Package 'otuSummary'

September 6, 2023

Type Package

Title Summarizing OTU Table Regarding the Composition, Abundance and
Beta Diversity of Abundant and Rare Biospheres

Version 0.1.2

Date 2023-8-13

Author Sizhong Yang

Maintainer Sizhong Yang <yanglzu@163.com>

Description Summarizes the taxonomic composition, diversity contribution of the rare and abundant community by using OTU (operational taxonomic unit) table which was generated by analyzing pipeline of 'QIIME' or 'mothur'. The rare biosphere in this package is subset by the relative abundance threshold (for details about rare biosphere please see Lynch and Neufeld (2015) <doi:10.1038/nrmicro3400>).

Depends R (>= 3.5.0)

Imports reshape2 (>= 1.4)

URL https://github.com/cam315/otuSummary

License GPL (>= 3)

Encoding UTF-8

LazyData true

NeedsCompilation no

Repository CRAN

Date/Publication 2023-09-05 22:00:03 UTC

R topics documented:

---

alphaDiversity	*Calculate the alpha diversity indices*

Description

This function will calculate the alpha diversity indices for the total, abundant and rare biospheres.

Usage

```
alphaDiversity(otutab, siteInCol = FALSE, taxhead = NULL, threshold = 1,
    percent = FALSE, write = FALSE, ...)
```

Arguments

otutab	A OTU table of microbial community, which can contain a taxonomic column (if siteInCol) or row (if site in rows). The OTU table should be given in numeric (integer) counts.
siteInCol	Logical, if "TRUE", the OTU table contains samples in columns and taxa in rows. By default in this function, the siteInCol is FALSE, meaning the samples in rows.
taxhead	Character, specify the header of taxonomy if there is a taxonomic column in your data. By default this argument is NULL.
threshold	Numeric, the threshold of relative abundance upon which the rare biosphere will be subset.
percent	Logical, whether the input OTU table are given in relative abundance. FALSE means that the input OTU table is in numeric counts.
write	Logical, if TRUE, the result will be written out in a Tab separated data frame.
...	arguments to be passed to write.table().

Details

The rare biosphere is defined by the relative abundance cutoffs (which is the "threshold" argument in this function) (Lynch and Neufeld, 2015). This update (version 0.1.2) removed the dependencies on functions "specnumber", "diversity" and "estimateR" from the R package "vegan" (Oksanen et al, 2013), and removed the "gini" function from package 'reldist' (http://www.stat.ucla.edu/~handcock/RelDist).

Value

The function will return a list of length 3, including indices of observed, shannon, simpson, invsimpson, chao1, chao2, and evenness.

allBio	The alpha diversity indices for the whole community
abundBio	The alpha diversity indices for the abundant population
rareBio	The alpha diversity indices for the rare biosphere

Author(s)

Sizhong Yang <yanglzu@163.com>

References

Lynch MDJ, Neufeld JD (2015). Ecology and exploration of the rare biosphere. Nature Reviews Microbiology 13: 217-229.

Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB et al (2013). vegan: Community Ecology Package. R package version 2.0-7. http://CRAN.R-project.org/package=vegan.

Examples

```
data(otumothur)

test1 <- alphaDiversity(otutab = otumothur, siteInCol = TRUE,
    taxhead = "taxonomy", threshold = 1, percent = FALSE, write = FALSE)

test2 <- alphaDiversity(otutab = otumothur[,-ncol(otumothur)], siteInCol = TRUE,
    taxhead = NULL, threshold = 1, percent = FALSE, write = FALSE)
```

---

bplot	*Generate barplot with custom controls on the x axis labels*

Description

barplot with custom controls on the x axis labels, e.g. rotation.

Usage

```
bplot(data, srt = 45, yoff = 0.05, dataoff = 0.025, barcol = "grey", grid = TRUE, ...)
```

Arguments

data	Numeric, vector to plot in barplot.
srt	Numeric, rotation degree of the x axis labels.
yoff	Numeric, vertical offset of x axis labels.
dataoff	Numeric, vertical offset of data labels in relation to the bar height.
barcol	Character, color of bar. By default is grey.
grid	Logical, whether show the grids in the plot.
...	arguments to be passed to/from other methods.

Examples

```
data(otumothur)

summaryInfo <- otuReport(otutab = otumothur, siteInCol = TRUE, taxhead = "taxonomy",
    platform = "qiime", percent = FALSE, taxlevel = "phylum", collap = ";")

length(summaryInfo)
names(summaryInfo)
summaryInfo[[1]]

bplot(summaryInfo[["taxaFreqs"]])
```

---

calc_bc   *calculate the Bray-Curtis dissimilarity*

Description

This function will calculate the Bray-Curtis distance matrix for community data.

Usage

```
calc_bc(df)
```

Arguments

df	data frame or matrix, i.e. OTU or ASV table with sites in rows and species in columns.

Details

This function will calculate the dissimilarity and output as lower triangular distance matrix.

Value

The function returns lower triangular distrance matrix.

Author(s)

Sizhong Yang <yanglzu@163.com>

Examples

```
data(otuqiime)
mat <- calc_bc(t(otuqiime[,-ncol(otuqiime)]))
mat2 <- as.matrix(mat)
dim(mat2)
```

---

contrib	*Contribution of rare/abundant biosphere to the total Bray-Curtis dis-similarity*

Description

function to calculate the contribution (in fraction) of the abundant or rare biosphere to the Bray-Curtis dissimilarity of the whole community.

Usage

```
contrib(otutab, siteInCol = TRUE, taxhead = NULL, threshold = 1, percent = FALSE,
    check = "rare", write = FALSE, plot = FALSE, ...)
```

Arguments

otutab	An OTU table of microbial community, which can contain taxonomy in a column or a row.
siteInCol	Logical, if "TRUE", the OTU table contains samples in columns and taxa in rows. The function will decide whether to transpose the OTU table based on this parameter.
taxhead	Character, specify the header of taxonomy, e.g. "taxonomy" if there is a taxonomy column or row. It is NULL by default.
threshold	Numeric, the threshold relative abundance cutoff upon which the rare biosphere will be subset.
percent	Logical, whether the input OTU table are in relative abundance.
check	Character, either "rare" or "abundant", telling the function which biosphere to be check.
write	Logical, if TRUE, the result will be written out as "txt" file, default is FALSE.
plot	Logical, whether the contribution result to be visualized in boxplot. By default is FALSE.
...	arguments to be passed to/from other methods.

Details

In this function, the rare biosphere is defined by the relative abundance cutoffs (argument threshold). The Bray-Curtis distance between pairwise samples was partitioned. The Bray-Curtis measure is a scaled summation of abundance differences between two communities and can thus be partitioned for a subset population from the community (Shade et al 2014, Yang et al 2017).

Value

The function will return a data frame of five columns. The first two columns specify the sample names whose Bray-Curtis distance were calculated. The third and forth columns give the distances respectively based on the whole community OTU data or the subset data. The last column gives the contribution (in fraction, not percentage) of the subset data for each pair of samples.

Author(s)

Sizhong Yang <yanglzu@163.com>

References

Shade A, Jones SE, Caporaso JG, Handelsman J, Knight R, Fierer N et al (2014). Conditionally rare taxa disproportionately contribute to temporal changes in microbial diversity. Mbio 5: e01371-01314.

Yang S, Winkel M, Wagner D, Liebner S (2017). Community structure of rare methanogenic archaea: insight from a single functional group. FEMS Microbiology Ecology: fix126.

Examples

```
data(otuqiime)

result <- contrib(otutab = otuqiime, siteInCol = TRUE, taxhead = "taxonomy",
    threshold = 1, percent = FALSE, check = "abund", plot = FALSE)

names(result)
head(result)
```

---

  fillTax   *reshape the heirarchical taxonomy*

Description

reformat the taxonomy with the last clear assignment and changing the prefix

Usage

```
fillTax(x, split = ';', prefix=TRUE, fillAll=TRUE)
```

Arguments

`x`	Character, structured strings giving the heirarchical rank of taxonomy, e.g., d__Bacteria; p__Planctomycetota;c__Planctomycetes;o__Gemmatales;f__Gemmataceae;g__;s__
`split`	Character, the seperator for the heirarchical taxonomy.
`prefix`	Logical, whether contains prefix in the taxonomic strings, default TRUE.
`fillAll`	Logical, whether to fill all taxonic level, default TRUE.

See Also

fillTax2, slimTax

Examples

```
test = 'd__Bacteria;p__Planctomycetota;c__Planctomycetes;o__Gemmatales;f__Gemmataceae;g__;s__'
 fillTax(x = test, split = ';', prefix=TRUE, fillAll=TRUE)
```

---

fillTax2	*reshape the heirarchical taxonomy*

Description

this function is similar to fillTax, except that this function has no parameter of "fillAll"

Usage

```
fillTax2(x, split = ';', prefix=TRUE)
```

Arguments

`x`	Character, structured strings giving the heirarchical rank of taxonomy. Please convert to character object for your input taxonomic information before using.
`split`	Character, the seperator for the heirarchical taxonomy.
`prefix`	Logical, whether contains prefix in the taxonomic strings, default TRUE.

See Also

fillTax, slimTax

Examples

```
test = 'd__Bacteria;p__Planctomycetota;c__Planctomycetes;o__Gemmatales;f__Gemmataceae;g__;s__'
 fillTax2(x = test, split = ';', prefix=TRUE)
```

---

matrixConvert   *Convert lower triangular distance matrix into data frame*

Description

This function will convert lower triangular distance matrix into a 3-column, long-format data frame.

Usage

```
matrixConvert(triMatrix, colname = c("sp1", "sp2", "dist"))
```

Arguments

triMatrix   Matrix, the input matrix should be lower triangular matrix.

colname   Character, a vector of length 3 to specify the column names of the converted
                data frame.

Details

This function will call the "melt" function in the reshape2 package, and convert the pairwise values
in the lower triangular distance matrix into long-format data frame.

Value

The function returns long format of data frame, with 3 columns. The first two columns give the
pairwise names and the third column contains values in the matrix.

Author(s)

Sizhong Yang <yanglzu@163.com>

Examples

```
data(otuqiime)
mat <- calc_bc(t(otuqiime[,-ncol(otuqiime)]))
mat.m <- matrixConvert(mat, colname = c("sp1", "sp2", "bray"))
```

---

otu4type | *An example OTU table with samples and taxonomy in rows and otus in columns*

Description

A data set containing bacterial counts from the North Temperate Lakes Microbial Observatory. Due to the last row contain taxonomy, read.table function with default setting will treat the type of each column as factor.

Usage

```
data("otu4type")
```

Format

A data frame with 591 columns (OTUs) and 454 rows (453 samples plus 1 taxonomy).

Details

This data show examples of "SiteInRow" if there is a taxonomy rows. When read in, the column with numeric counts will be marked as "factor". This data set could be transposed to correct type with "typeConvert" function.

Source

see the entire dataset at https://github.com/cran/OTUtable/tree/master/data

Examples

```
data(otu4type)
sapply(otu4type, class)
new <- typeConvert(otu4type)
sapply(new, class)
```

---

otuCollap | *collapse a OTU table at given level.*

Description

The function will collapse a structured OTU table at given taxonomic level.

Usage

```
otuCollap(otutab, taxto, siteInCol = TRUE, taxhead = "taxonomy",
    pattern = ";", collap = ";")
```

Arguments

otutab	A OTU table of microbial community, which must contain a taxonomic column (if siteInCol) or row (if site in rows). The otu table can be given in numeric counts or in relative abundance.
taxto	numeric, collapse the otutable at the taxonomic level which start position (index) from the left of taxonomy. For example, in "Archaea; Euryarchaeota; Methanomicrobia; Methanomicrobiales; Methanoregulaceae", the indices of "Euryarchaeota" and "Methanomicrobiales" are 2 and 4, respectively.
siteInCol	Logical, if "TRUE", the OTU table contains samples in columns and taxa in rows. The function will decide whether to transpose the otu table based on this parameter.
taxhead	Character, specify the header of taxonomy. By default we assume your taxonomic column is entitled "taxonomy".
pattern	Character, specify the separation of taxonomy. By default, the taxonomy is separated by semicolon (";").
collap	Character, tell the function about the separation for the hierarchical order in the output.

Details

This function will directly collapse the otu table according to numeric position of structured taxonony. This function can also collapse data with structured format like the example OTU table.

Value

The function will return a collapsed OTU table.

Author(s)

Sizhong Yang <yanglzu@163.com>

See Also

otuReport

Examples

```
data(otuqiime)
dim(otuqiime)

result <- otuCollap(otutab = otuqiime, taxto = 2, siteInCol = TRUE,
    taxhead = "taxonomy", pattern = ";")

dim(result)
```

---

otumothur	*OTU table generated from 8 lakes over 4 years*

Description

A data set containing bacterial counts from the North Temperate Lakes Microbial Observatory, the taxonomy column is in mothur format. The data set is published in Msphere (Linz et al, 2017). The data set is identical to otuqiime except for the taxonomy column in qiime format.

Usage

```
data("otumothur")
```

Format

A data frame with 454 columns (453 samples plus 1 taxonomy) and 591 rows (OTUs). The taxonomy is given in mothur format, with hierarchical taxonomy from kingdom to species separated with semiclone.

Details

The full version of the data set is published in Msphere (Linz et al 2017). The first two letters of sample names denote the sampling site (e.g. "CB"), followed with epilimnion or hypolimnion ("E" or "H") and sampling date ("01OCT07"). The original data set have replicates ("R1" and "R2"), this data set only extracted the subset of "R2" (with extension ".R2" in sample names). The data set is identical to otuqiime except that the taxonomy is in format generated by software 'mothur' (Schloss et al 2009).

Source

see the entire dataset at https://github.com/cran/OTUtable/tree/master/data

References

Linz AM, Crary BC, Shade A, Owens S, Gilbert JA, Knight R et al (2017). Bacterial community composition and dynamics spanning five years in freshwater bog lakes. Msphere 2: e00169-00117. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB et al (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Applied and Environmental Microbiology 75: 7537-7541.

Examples

```
data(otumothur)
dim(otumothur)
sapply(otumothur, class)
head(otumothur$taxonomy)
```

---

otuqiime   *OTU table generated from 8 lakes over 4 years*

Description

A data set containing bacterial counts from the North Temperate Lakes Microbial Observatory, the taxonomy column is in format generated by software platfor 'QIIME' (Caporaso et al 2010). The data set is published in Msphere (Linz et al 2017).

Usage

```
data("otuqiime")
```

Format

A data frame with 454 columns (453 samples plus 1 taxonomy) and 591 rows (OTUs).

Details

The first two letters of sample names denote the sampling site (e.g. "CB"), followed with epilimnion or hypolimnion ("E" or "H") and sampling date ("01OCT07"). The original data set have replicates ("R1" and "R2"), this data set only extracted the subset of "R2" (with extension ".R2" in sample names).

Source

see https://github.com/cran/OTUtable/tree/master/data

References

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK et al (2010). QIIME allows analysis of high-throughput community sequencing data. Nature Methods 7: 335-336. Linz AM, Crary BC, Shade A, Owens S, Gilbert JA, Knight R et al (2017). Bacterial community composition and dynamics spanning five years in freshwater bog lakes. Msphere 2: e00169-00117.

Examples

```
data(otuqiime)
dim(otuqiime)
names(otuqiime)[1:10]
rownames(otuqiime)[1:10]
head(otuqiime$taxonomy)
```

---

otuReport   *Summarize the community structure and abundance with OTU table*

Description

The function will summarize the frequency, abundance at given taxonomic level for the input OTU table. This function could quickly give the summary information for user when adding these values in describing the community structure in a paper.

Usage

```
otuReport(otutab, siteInCol = TRUE, taxhead = "taxonomy", platform = "mothur",
    pattern = ";", prefix = TRUE, percent = FALSE, taxlevel = "phylum",
    collap = ";")
```

Arguments

otutab	A OTU table of microbial community, which must contain a taxonomic column (if siteInCol) or row (if site in rows). The otu table can be given in numeric counts or in relative abundance.
siteInCol	Logical, if "TRUE", the OTU table contains samples in columns and taxa in rows. The function will decide whether to transpose the otu table based on this parameter.
taxhead	Character, specify the header of taxonomy. By default we assume your taxonomic column is entitled "taxonomy".
platform	Character, argument to specify the platform generating the otu table. Currently, the function support otu table generate by "mothur" or "qiime".
pattern	Character, specify the separation of taxonomy. By default, the taxonomy is separated by semicolon (";").
prefix	Logical, tell the function whether the output will include the prefix like "p__", "c__" for the corresponding taxonomic levels.
percent	Logical, whether the input otu table are given in relative abundance. FALSE means that the input otu table is in numeric counts.
taxlevel	Character, specify the taxonomic level at which you want to know for the otu table. The valid choice are c("kingdom", "phylum", "class", "order", "family", "genus", "species").
collap	Character, tell the function about the separation for the hierarchical order in the output.

Details

This function was designed according to the structured taxonomy generated by mothur or qiime. So far, the function support the 7 levels of hierarchical taxonomy from kingdom to species.

Value

If the input otu table is in counts, the function will return a list of results summarizing 9 different aspects for the microbial community as follows:

whatTaxa	The "whatTaxa" will give which lineages are present at given taxonomic level in the community.
taxaFreqs	The element of "taxaFreqs" in the list is the frequency table of each lineage when community table were collapsed at a given taxonomic level.
taxaFrac	The "taxaFrac" element summarizes the fraction of each lineage among the total lineages at a given taxonomic level.
reads	The "reads" table is the otu table in absolute counts which has been collapsed at a given taxonomic level.
readSum	The "readSum" gives the total amount of reads in each sample after the community was collapsed at given taxonomic level.
readFrac	The "readFrac" table give the fraction of reads in relation to the total counts of the whole community at given taxonomic level.
readFracSum	The "readFracSum" table give sum of reads fraction by different lineages at given taxonomic level.
Relabund	The "Relabund" is the relative abundance table in percentage at given taxonomic level. If the input otu table is in absolute counts, the reads will be normalized by the total amount of reads of each sample, not the total amount of the whole community.
RelabundMean	The data "RelabundMean" in the list return the mean relative abundance of each lineage at given taxonomic level across all samples.

For the relative abundance otu input, the function will omit four summary table regarding absolute reads ("reads", "readSum", "readFrac" and "readFracSum").

Author(s)

Sizhong Yang <yanglzu@163.com>

See Also

subOTU, otuCollap

Examples

```
# summary the otu table in qiime format

data(otuqiime)

summaryInfo <- otuReport(otutab = otuqiime, siteInCol = TRUE, taxhead = "taxonomy",
    platform = "qiime", pattern = ";", prefix = TRUE, percent = FALSE, taxlevel = "class")

length(summaryInfo)
names(summaryInfo)
```

```
summaryInfo[[1]]

# summary otu table in mothur format

data(otumothur)

summaryInfo <- otuReport(otutab = otumothur, siteInCol = TRUE, taxhead = "taxonomy",
    platform = "mothur", pattern = ";", percent = FALSE, taxlevel = "phylum", collap = ";")

length(summaryInfo)
names(summaryInfo)
summaryInfo[[1]]

op <- par(mar = c(8,6,2,1)+0.1)
bplot(summaryInfo[["taxaFreqs"]])
par(op)

# summary otu table of relative abundance

per <- subOTU(otutab = otuqiime, siteInCol = TRUE, taxhead = "taxonomy",
    percent = FALSE, choose = "all", outype = "Relabund", sort = TRUE)

summaryInfo <- otuReport(otutab = per, siteInCol = TRUE, taxhead = "taxonomy",
    platform = "qiime", pattern = ";", percent = TRUE, taxlevel = "class")

length(summaryInfo)
names(summaryInfo)
```

---

pbray	*Partition the Bray-Curtis distance dissimilarity*

Description

The pbray function partitions the Bray-Curtis distance matrix based on the who community and the subset of the community data.

Usage

```
pbray(allComm, subComm, tolower = TRUE)
```

Arguments

allComm	A otu table of microbial community which contains sample in rows and taxa in column. The otu table should not contain a taxonomic column.
subComm	A subset of the otu table, with the same samples as allComm.
tolower	Logical, by default the function returns the lower triangular matrix.

Details

The Bray-Curtis dissimilarity is a scaled summation of abundance differences between two communities, it is thus could be partitioned between two samples attributable to a subset of the community (Shade et al., 2014; Yang et al., 2017). Note, the pbray function requires the input for "allComm" and "subComm" keep consistent either in either counts or relative abundance. If the input for "allComm" and "subComm" are the same data, the function will return the Bray-Curtis matrix for the whole community.

Value

The function returns a distance matrix by using the subset community against the whole community data.

Author(s)

Sizhong Yang <yanglzu@163.com>

References

Shade A, Jones SE, Caporaso JG, Handelsman J, Knight R, Fierer N, and Gilbert JA. Conditionally rare taxa disproportionately contribute to temporal changes in microbial diversity. Mbio, 2014, 5(4): e01371-01314.

Yang S, Winkel M, Wagner D, and Liebner S. Community structure of rare methanogenic archaea: insight from a single functional group. FEMS Microbiology Ecology, 2017: fix126.

See Also

function contrib(), matrixConvert().

Examples

```
data(otumothur)

subotus <- subOTU(otutab = otumothur, taxhead = "taxonomy", siteInCol = TRUE,
    percent = FALSE, choose = "rare", outype = "counts", sort = FALSE)
pRare <- pbray(allComm = t(otumothur[,-454]), subComm = t(subotus[,-454]))
class(pRare)

# convert to long format data frame

longdist <- matrixConvert(pRare, colname = c("sp1", "sp2", "bray"))
```

---

rareBiosphere	*Summarize different subgroups in the rare biosphere*

Description

The function discriminates different fractions of rare biosphere from the input community data, based on the ratio between the maximum and minimum abundance (for details, please see Yang et al. 2017).

Usage

```
rareBiosphere(otutab, siteInCol = TRUE, taxhead= NULL, percent = FALSE,
    threshold = 1, cutRatio = 100, cutPERare = 5, ...)
```

Arguments

otutab	An OTU table of microbial community, which can contain a taxonomic column (if siteInCol) or row (if site in rows). Note this function requires the otu table must be given in absolute counts, not in relative abundance.
siteInCol	Logical, by default it is "TRUE", meaning the OTU table contains samples in columns and taxa in rows. Otherwise, the otu table will be transposed.
taxhead	Character, specify the header of taxonomy. By default the taxonomic column is NULL.
percent	Logical, whether the input otu table is in relative abundance. The default is FALSE.
threshold	Numeric, the threshold specify the relative abundance cutoff upon which the rare biosphere is subset.
cutRatio	Numeric, the cutRatio parameter is the ratio between the maximum and minimum non-zero absolute abundance, which specify the threshold of conditionally rare taxa (CRT).
cutPERare	Numeric, the argument cutPERare specify the threshold of permanently rare taxa (PERare) in the community according to the ratio as mentioned above.
...	arguments to be passed to/from other methods.

Details

The rare biopshere constitutes different fractions of rarity. The conditionally rare taxa (CRT) are generally rare across samples but can become abundant in some samples. In contrast, the permanently rare taxa (PERare) are persistently low in abundance. Different levels of rarity may have practical implications and suggest potential roles in metabolism and ecological functions (Lynch and Neufeld, 2015; Yang et al., 2017).

This function will filter the rare biophere from the input OTU table and then discriminate different rare fractions based on the ratio between the maximum and minimum non-zero absolute abundance, which has been as an technical alternative of skewness (please see Yang et al. 2017 for details).

The function only works with otu table with absolute abundance, for the absolute reads enable to identify of absolute singletons and doubletons.

Value

The function returns a list of 4 data frames.

summaryTable	The element "summaryTable" returns a table containing the maximum and the minimum relative abundance, the ratio of the two abundance, the grouping of rarity, whether the taxa is singleton or doubletons, with additional column of taxonomy if it is included in the input otu table.
CRT	Table "CRT" gives the subset of the above "summaryTable" only for the conditionally rare taxa.
PERare	Table "PERare" shows the information of permanently rare taxa.
otherRare	Table "otherRare" summarizes the rare taxa outside the "CRT" and "PERare".

Author(s)

Sizhong Yang <yanglzu@163.com>

References

Lynch MDJ, Neufeld JD (2015). Ecology and exploration of the rare biosphere. Nature Reviews Microbiology 13: 217-229.

Yang S, Winkel M, Wagner D, Liebner S (2017). Community structure of rare methanogenic archaea: insight from a single functional group. FEMS Microbiology Ecology: fix126.

Examples

```
data(otumothur)

example <- rareBiosphere(otutab = otumothur, siteInCol = TRUE, taxhead = "taxonomy",
    percent = FALSE, threshold = 1, cutRatio = 100, cutPERare = 5)
length(example)
names(example)
head(example[["summaryTable"]])
head(example[["CRT"]])

example2 <- rareBiosphere(otutab = otumothur[,-454], siteInCol = TRUE,
    taxhead = NULL, percent=FALSE, threshold = 1, cutRatio = 100, cutPERare = 5)
length(example2)
names(example2)
```

---

slimTax   *reshape heirarchical taxonomy*

Description

The function will trunck the heirarchical taxonomy by "from" and "to".

Usage

```
slimTax(x, from, to, separator =';', jump=FALSE)
```

Arguments

x	Character, structured strings to give the heirarchical taxonomy. If x is not character, please use as.character() function to convert first.
from	numeric, valid values within 1,2,...7, which represent the tax rank of domain, phylum, class, order, family, genus and species, respectively.
to	numeric, similar to parameter "from". Please note that value of "to" must not be smaller than "from".
separator	Character, for example semicolon ";" which states the seperator of taxonomic heirarchy.
jump	Logic, Setting "jump" to TRUE means only take the tax at indices "from" and "to", excluding those in between.

Details

This function will reformat the taxonomy by the taxonomic ranks specified by "from" and "to".

Value

The function will return the reshaped taxonomy.

Author(s)

Sizhong Yang <yanglzu@163.com>

See Also

fillTax

Examples

```
test = 'd__Bacteria;p__Planctomycetota;c__Planctomycetes;o__Gemmatales;f__Gemmataceae;g__;s__'
slimTax(test, from=2, to=5)

data(otuqiime)
dim(otuqiime)

result <- slimTax(x = as.character(otuqiime$taxonomy), from=2, to=5, separator =';', jump=FALSE)

dim(result)
```

---

subOTU *Subset a OTU table*

Description

This function subset a OTU table according to the specified threshold of mean relative abundance.

Usage

```
subOTU(otutab, siteInCol = TRUE, taxhead = NULL, percent = TRUE, choose = "rare",
        threshold = 1, outype = "Relabund", sort = TRUE, write = FALSE)
```

Arguments

otutab	An OTU table of microbial community, which can contain a taxonomic column (if siteInCol) or row (if site in rows). The OTU table can be given in numeric counts or in relative abundance.
siteInCol	Logical, if "TRUE", the OTU table contains samples in columns and taxa in rows. The function will decide whether to transpose the OTU table based on this parameter.
taxhead	Character, specify the header of taxonomy. By default the taxonomic column is NULL.
percent	Logical, whether the input OTU table is in relative abundance. The default is TRUE.
choose	Character, tell the function which part to subset out. The valid choice are "rare","abundant" and "all", which will specify the rare biosphere, the abundant population and keep the whole community data.
threshold	Numeric, the threshold specify the relative abundance cutoff upon which the rare biosphere is subset.
outype	Character, specify whether the output otu subset should in relative abundance (outype="Relabund" or "relabund") or in absolute counts (outype="counts" or "Counts"). This function also support partial match of the parameter.
sort	Logical, be default, the output OTU subset is sorted according to the descending order of mean relative abundance across samples.
write	Logical, whether the output OTU table will be written out when running this function. The default is FALSE.

Details

The function will subset the OTU table to abundant (choose = "abundant") or rare biosphere (choose = "rare") according to the given relative abundance threshold of rare biosphere (Lynch and Neufeld, 2015). It will also keep the whole community without subsetting, if choose = "all". The output could be relative abundance (outype = "Relabund") or absolute counts (outype = "counts"). If sort is TRUE, the output result will be sorted by the descending order of mean relative abundance across samples.

Value

This function will return an OTU table (data frame) according to the specified arguments.

Author(s)

Sizhong Yang <yanglzu@163.com>

References

Lynch MDJ, Neufeld JD (2015). Ecology and exploration of the rare biosphere. Nature Reviews Microbiology 13: 217-229.

Examples

```
data(otuqiime)

example1 <- subOTU(otutab = otuqiime, siteInCol = TRUE, taxhead ="taxonomy",
    percent = FALSE, choose = "abundant", threshold = 1, outype = "Relabund")
dim(example1)

example2 <- subOTU(otutab = otuqiime[,-454], siteInCol = TRUE, taxhead = NULL,
    percent = FALSE, choose = "rare", threshold = 1, outype = "counts")
dim(example2)
```

---

typeConvert   *Transpose the data frame if there is data type conversion.*

Description

function to transpose the data frame if there is data type conversion.

Usage

```
typeConvert(otutab,taxhead = NULL)
```

Arguments

otutab   An OTU table, which can contain no taxonomy. The OTU table can be given in numeric counts or in relative abundance.

taxhead   Character, specify the header of taxonomy. By default the taxonomic column is NULL.

Details

This function is to convert the numeric values into right type so that the downstream numeric calculation could be processed without type error.

Value

This function returns a transposed OTU table. In the source OTU table the numeric values showing "character" or "factor" will be converted to right types.

Author(s)

Sizhong Yang <yanglzu@163.com>

Examples

```
data(otu4type)
sapply(otu4type, class)
new <- typeConvert(as.data.frame(t(otu4type)), taxhead = "taxonomy")
sapply(new, class)
```

Index