# Package 'MicroNiche'

October 12, 2022

**Type** Package

**Title** Microbial Niche Measurements

**Version** 1.0.0

**Depends** ggplot2, reshape2, stats

**Date** 2020-01-19

**Maintainer** Damien Finn <damien.finn@uqconnect.edu.au>

**Description** Measures niche breadth and overlap of microbial taxa from large matrices. Niche breadth measurements include Levins' niche breadth (Bn) index, Hurlbert's Bn and Feinsinger's proportional similarity (PS) index. (Feinsinger, P., Spears, E.E., Poole, R.W. (1981) <doi:10.2307/1936664>). Niche overlap measurements include Levin's Overlap (Ludwig, J.A. and Reynolds, J.F. (1988, ISBN:0471832359)) and a Jaccard similarity index of Feinsinger's PS values between taxa pairs, as Proportional Overlap.

**License** GPL-2

**NeedsCompilation** no

**Author** Damien Finn [aut, cre]

**Repository** CRAN

**Date/Publication** 2020-01-30 09:40:10 UTC

## R topics documented:

---

MicroNiche-package    *Microbial Niche Measurements*

---

**Description**

Measures niche breadth and overlap of microbial taxa from large matrices. Niche breadth measurements include Levins' niche breadth (Bn) index, Hurlbert's Bn and Feinsinger's proportional similarity (PS) index. (Feinsinger, P., Spears, E.E., Poole, R.W. (1981) <doi:10.2307/1936664>). Niche overlap measurements include Levin's Overlap (Ludwig, J.A. and Reynolds, J.F. (1988, ISBN:0471832359)) and a Jaccard similarity index of Feinsinger's PS values between taxa pairs, as Proportional Overlap.

**Details**

The DESCRIPTION file:

| | |
|---|---|
| Package: | MicroNiche |
| Type: | Package |
| Title: | Microbial Niche Measurements |
| Version: | 1.0.0 |
| Depends: | ggplot2, reshape2, stats |
| Date: | 2020-01-19 |
| Authors@R: | person("Damien", "Finn", role = c("aut", "cre"), email = "damien.finn@uqconnect.edu.au") |
| Maintainer: | Damien Finn <damien.finn@uqconnect.edu.au> |
| Description: | Measures niche breadth and overlap of microbial taxa from large matrices. Niche breadth measurements incluc |
| License: | GPL-2 |
| Author: | Damien Finn [aut, cre] |

Index of help topics:

```
MicroNiche-package    Microbial Niche Measurements
df                    Data frame for package MicroNiche
feinsingers.PS        Feinsinger's proportional similarity index
hurlberts.Bn          Hurlbert's niche breadth index
levins.Bn             Levins' niche breadth index
levins.overlap        Levin's niche overlap index
proportional.overlap  Proportional similarity overlap index
```

**Author(s)**

NA

Maintainer: Damien Finn <damien.finn@uqconnect.edu.au>

**References**

Feinsinger et al. 1981. A simple measure of niche breadth. Ecology 62(1):27-32

Ludwig and Reynolds. 1988. Statistical Ecology. Wiley and Sons, Milton Australia

## See Also

levins.Bn hurlberts.Bn feinsingers.PS levins.overlap proportional.overlap

## Examples

```
NA
```

---

df                          *Data frame for package MicroNiche*

---

## Description

This data frame consists of taxa (rows) by samples (columns). Taxa are separated into six distributions (D1 - 6) with 10 'species' per distribution. Samples are separated into four 'environments' (R1 - 4) with 10 samples per environment.

## Usage

```
data("df")
```

## Format

A data frame with 60 observations on the following 41 variables.

Taxon a factor with levels D1S1 D1S10 D1S2 D1S3 D1S4 D1S5 D1S6 D1S7 D1S8 D1S9 D2S1 D2S10 D2S2 D2S3 D2S4 D2S5 D2S6 D2S7 D2S8 D2S9 D3S1 D3S10 D3S2 D3S3 D3S4 D3S5 D3S6 D3S7 D3S8 D3S9 D4S1 D4S10 D4S2 D4S3 D4S4 D4S5 D4S6 D4S7 D4S8 D4S9 D5S1 D5S10 D5S2 D5S3 D5S4 D5S5 D5S6 D5S7 D5S8 D5S9 D6S1 D6S10 D6S2 D6S3 D6S4 D6S5 D6S6 D6S7 D6S8 D6S9

R1S1 a numeric vector

R1S2 a numeric vector

R1S3 a numeric vector

R1S4 a numeric vector

R1S5 a numeric vector

R1S6 a numeric vector

R1S7 a numeric vector

R1S8 a numeric vector

R1S9 a numeric vector

R1S10 a numeric vector

R2S1 a numeric vector

R2S2 a numeric vector

R2S3 a numeric vector

R2S4  a numeric vector

R2S5  a numeric vector

R2S6  a numeric vector

R2S7  a numeric vector

R2S8  a numeric vector

R2S9  a numeric vector

R2S10  a numeric vector

R3S1  a numeric vector

R3S2  a numeric vector

R3S3  a numeric vector

R3S4  a numeric vector

R3S5  a numeric vector

R3S6  a numeric vector

R3S7  a numeric vector

R3S8  a numeric vector

R3S9  a numeric vector

R3S10  a numeric vector

R4S1  a numeric vector

R4S2  a numeric vector

R4S3  a numeric vector

R4S4  a numeric vector

R4S5  a numeric vector

R4S6  a numeric vector

R4S7  a numeric vector

R4S8  a numeric vector

R4S9  a numeric vector

R4S10  a numeric vector

## Details

An object of class "data.frame" of 60 rows x 40 columns, representing taxa and samples, respectively.

## Source

Generated in silico to compare taxa of known distributions: D1 are generalists roughly in equal abundance across the four environments. D2 decrease linearly across the four environments. D3 are specialists that decrease exponentially across the four environments. D4 are specialists unique to one of the four environments. D5 are specialists in environments two and four, with no relationship with an environmental gradient supplied in the vignette. Finally, D6 represent low, spurious counts haphazardly distributed across all samples often present in microbial count tables.

**References**

NA

**Examples**

```
data(df)
```

---

feinsingers.PS             *Feinsinger's proportional similarity index*

---

**Description**

This index is a measurement of proportional similarity (PS) of a taxon in relation to an environmental parameter, such as pH or temperature. A value of 0 indicates an inverse relationship between taxon abundance and the environmental parameter, whilst 1 indicates a positive relationship between taxon abundance and the environmental parameter. A value of 0.5 indicates no relationship.

**Usage**

```
feinsingers.PS(df, R, sampleInfo, envInfo, q = 1.65)
```

**Arguments**

| | |
|---|---|
| df | A matrix of taxa (rows) by samples (columns) as discrete counts per sample. Col 1 must be a taxon identifier. |
| R | The number of different environments being compared. |
| sampleInfo | A categorical variable identifying which samples correspond to which environment. |
| envInfo | A quantitative variable of some environmental parameter measured per sample. |
| q | A coefficient for the LOQ, set to 1.65. Decreasing or increasing q will affect which taxa are flagged as being below the LOQ. |

**Details**

Taxon inputs must be as discrete counts across samples, with a taxon identifier in column 1. This function also performs null model testing against a distribution of PS generated from R random proportions permuted 999 times. This allows the user to perform significance testing and to determine a probability that the taxa PS differs from the null PS. The output includes each taxon's PS, a P value of significance testing against the null model distribution, and a Benjamin-Hochberg adjusted P value to account for false discovery rate.

A plot of the null distribution with 0.05 and 0.95 quantiles highlighted as red dotted lines is also provided.

Finally, this function also calculates the limit of quantification (LOQ) to identify taxa whose niche cannot be confidently measured. A plot of log abundance x taxon rank is provided to visualise the

distribution of the data, a fitted lognormal taxa rank model, and the LOQ determined based on the standard deviation of the lognormal model.

Feinsinger's PS is calculated as follows: PS[j] = 1 - 0.5*sum(|p[i] - r[i]|) whereby p[i] is the proportion of taxon j in environment i and r[i] is the proportion of the environmental parameter in i.

This function depends on ggplot2 to generate a plot of the null distribution and reshape2 to identify the taxa below the LOQ.

**Value**

An object of class "data.frame" that gives the PS, P value and adjusted P value for each taxon. A column noting taxa that fall below the LOQ are identified as T.

**Note**

The null model testing becomes increasingly robust with increasing R. If R == 2 the results should be interpreted with caution.

**Author(s)**

Damien Finn

**References**

Feinsinger et al. 1981. A simple measure of niche breadth. Ecology 62(1):27-32

**See Also**

proportional.overlap

**Examples**

```
data(df)
sampleInfo <- c(rep("R1",10), rep("R2",10), rep("R3",10), rep("R4",10))
pH.grad <- c(2.1, 2.2, 2, 1.9, 2.1, 1.8, 1.9, 2, 2.1, 1.9, 3.5, 3.6, 3.5,
             3.4, 3.6, 3.5, 3.5, 3.4, 3.7, 3.4, 6.6, 6.5, 6.4, 6.8, 7, 6.6,
             6.8, 6.9, 7, 7.1, 8, 8.2, 7.9, 8.1, 7.8, 7.9, 8.3, 8.2, 8.1, 7.9)
res <- feinsingers.PS(df, 4, sampleInfo, pH.grad)
```

---

hurlberts.Bn                    *Hurlbert's niche breadth index*

---

**Description**

This index is a measurement of niche breadth (Bn) of a taxon in relation to an environmental parameter, such as pH or temperature. A value of 0 indicates an inverse relationship between taxon abundance and the environmental parameter, whilst 1 indicates a positive relationship between taxon abundance and the environmental parameter. A value of 0.5 indicates no relationship.

## Usage

```
hurlberts.Bn(df, R, sampleInfo, envInfo, q = 1.65)
```

## Arguments

| | |
|---|---|
| df | A matrix of taxa (rows) by samples (columns) as discrete counts per sample. Col 1 must be a taxon identifier. |
| R | The number of different environments being compared. |
| sampleInfo | A categorical variable identifying which samples correspond to which environment. |
| envInfo | A quantitative variable of some environmental parameter measured per sample. |
| q | A coefficient for the LOQ, set to 1.65. Decreasing or increasing q will affect which taxa are flagged as being below the LOQ. |

## Details

Taxon inputs must be as discrete counts across samples, with a taxon identifier in column 1. This function also performs null model testing against a distribution of Bn generated from R random proportions permuted 999 times. This allows the user to perform significance testing and to determine a probability that the taxa Bn differs from the null Bn. The output includes each taxon's Bn, a P value of significance testing against the null model distribution, and a Benjamin-Hochberg adjusted P value to account for false discovery rate.

A plot of the null distribution with 0.05 and 0.95 quantiles highlighted as red dotted lines is also provided.

Finally, this function also calculates the limit of quantification (LOQ) to identify taxa whose niche cannot be confidently measured. A plot of log abundance x taxon rank is provided to visualise the distribution of the data, a fitted lognormal taxa rank model, and the LOQ determined based on the standard deviation of the lognormal model.

Hurlbert's niche breadth is calculated as follows: $Bn[j] = 1/sum(p[i]^2/r[i])$ whereby $p[i]$ is the proportion of taxon j in environment i and $r[i]$ is the proportion of the environmental parameter in i.

This function depends on ggplot2 to generate a plot of the null distribution and reshape2 to identify the taxa below the LOQ.

## Value

An object of class "data.frame" that gives the Bn, P value and adjusted P value for each taxon. A column noting taxa that fall below the LOQ are identified as T.

## Note

The null model testing becomes increasingly robust with increasing R. If R == 2 the results should be interpreted with caution.

## Author(s)

Damien Finn

## References

Feinsinger et al. 1981. A simple measure of niche breadth. Ecology 62(1):27-32

## See Also

feinsingers.PS

## Examples

```
data(df)
sampleInfo <- c(rep("R1",10), rep("R2",10), rep("R3",10), rep("R4",10))
pH.grad <- c(2.1, 2.2, 2, 1.9, 2.1, 1.8, 1.9, 2, 2.1, 1.9, 3.5, 3.6, 3.5,
             3.4, 3.6, 3.5, 3.5, 3.4, 3.7, 3.4, 6.6, 6.5, 6.4, 6.8, 7, 6.6,
             6.8, 6.9, 7, 7.1, 8, 8.2, 7.9, 8.1, 7.8, 7.9, 8.3, 8.2, 8.1, 7.9)
res <- hurlberts.Bn(df, 4, sampleInfo, pH.grad)
```

---

| levins.Bn | *Levins' niche breadth index* |
|---|---|

---

## Description

This index is a measurement of niche breadth (Bn) of a taxon weighted by the number of differing environments (R), whereby a value of 1 indicates generalist taxa that are equally abundant across environments, and 0 indicates specialist taxa that favour specific environments.

## Usage

```
levins.Bn(df, R, sampleInfo, q = 1.65)
```

## Arguments

| | |
|---|---|
| df | A matrix of taxa (rows) by samples (columns) as discrete counts per sample. Col 1 must be a taxon identifier. |
| R | The number of different environments being compared. |
| sampleInfo | A categorical variable identifying which samples correspond to which environment. |
| q | A coefficient for the LOQ, set to 1.65. Decreasing or increasing q will affect which taxa are flagged as being below the LOQ. |

## Details

Taxon inputs must be as discrete counts across samples, with a taxon identifier in column 1. This function also performs null model testing against a distribution of Bn generated from R random proportions permuted 999 times. This allows the user to perform significance testing and to determine a probability that the taxa Bn differs from the null Bn. The output includes each taxon's Bn, a P value of significance testing against the null model distribution, and a Benjamin-Hochberg adjusted P value to account for false discovery rate.

A plot of the null distribution with 0.05 and 0.95 quantiles highlighted as red dotted lines is also provided.

Finally, this function also calculates the limit of quantification (LOQ) to identify taxa whose niche cannot be confidently measured. A plot of log abundance x taxon rank is provided to visualise the distribution of the data, a fitted lognormal taxa rank model, and the LOQ determined based on the standard deviation of the lognormal model.

Levin's niche breadth is calculated as follows: $Bn[j] = (1/R)/sum(p[i]^2)$ whereby R is the number of differing environments and $p[i]$ is the proportion of taxon j in environment i.

This function depends on ggplot2 to generate a plot of the null distribution and reshape2 to identify the taxa below the LOQ.

### Value

An object of class "data.frame" that gives the Bn, P value and adjusted P value for each taxon. A column noting taxa that fall below the LOQ are identified as T.

### Note

The null model testing becomes increasingly robust with increasing R. If R == 2 the results should be interpreted with caution.

### Author(s)

Damien Finn

### References

Feinsinger et al. 1981. A simple measure of niche breadth. Ecology 62(1):27-32

### See Also

levins.overlap

### Examples

```
data(df)
sampleInfo <- c(rep("R1",10), rep("R2",10), rep("R3",10), rep("R4",10))
res <- levins.Bn(df, 4, sampleInfo)
```

---

levins.overlap                          *Levin's niche overlap index*

---

### Description

This index is a measurement of the overlap between pairs of taxa across samples. A value of 1 indicates perfect overlap, and 0 indicates no overlap.

### Usage

```
levins.overlap(df, q = 1.65)
```

### Arguments

df
: A matrix of taxa (rows) by samples (columns) as discrete counts per sample. Col 1 must be a taxon identifier.

q
: A coefficient for the LOQ, set to 1.65. Decreasing or increasing q will affect which taxa are flagged as being below the LOQ.

### Details

Taxon inputs must be as discrete counts across samples, with a taxon identifier in column 1. The output is a matrix of paired taxon overlap indices. Specifically, Levin's Overlap (LO) indices of LO[1,2] are rows and LO[2,1] are columns, for taxa 1 and 2.

Levin's overlap is calculated as follows: $LO[i,j] = sum(p[i]*p[j])/sum(p[i]^2)$ whereby $p[i]$ is the proportion of taxon i per sample and $p[j]$ is the proportion of taxon j per sample.

Taxa below the LOQ are flagged with an asterisk.

This function depends on reshape2 to produce a matrix of overlap comparisons.

### Value

An object of class "data.frame" that gives LO[1,2] as rows and LO[2,1] as columns.

### Note

NA

### Author(s)

Damien Finn

### References

Ludwig and Reynolds. 1988 Statistical Ecology. Wiley and Sons, Milton Australia

## See Also

levins.Bn

## Examples

```
data(df)
overlapdf <- df[48:53,]
res <- levins.overlap(overlapdf)
```

---

proportional.overlap    *Proportional similarity overlap index*

---

## Description

This index is a measurement of the overlap of proportional similarity (PS) indices between pairs of taxa across samples. As the PS is a niche breadth index that measures the similarity between a taxon's distribution and an environmental parameter, the PS overlap also considers environmental information. A value of 1 indicates perfect overlap, and 0 indicates no overlap.

## Usage

```
proportional.overlap(df, sampleInfo, envInfo, q = 1.65)
```

## Arguments

| | |
|---|---|
| df | A matrix of taxa (rows) by samples (columns) as discrete counts per sample. Col 1 must be a taxon identifier. |
| sampleInfo | A categorical variable identifying which samples correspond to which environment. |
| envInfo | A quantitative variable of some environmental parameter measured per sample. |
| q | A coefficient for the LOQ, set to 1.65. Decreasing or increasing q will affect which taxa are flagged as being below the LOQ. |

## Details

Taxon inputs must be as discrete counts across samples, with a taxon identifier in column 1. The output is a matrix of paired taxon overlap indices. Specifically, Proportional Overlap (PO) indices of PO[1,2] are rows and PO[2,1] are columns, for taxa 1 and 2, however these values are identical. Feinsinger's PS is calculated as per feinsingers.PS.

Proportional overlap is calculated as follows: PO[i,j] = 1 - (|PS[i] - PS[j]|)/(PS[i] + PS[j]) whereby PS[i] is the PS of taxon i and PS[j] is the PS of taxon j.

Taxa below the LOQ are flagged with an asterisk.

This function depends on reshape2 to produce a matrix of overlap comparisons.

## Value

An object of class "data.frame" that gives PO[1,2] as rows and PO[2,1] as columns.

## Note

NA

## Author(s)

Damien Finn

## References

Feinsinger et al. 1981. A simple measure of niche breadth. Ecology 62(1):27-32

## See Also

feinsingers.PS

## Examples

```
data(df)
overlapdf <- df[48:53,]
sampleInfo <- c(rep("R1",10), rep("R2",10), rep("R3",10), rep("R4",10))
pH.grad <- c(2.1, 2.2, 2, 1.9, 2.1, 1.8, 1.9, 2, 2.1, 1.9, 3.5, 3.6, 3.5,
             3.4, 3.6, 3.5, 3.5, 3.4, 3.7, 3.4, 6.6, 6.5, 6.4, 6.8, 7, 6.6,
             6.8, 6.9, 7, 7.1, 8, 8.2, 7.9, 8.1, 7.8, 7.9, 8.3, 8.2, 8.1, 7.9)
res <- proportional.overlap(overlapdf, sampleInfo, pH.grad)
```

# Index