix(sample(x = c(rep(0, 1500), rep(1:10, 500), 1:1000),
   size = 120, replace = TRUE), 10)
# find the column with the lowest abundance
samplesize <- min(colSums(data))
# rarefy the dataset, so each column contains the same number of samples
data.rarefied <- rtk(input = data, depth = samplesize, ReturnMatrix = 1)
rtk

richness <- get.diversity(data.rarefied, div = "richness")
eveness <- get.diversity(data.rarefied, div = "eveness")
Index

collectors.curve, 3, 3, 6, 8
get.diversity, 4
get.mean.diversity(get.diversity), 4
get.median.diversity(get.diversity), 4
plot, 5
plot.rtk, 3, 4, 8
rarefactioncurve(plot), 5
rt, 5
rtk, 3, 6, 7
rtk-package, 2
SSlogis, 6
SSmicmen, 6