

# Package ‘microbial’

December 7, 2020

**Type** Package

**Title** Do 16s Data Analysis and Generate Figures

**Version** 0.0.17

**Description** Provides functions to enhance the available statistical analysis procedures in R by providing simple functions to analysis and visualize the 16S rRNA data. Here we present a tutorial with minimum working examples to demonstrate usage and dependencies.

**License** GPL-3

**Depends** R (>= 3.5.0)

**Imports** dplyr, plyr, magrittr, broom, phyloseq, vegan, rlang, tools, ggplot2, ggpubr, DESeq2, dada2, SummarizedExperiment, S4Vectors, Biostrings, rstatix, tidyr, DECIPHER, phangorn, MASS, randomForest, edgeR, knitr

**Encoding** UTF-8

**LazyData** true

**Suggests** markdown, rmarkdown

**VignetteBuilder** knitr

**biocViews** Software, GraphAndNetwork

**RoxygenNote** 7.1.1

**NeedsCompilation** no

**Author** Kai Guo [aut, cre],  
Pan Gao [aut]

**Maintainer** Kai Guo <guokai8@gmail.com>

**Repository** CRAN

**Date/Publication** 2020-12-07 07:40:02 UTC

## R topics documented:

|                      |   |
|----------------------|---|
| .checkfile . . . . . | 2 |
| .getstar . . . . .   | 3 |

|                          |           |
|--------------------------|-----------|
| .lda.fun . . . . .       | 3         |
| betadiv . . . . .        | 3         |
| betatest . . . . .       | 4         |
| biomarker . . . . .      | 5         |
| buildTree . . . . .      | 6         |
| data-physeq . . . . .    | 7         |
| diffest . . . . .        | 7         |
| distcolor . . . . .      | 8         |
| do_aov . . . . .         | 9         |
| do_ttest . . . . .       | 9         |
| do_wilcox . . . . .      | 10        |
| ldamarker . . . . .      | 11        |
| lightcolor . . . . .     | 11        |
| normalize . . . . .      | 12        |
| otu_table . . . . .      | 13        |
| phy_tree . . . . .       | 13        |
| plotalpha . . . . .      | 14        |
| plotbar . . . . .        | 15        |
| plotbeta . . . . .       | 16        |
| plotdiff . . . . .       | 17        |
| plotLDA . . . . .        | 18        |
| plotmarker . . . . .     | 19        |
| plotquality . . . . .    | 20        |
| prefilter . . . . .      | 21        |
| preRef . . . . .         | 22        |
| processSeq . . . . .     | 22        |
| richness . . . . .       | 23        |
| sample_data . . . . .    | 24        |
| subset_samples . . . . . | 25        |
| subset_taxa . . . . .    | 25        |
| tax_table . . . . .      | 26        |
| <b>Index</b>             | <b>27</b> |

---

|            |                          |
|------------|--------------------------|
| .checkfile | <i>check file format</i> |
|------------|--------------------------|

---

### Description

check file format

### Usage

```
.checkfile(file)
```

### Arguments

|      |          |
|------|----------|
| file | filename |
|------|----------|

---

.getstar                      *replace p value with star*

---

**Description**

replace p value with star

**Usage**

.getstar(x)

**Arguments**

x                      a (non-empty) numeric data values

---

.lda.fun                      *LEfse function*

---

**Description**

LEfse function

**Usage**

.lda.fun(df)

**Arguments**

df                      a dataframe with groups and bacteria abundance

---

betadiv                      *calcaute beta diversity*

---

**Description**

calcaute beta diversity

**Usage**

betadiv(physeq, distance = "bray", method = "PCoA")

**Arguments**

|          |   |
|----------|---|
| physeq   | A phyloseq object containing merged information of abundance, taxonomic assignment, sample data including the measured variables and categorical information of the samples, and / or phylogenetic tree if available.   |
| distance | A string character specifying dissimilarity index to be used in calculating pairwise distances (Default index is "bray"). "unifrac", "wunifrac", "manhattan", "euclidean", "canberra", "bray", "kulczynski", "jaccard", "gower", "altGower", "morisita", "horn", "mountford", "raup", "binomial", "chao", "cao" or "mahalanobis". |
| method   | A character string specifying ordination method. All methods available to the ordinate function of phyloseq are acceptable here as well.  |

**Value**

list with beta diversity data.frame and PCs

**Author(s)**

Kai Guo

**Examples**

```
{
  data("Physeq")
  phy<-normalize(physeq)
  res <- betadiv(phy)
}
```

---

|          |                                    |
|----------|------------------------------------|
| betatest | <i>PERMANOVA test for phyloseq</i> |
|----------|------------------------------------|

---

**Description**

PERMANOVA test for phyloseq

**Usage**

```
betatest(physeq, group, distance = "bray")
```

**Arguments**

|        |   |
|--------|---|
| physeq | A phyloseq object containing merged information of abundance, taxonomic assignment, sample data including the measured variables and categorical information of the samples, and / or phylogenetic tree if available. |
| group  | (Required). Character string specifying name of a categorical variable that is preferred for grouping the information. information.   |

distance A string character specifying dissimilarity index to be used in calculating pairwise distances (Default index is "bray"). "unifrac", "wunifrac", "manhattan", "euclidean", "canberra", "bray", "kulczynski", "jaccard", "gower", "altGower", "morisita", "horn", "mountford", "raup", "binomial", "chao", "cao" or "mahalanobis".

### Value

PERMANOVA test result

### Author(s)

Kai Guo

### Examples

```
{
data("Physeq")
phy<-normalize(physeq)
beta <-betatest(phy,group="SampleType")
}
```

---

biomarker

*Identify biomarker by using randomForest method*

---

### Description

Identify biomarker by using randomForest method

### Usage

```
biomarker(
  physeq,
  group,
  ntree = 500,
  pvalue = 0.05,
  normalize = TRUE,
  method = "relative"
)
```

### Arguments

physeq A phyloseq object containing merged information of abundance, taxonomic assignment, sample data including the measured variables and categorical information of the samples, and / or phylogenetic tree if available.

group group. A character string specifying the name of a categorical variable containing grouping information.

|   |          |  |
|---|----------|--|
| n | tree     | Number of trees to grow. This should not be set to too small a number, to ensure that every input row gets predicted at least a few times.         |
| p | value    | pvalue threshold for significant results from kruskal.test   |
| n | ormalize | to normalize the data before analysis(TRUE/FALSE)  |
| m | ethod    | A list of character strings specifying method to be used to normalize the phyloseq object Available methods are: "relative", "TMM", "vst", "log2". |

**Value**

data frame with significant biomarker

**Author(s)**

Kai Guo

**Examples**

```
data("Physeq")
res <- biomarker(physeq,group="group")
```

---

|           |  |
|-----------|--|
| buildTree | <i>contruction of plylogenetic tree (extreme slow)</i> |
|-----------|--|

---

**Description**

contruction of plylogenetic tree (extreme slow)

**Usage**

```
buildTree(seqs)
```

**Arguments**

|      |               |
|------|---------------|
| seqs | DNA sequences |
|------|---------------|

**Value**

tree object

**Author(s)**

Kai Guo

---

`data-physeq`*The physeq data was modified from the (Data) Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample (2011)*

---

### Description

Published in PNAS in early 2011. This work compared the microbial communities from 25 environmental samples and three known “mock communities” – a total of 9 sample types – at a depth averaging 3.1 million reads per sample. Authors were able to reproduce diversity patterns seen in many other published studies, while also investigating technical issues/bias by applying the same techniques to simulated microbial communities of known

### References

Caporaso, J. G., et al. (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. PNAS, 108, 4516-4522. PMID: PMC3063599

### Examples

```
data(Physeq)
```

---

`diffTest`*Calculate differential bacteria with DESeq2*

---

### Description

Calculate differential bacteria with DESeq2

### Usage

```
diffTest(  
  physeq,  
  group,  
  pvalue = 0.05,  
  padj = NULL,  
  log2FC = 0,  
  gm_mean = TRUE,  
  fitType = "local",  
  quiet = FALSE  
)
```

**Arguments**

|         |   |
|---------|---|
| physeq  | A phyloseq object containing merged information of abundance, taxonomic assignment, sample data including the measured variables and categorical information of the samples, and / or phylogenetic tree if available. |
| group   | group (DESeq2). A character string specifying the name of a categorical variable containing grouping information.   |
| pvalue  | pvalue threshold for significant results  |
| padj    | adjust p value threshold for significant results  |
| log2FC  | log2 Fold Change threshold  |
| gm_mean | TRUE/FALSE calculate geometric means prior to estimate size factors   |
| fitType | either "parametric", "local", or "mean" for the type of fitting of dispersions to the mean intensity.   |
| quiet   | whether to print messages at each step  |

**Value**

dataframe with differential test with DESeq2

**Author(s)**

Kai Guo

**Examples**

```
data("Physeq")
res <- diffTest(physeq, group="group")
```

---

distcolor

*distinguish colors for making figures*

---

**Description**

distinguish colors for making figures

**Usage**

```
distcolor
```

**Format**

An object of class character of length 22.

**Author(s)**

Kai Guo



---

do\_aov                      *do anova test and return results as data.frame*

---

**Description**

do anova test and return results as data.frame

**Usage**

```
do_aov(x, group, ...)
```

**Arguments**

|       |  |
|-------|--|
| x     | data.frame with sample id as the column name, genes or otu as rownames |
| group | group factor used for comparison                                       |
| ...   | parameters to anova_test   |

**Author(s)**

Kai Guo

**Examples**

```
{
  data("ToothGrowth")
  do_aov(ToothGrowth,group="supp")
}
```

---

do\_ttest                      *do t.test*

---

**Description**

do t.test

**Usage**

```
do_ttest(x, group, ref = NULL, ...)
```

**Arguments**

|       |  |
|-------|--|
| x     | data.frame with sample id as the column name, genes or otu as rownames |
| group | group factor used for comparison                                       |
| ref   | reference group  |
| ...   | parameters to t_test   |

**Author(s)**

Kai Guo

**Examples**

```
{
  data("mtcars")
  do_ttest(mtcars,group="vs")
  do_ttest(mtcars,group="cyl",ref="4")
}
```

---

`do_wilcox`*do wilcox test*

---

**Description**

do wilcox test

**Usage**`do_wilcox(x, group, ref = NULL, ...)`**Arguments**

|                    |  |
|--------------------|--|
| <code>x</code>     | data.frame with sample id as the column name, genes or otu as rownames |
| <code>group</code> | group factor used for comparison                                       |
| <code>ref</code>   | reference group  |
| <code>...</code>   | parameters to <code>wilcox_test</code>                                 |

**Author(s)**

Kai Guo

**Examples**

```
{
  data("mtcars")
  do_wilcox(mtcars,group="vs")
  do_wilcox(mtcars,group="cyl",ref="4")
}
```

---

|           |   |
|-----------|---|
| ldamarker | <i>Identify biomarker by using LEfSe method</i> |
|-----------|---|

---

**Description**

Identify biomarker by using LEfSe method

**Usage**

```
ldamarker(physeq, group, pvalue = 0.05, normalize = TRUE, method = "relative")
```

**Arguments**

|           |   |
|-----------|---|
| physeq    | A phyloseq object containing merged information of abundance, taxonomic assignment, sample data including the measured variables and categorical information of the samples, and / or phylogenetic tree if available. |
| group     | group. A character string specifying the name of a categorical variable containing grouping information.  |
| pvalue    | pvalue threshold for significant results from kruskal.test  |
| normalize | to normalize the data before analysis(TRUE/FALSE)   |
| method    | A list of character strings specifying method to be used to normalize the phyloseq object Available methods are: "relative", "TMM", "vst", "log2".  |

**Author(s)**

Kai Guo

**Examples**

```
data("Physeq")
res <- ldamarker(physeq,group="group")
```

---

|            |  |
|------------|--|
| lightcolor | <i>light colors for making figures</i> |
|------------|--|

---

**Description**

light colors for making figures

**Usage**

```
lightcolor
```

**Format**

An object of class character of length 56.

**Author(s)**

Kai Guo

---

normalize

*Normalize the phyloseq object with different methods*

---

**Description**

Normalize the phyloseq object with different methods

**Usage**

```
normalize(physeq, group, method = "relative", table = FALSE)
```

**Arguments**

|        |   |
|--------|---|
| physeq | A phyloseq object containing merged information of abundance, taxonomic assignment, sample data including the measured variables and categorical information of the samples, and / or phylogenetic tree if available. |
| group  | group (DESeq2). A character string specifying the name of a categorical variable containing grouping information.   |
| method | A list of character strings specifying method to be used to normalize the phyloseq object Available methods are: "relative", "TMM", "vst", "log2".  |
| table  | return a data.frame or not  |

**Value**

phyloseq object with normalized data

**Author(s)**

Kai Guo

**Examples**

```
{
  data("Physeq")
  phy<-normalize(physeq)
}
```

---

|           |                          |
|-----------|--------------------------|
| otu_table | <i>extract otu table</i> |
|-----------|--------------------------|

---

**Description**

extract otu table

**Usage**

```
otu_table(physeq, ...)
```

**Arguments**

|        |  |
|--------|--|
| physeq | (Required). An integer matrix, otu_table-class, or phyloseq-class. |
| ...    | parameters for the otu_table function in phyloseq package          |

---

|          |  |
|----------|--|
| phy_tree | <i>Retrieve phylogenetic tree (phylo-class) from object.</i> |
|----------|--|

---

**Description**

Retrieve phylogenetic tree (phylo-class) from object.

**Usage**

```
phy_tree(physeq, ...)
```

**Arguments**

|        |  |
|--------|--|
| physeq | (Required). An instance of phyloseq-class that contains a phylogenetic tree. If physeq is a phylogenetic tree (a component data class), then it is returned as-is. |
| ...    | parameters for the phy_tree function in phyloseq package   |

---

plotalpha

*plot alpha diversity*


---

### Description

plot alpha diversity

### Usage

```
plotalpha(
  physeq,
  group,
  method = c("Observed", "Simpson", "Shannon"),
  color = NULL,
  geom = "boxplot",
  pvalue = 0.05,
  padj = NULL,
  sig.only = TRUE,
  wilcox = FALSE,
  show.number = FALSE
)
```

### Arguments

|             |   |
|-------------|---|
| physeq      | A phyloseq object containing merged information of abundance, taxonomic assignment, sample data including the measured variables and categorical information of the samples, and / or phylogenetic tree if available.           |
| group       | group (Required). A character string specifying the name of a categorical variable containing grouping information.   |
| method      | A list of character strings specifying method to be used to calculate for alpha diversity in the data. Available methods are: "Observed", "Chao1", "ACE", "Richness", "Fisher", "Simpson", "Shannon", "Evenness", "InvSimpson". |
| color       | A vector of character use specifying the color  |
| geom        | different geom to display("boxplot", "violin", "dotplot")   |
| pvalue      | pvalue threshold for significant dispersion results   |
| padj        | adjust p value threshold for significant dispersion results   |
| sig.only    | display the significant comparison only(TRUE/ FALSE)  |
| wilcox      | use wilcoxon test or not  |
| show.number | to show the pvalue instead of significant symbol(TRUE/FALSE)  |

### Value

Returns a ggplot object. This can further be manipulated as preferred by user.

**Author(s)**

Kai Guo

**Examples**

```
{
  data("Physeq")
  plotalpha(physeq, group="SampleType")
}
```

---

`plotbar`*plot bar for relative abundance for bacteria*

---

**Description**

plot bar for relative abundance for bacteria

**Usage**

```
plotbar(
  physeq,
  level = "Phylum",
  color = NULL,
  group = NULL,
  top = 5,
  fontsize.x = 5,
  fontsize.y = 12
)
```

**Arguments**

|                         |   |
|-------------------------|---|
| <code>physeq</code>     | A phyloseq object containing merged information of abundance, taxonomic assignment, sample data including the measured variables and categorical information of the samples, and / or phylogenetic tree if available. |
| <code>level</code>      | the level to plot   |
| <code>color</code>      | A vector of character use specifying the color  |
| <code>group</code>      | group (Optional). A character string specifying the name of a categorical variable containing grouping information.   |
| <code>top</code>        | the number of most abundance bacteria to display  |
| <code>fontsize.x</code> | the size of x axis label  |
| <code>fontsize.y</code> | the size of y axis label  |

**Value**

Returns a ggplot object. This can further be manipulated as preferred by user.

**Author(s)**

Kai Guo

**Examples**

```
data("Physeq")
phy<-normalize(physeq)
plotbar(phy,level="Phylum")
```

plotbeta

*plot beta diversity***Description**

plot beta diversity

**Usage**

```
plotbeta(
  physeq,
  group,
  shape = NULL,
  distance = "bray",
  method = "PCoA",
  color = NULL,
  size = 3,
  ellipse = FALSE
)
```

**Arguments**

|          |   |
|----------|---|
| physeq   | A phyloseq object containing merged information of abundance, taxonomic assignment, sample data including the measured variables and categorical information of the samples, and / or phylogenetic tree if available.   |
| group    | (Required). Character string specifying name of a categorical variable that is preferred for grouping the information.  |
| shape    | shape(Optional) Character string specifying shape of a categorical variable   |
| distance | A string character specifying dissimilarity index to be used in calculating pairwise distances (Default index is "bray"). "unifrac","wunifrac","manhattan", "euclidean", "canberra", "bray", "kulczynski", "jaccard", "gower", "altGower", "morisita", "horn", "mountford", "raup", "binomial", "chao", "cao" or "mahalanobis". |
| method   | A character string specifying ordination method. All methods available to the ordinate function of phyloseq are acceptable here as well.  |



|         |                              |
|---------|------------------------------|
| color   | user defined color for group |
| size    | the point size               |
| ellipse | draw ellipse or not          |

**Value**

ggplot2 object

**Author(s)**

Kai Guo

**Examples**

```
{
  data("Physeq")
  phy<-normalize(physeq)
  plotbeta(phy,group="SampleType")
}
```

---

plotdiff

*plot differential results*

---

**Description**

plot differential results

**Usage**

```
plotdiff(
  res,
  level = "Genus",
  color = NULL,
  pvalue = 0.05,
  padj = NULL,
  log2FC = 0,
  size = 3,
  fontsize.x = 5,
  fontsize.y = 10,
  horiz = TRUE
)
```

**Arguments**

|            |  |
|------------|--|
| res        | differential test results from diff_test         |
| level      | the level to plot                                |
| color      | A vector of character use specifying the color   |
| pvalue     | pvalue threshold for significant results         |
| padj       | adjust p value threshold for significant results |
| log2FC     | log2 Fold Change threshold                       |
| size       | size for the point                               |
| fontsize.x | the size of x axis label                         |
| fontsize.y | the size of y axis label                         |
| horiz      | horizontal or not (TRUE/FALSE)                   |

**Value**

ggplot object

**Author(s)**

Kai Guo

**Examples**

```
data("Physeq")
res <- difftest(physeq,group="group")
plotdiff(res,level="Genus",padj=0.001)
```

---

plotLDA

*plot LEfSe results from ldamarker function*

---

**Description**

plot LEfSe results from ldamarker function

**Usage**

```
plotLDA(  
  x,  
  group,  
  lda = 2,  
  pvalue = 0.05,  
  padj = NULL,  
  color = NULL,  
  fontsize.x = 4,  
  fontsize.y = 5  
)
```

**Arguments**

|            |   |
|------------|---|
| x          | LEfse results from ldamarker                                |
| group      | a vector include two character to show the group comparison |
| lda        | LDA threshold for significant biomarker                     |
| pvalue     | pvalue threshold for significant results                    |
| padj       | adjust p value threshold for significant results            |
| color      | A vector of character use specifying the color              |
| fontsize.x | the size of x axis label                                    |
| fontsize.y | the size of y axis label                                    |

**Value**

ggplot2 object

**Author(s)**

Kai Guo

**Examples**

```
data("Physeq")
res <- ldamarker(physeq,group="group")
plotLDA(res,group=c("A","B"),lda=5,pvalue=0.05)
```

---

plotmarker

*plot the biomarker from the biomarker function with randomForest*

---

**Description**

plot the biomarker from the biomarker function with randomForest

**Usage**

```
plotmarker(  
  x,  
  level = "Genus",  
  top = 30,  
  rotate = FALSE,  
  dot.size = 8,  
  label.color = "black",  
  label.size = 6  
)
```

**Arguments**

|             |   |
|-------------|---|
| x           | biomarker results from randomForest       |
| level       | the bacteria level to display             |
| top         | the number of important biomarker to draw |
| rotate      | TRUE/FALSE                                |
| dot.size    | size for the dot                          |
| label.color | label color                               |
| label.size  | label size                                |

**Value**

ggplot2 object

**Author(s)**

Kai Guo

**Examples**

```
data("Physeq")
res <- biomarker(physeq, group="group")
plotmarker(res, level="Genus")
```

---

plotquality      *plot the quality for the fastq file*

---

**Description**

plot the quality for the fastq file

**Usage**

```
plotquality(file, n = 5e+05, aggregate = FALSE)
```

**Arguments**

|           |  |
|-----------|--|
| file      | (Required). character. File path(s) to fastq or fastq.gz file(s).                                      |
| n         | (Optional). Default 500,000. The number of records to sample from the fastq file.                      |
| aggregate | (Optional). Default FALSE. If TRUE, compute an aggregate quality profile for all fastq files provided. |

**Value**

figure

**Examples**

```
plotquality(system.file("extdata", "sam1F.fastq.gz", package="dada2"))
```

---

prefilter

*filter the phyloseq*

---

**Description**

filter the phyloseq

**Usage**

```
prefilter(physeq, min = 10, perc = 0.05)
```

**Arguments**

|        |   |
|--------|---|
| physeq | A phyloseq object containing merged information of abundance, taxonomic assignment, sample data including the measured variables and categorical information of the samples, and / or phylogenetic tree if available. |
| min    | Numeric, the threshold for minimal Phylum shown in samples  |
| perc   | Numeric, input the percentage of samples for which to filter low counts.  |

**Value**

filter phyloseq object

**Author(s)**

Kai Guo

**Examples**

```
data("Physeq")  
physeqs<-prefilter(physeq)
```

---

|        |  |
|--------|--|
| preRef | <i>Download the reference database</i> |
|--------|--|

---

**Description**

Download the reference database

**Usage**

```
preRef(ref_db, path = ".")
```

**Arguments**

|        |                        |
|--------|------------------------|
| ref_db | the reference database |
| path   | path for the database  |

**Value**

the path of the database

**Author(s)**

Kai Guo

**Examples**

```
preRef(ref_db="silva", path=tempdir())
```

---

|            |                               |
|------------|-------------------------------|
| processSeq | <i>Perform dada2 analysis</i> |
|------------|-------------------------------|

---

**Description**

Perform dada2 analysis

**Usage**

```
processSeq(  
  path = ".",  
  truncLen = c(240, 160),  
  trimLeft = 0,  
  trimRight = 0,  
  minLen = 20,  
  maxlen = Inf,
```

```

sample_info = NULL,
train_data = "silva_nr99_v138_train_set.fa.gz",
train_species = "silva_species_assignment_v138.fa.gz",
outpath = NULL,
buildtree = FALSE,
verbose = TRUE
)

```

### Arguments

|               |  |
|---------------|--|
| path          | working dir for the input reads  |
| truncLen      | (Optional). Default 0 (no truncation). Truncate reads after truncLen bases. Reads shorter than this are discarded.   |
| trimLeft      | (Optional). The number of nucleotides to remove from the start of each read.   |
| trimRight     | (Optional). Default 0. The number of nucleotides to remove from the end of each read. If both truncLen and trimRight are provided, truncation will be performed after trimRight is enforced. |
| minLen        | (Optional). Default 20. Remove reads with length less than minLen. minLen is enforced after trimming and truncation.   |
| maxLen        | (Optional). Default Inf (no maximum). Remove reads with length greater than maxLen. maxLen is enforced before trimming and truncation.   |
| sample_info   | (Optional).sample information for the sequence   |
| train_data    | (Required).training database   |
| train_species | (Required). species database   |
| outpath       | (Optional).the path for the filtered reads and th out table  |
| buildtree     | build phylogenetic tree or not(default: FALSE)   |
| verbose       | (Optional). Default TRUE. Print verbose text output.   |

### Value

list include count table, summary table, taxonomy information and phyloseq object

### Author(s)

Kai Guo

---

|          |  |
|----------|--|
| richness | <i>calculat the richness for the phyloseq object</i> |
|----------|--|

---

### Description

calculat the richness for the phyloseq object

**Usage**

```
richness(physeq, method = c("Observed", "Simpson", "Shannon"))
```

**Arguments**

**physeq** A phyloseq object containing merged information of abundance, taxonomic assignment, sample data including the measured variables and categorical information of the samples, and / or phylogenetic tree if available.

**method** A list of character strings specifying method to be used to calculate for alpha diversity in the data. Available methods are: "Observed", "Chao1", "ACE", "Richness", "Fisher", "Simpson", "Shannon", "Evenness", "InvSimpson".

**Value**

data.frame of alpha diversity

**Author(s)**

Kai Guo

**Examples**

```
{
  data("Physeq")
  rich <- richness(physeq, method=c("Simpson", "Shannon"))
}
```

---

sample\_data

*extract sample information*

---

**Description**

extract sample information

**Usage**

```
sample_data(physeq, ...)
```

**Arguments**

**physeq** (Required). A data.frame-class, or a phyloseq-class object.

**...** parameters for the sample\_data function in phyloseq package



---

|                |  |
|----------------|--|
| subset_samples | <i>Subset the phyloseq based on sample</i> |
|----------------|--|

---

**Description**

Subset the phyloseq based on sample

**Usage**

```
subset_samples(physeq, ...)
```

**Arguments**

|        |  |
|--------|--|
| physeq | A sample_data-class, or a phyloseq-class object with a sample_data. If the sample_data slot is missing in physeq, then physeq will be returned as-is, and a warning will be printed to screen. |
| ...    | parameters for the subset_samples function in phyloseq package   |

---

|             |   |
|-------------|---|
| subset_taxa | <i>Subset species by taxonomic expression</i> |
|-------------|---|

---

**Description**

Subset species by taxonomic expression

**Usage**

```
subset_taxa(physeq, ...)
```

**Arguments**

|        |  |
|--------|--|
| physeq | A sample_data-class, or a phyloseq-class object with a sample_data. If the sample_data slot is missing in physeq, then physeq will be returned as-is, and a warning will be printed to screen. |
| ...    | parameters for the subset_taxa function in phyloseq package  |

---

|           |                               |
|-----------|-------------------------------|
| tax_table | <i>extract taxonomy table</i> |
|-----------|-------------------------------|

---

**Description**

extract taxonomy table

**Usage**

```
tax_table(physeq, ...)
```

**Arguments**

|        |  |
|--------|--|
| physeq | An object among the set of classes defined by the phyloseq package that contain taxonomyTable. |
| ...    | parameters for the tax_table function in phyloseq package                                      |

# Index

- \* **datasets**
  - distcolor, 8
  - lightcolor, 11
- \* **data**
  - data-physeq, 7
  - .checkfile, 2
  - .getstar, 3
  - .lda.fun, 3
  
- betadiv, 3
- betatest, 4
- biomarker, 5
- buildTree, 6
  
- data-physeq, 7
- difftest, 7
- distcolor, 8
- do\_aov, 9
- do\_ttest, 9
- do\_wilcox, 10
  
- ldamarker, 11
- lightcolor, 11
  
- normalize, 12
  
- otu\_table, 13
  
- phy\_tree, 13
- physeq (data-physeq), 7
- plotalpha, 14
- plotbar, 15
- plotbeta, 16
- plotdiff, 17
- plotLDA, 18
- plotmarker, 19
- plotquality, 20
- prefilter, 21
- preRef, 22
- processSeq, 22
  
- richness, 23
  
- sample\_data, 24
- subset\_samples, 25
- subset\_taxa, 25
  
- tax\_table, 26