

Package ‘Nonpareil’

March 22, 2018

Title Metagenome Coverage Estimation and Projections for 'Nonpareil'

Type Package

Version 3.3.1

Author Luis M. Rodriguez-R [aut, cre]

Maintainer Luis M. Rodriguez-R <lmrodriguezr@gmail.com>

Description Plot, process, and analyze NPO files produced by 'Nonpareil' <<http://enve-omics.ce.gatech.edu/nonpareil>>.

URL <http://enve-omics.ce.gatech.edu/nonpareil>

Depends R (>= 2.9)

Imports methods

License Artistic-2.0

LazyLoad yes

Encoding UTF-8

NeedsCompilation no

Repository CRAN

Date/Publication 2018-03-22 19:44:27 UTC

R topics documented:

Nonpareil-package	2
+methods	2
Nonpareil.add.curve	3
Nonpareil.antif	3
Nonpareil.col	4
Nonpareil.coverage_factor	4
Nonpareil.curve	5
Nonpareil.Curve-class	6
Nonpareil.curve.batch	7
Nonpareil.f	9
Nonpareil.fit_model	9
Nonpareil.legend	10

Nonpareil.read_data	10
Nonpareil.read_metadata	11
Nonpareil.set	11
Nonpareil.Set-class	12
plot.Nonpareil.Curve	13
plot.Nonpareil.Set	14
predict.Nonpareil.Curve	15
print.Nonpareil.Curve	16
print.Nonpareil.Set	16
summary.Nonpareil.Curve	17
summary.Nonpareil.Set	18
\$-methods	18
\$<-methods	19

Index **20**

Nonpareil-package *Metagenome Coverage Estimation and Projections for 'Nonpareil'*

Description

Plot, process, and analyze NPO files produced by 'Nonpareil' <<http://enve-omics.ce.gatech.edu/nonpareil>>.

Author(s)

Luis M. Rodriguez-R <lmrodriguezr@gmail.com> [aut, cre]

+-methods *~~ Methods for Function + ~~*

Description

~~ Methods for function + ~~

Methods

signature(e1 = "Nonpareil.Set", e2 = "ANY")

Nonpareil.add.curve *Nonpareil add curve*

Description

Adds a 'Nonpareil.Curve' to a 'Nonpareil.Set'

Usage

```
Nonpareil.add.curve(nps, np)
```

Arguments

nps	'Nonpareil.Set' object
np	'Nonpareil.Curve' object

Value

Returns the 'Nonpareil.Set' including the added 'Nonpareil.Curve'

Author(s)

Luis M. Rodriguez-R [aut, cre]

Nonpareil.antif *Nonpareil antif*

Description

Complement function of 'Nonpareil.f'

Usage

```
Nonpareil.antif(y, a, b)
```

Arguments

y	Values of abundance-weighted average coverage
a	Parameter alpha of the gamma CDF.
b	Parameter beta of the gamma CDF.

Value

Estimated sequencing effort.

Author(s)

Luis M. Rodriguez-R [aut, cre]

Nonpareil.col *Nonpareil.col*

Description

Returns the color of the curve

Usage

```
Nonpareil.col(x, alpha = 1)
```

Arguments

x 'Nonpareil.Curve' or 'Nonpareil.Set' object
alpha Alpha level of the color from 0 to 1

Author(s)

Luis M. Rodriguez-R [aut, cre]

Nonpareil.coverage_factor
Nonpareil.coverage_factor

Description

Factor to transform redundancy into coverage (internal function).

Usage

```
Nonpareil.coverage_factor(x)
```

Arguments

x 'Nonpareil.Curve' object

Value

A numeric scalar.

Author(s)

Luis M. Rodriguez-R [aut, cre]

Nonpareil.curve	<i>Nonpareil curve</i>
-----------------	------------------------

Description

Generates a Nonpareil curve from an .npo file

Usage

```
Nonpareil.curve(file, plot = TRUE, label = NA, col = NA, enforce.consistency = TRUE,
  star = 95, correction.factor = TRUE, weights.exp = NA, skip.model = FALSE,
  ...)
```

Arguments

file	Path to the .npo file, containing the read redundancy
plot	Determines if the plot should be produced. If FALSE, it still computes the coverage and the model
label	Name of the dataset. If NA, it is determined by the file name
col	Color of the curve. If NA, a random color is assigned (even if plot=FALSE),
enforce.consistency	If TRUE, it fails verbosely on insufficient data, otherwise it warns about the inconsistencies and attempts the estimations
star	Objective coverage in percentage; i.e., coverage value considered near-complete
correction.factor	Should the overlap-dependent (or kmer-length-dependent) correction factor be applied? If FALSE, redundancy is assumed to equal coverage.
weights.exp	Vector of values to be tested (in order) as exponent of the weights distribution. If the model fails to converge, sometimes manual modifications in this parameter may help. By default (NA), five different values are tested in the following order: For linear sampling, -1.1, -1.2, -0.9, -1.3, -1. For logarithmic sampling (-d option in Nonpareil), 0, 1, -1, 1.3, -1.1, 1.5, -1.5.
skip.model	If set, skips the model estimation altogether.
...	Any additional parameters passed to 'plot.Nonpareil.Curve'

Value

Returns invisibly a 'Nonpareil.Curve' object

Author(s)

Luis M. Rodriguez-R [aut, cre]

Examples

```
# Generate a Nonpareil plot
file <- system.file("extdata", "LakeLanier.npo", package="Nonpareil")
np <- Nonpareil.curve(file)

# Show the estimated values
print(np)

# Predict coverage for 20Gbp
predict(np, 20e9)

# Obtain the Nd diversity index
np$diversity
```

Nonpareil.Curve-class *Nonpareil.Curve S4 class*

Description

A single Nonpareil curve. This object can be produced by ‘Nonpareil.curve’ and supports S4 methods ‘plot’, ‘summary’, ‘print’, and ‘predict’. For additional details, see help for ‘summary.Nonpareil.Curve’

Objects from the Class

Objects can be created by calls of the form `new(Nonpareil.Curve ...)`

Slots

file: (character) Input .npo file
label: (character) Name of the dataset
col: (character) Color of the dataset
L: (numeric) Read length
AL: (numeric) Adjusted read length (same as L for alignment)
R: (numeric) Number of reads
LR: (numeric) Effective sequencing effort used
overlap: (numeric) Minimum read overlap
ksize: (numeric) K-mer size (for kmer kernel only)
log.sample: (numeric) Multiplier of the log-sampling (or zero if linear)
kernel: (character) Read-comparison kernel
version: (character) Nonpareil version used
x.obs: (numeric) Rarefied sequencing effort
x.adj: (numeric) Adjusted rarefied sequencing effort
y.red: (numeric) Rarefied redundancy (observed)

y.cov: (numeric) Rarefied coverage (corrected)
y.sd: (numeric) Standard deviation of rarefied coverage
y.p25: (numeric) Percentile 25 (1st quartile) of rarefied coverage
y.p50: (numeric) Percentile 50 (median) of rarefied coverage
y.p75: (numeric) Percentile 75 (3rd quartile) of rarefied coverage
kappa: (numeric) Dataset redundancy
C: (numeric) Dataset coverage
consistent: (logical) Is the data sufficient for accurate estimation?
star: (numeric) Coverage considered 'nearly complete'
has.model: (logical) Was the model successfully estimated?
warning: (character) Warnings generated on consistency or model fitting
LRstar: (numeric) Projected seq. effort for nearly complete coverage
modelR: (numeric) Pearson's R for the estimated model
diversity: (numeric) Dataset Nd index of sequence diversity
model: (list) Fitted sigmoidal model
call: (call) Call producing this object

Methods

\$ signature(x = "Nonpareil.Curve"): ...
\$<- signature(x = "Nonpareil.Curve"): ...

Author(s)

Luis M. Rodriguez-R [aut, cre]

Nonpareil.curve.batch *Nonpareil curve batch*

Description

Generates a collection of Nonpareil curves (a 'Nonpareil.Set' object) and (optionally) plots all of them in a single canvas

Usage

```
Nonpareil.curve.batch(files, col = NA, labels = NA, plot = TRUE,  
plot.opts = list(), ...)
```

Arguments

files	Vector with the paths to the .npo files
col	Color of the curves (vector). If not passed, values are randomly assigned. Values are recycled
labels	Labels of the curves (vector). If not passed, values are determined by the file-name. Values are recycled
plot	If TRUE, it generates the Nonpareil curve plots
plot.opts	Any parameters accepted by 'plot.Nonpareil.Set' as a list
...	Any additional parameters accepted by 'Nonpareil.curve'

Value

Returns invisibly a 'Nonpareil.Set' object

Author(s)

Luis M. Rodriguez-R [aut, cre]

Examples

```
# Generate a Nonpareil plot with multiple curves
files <- system.file("extdata",
  c("HumanGut.npo", "LakeLanier.npo", "IowaSoil.npo"), package="Nonpareil")
col <- c("orange", "darkcyan", "firebrick4")
nps <- Nonpareil.set(files, col=col,
  plot.opts=list(plot.observed=FALSE, model.lwd=2))

# Show the estimated values
print(nps)

# Show current coverage (as %)
summary(nps)[,"C"]*100

# Extract Nd diversity index
summary(nps)[,"diversity"]

# Extract sequencing effort for nearly complete coverage (in Gbp)
summary(nps)[,"LRstar"]/1e9

# Predict coverage for a sequencing effort of 10Gbp
sapply(nps$np.curves, predict, 10e9)
```

Nonpareil.f	<i>Nonpareil f</i>
-------------	--------------------

Description

Function of the projected model

Usage

Nonpareil.f(x, a, b)

Arguments

x	Values of sequencing effort (in bp)
a	Parameter alpha of the Gamma CDF
b	Parameter beta of the Gamma CDF

Value

Predicted values of abundance-weighted average coverage.

Author(s)

Luis M. Rodriguez-R [aut, cre]

Nonpareil.fit_model	<i>Nonpareil fit model</i>
---------------------	----------------------------

Description

Fit the sigmoidal model to the rarefied coverage

Usage

Nonpareil.fit_model(np, weights.exp)

Arguments

np	'Nonpareil.Curve' object
weights.exp	Numeric; See 'Nonpareil.curve' for details

Author(s)

Luis M. Rodriguez-R [aut, cre]

Nonpareil.legend *Nonpareil legend*

Description

Generates a legend for Nonpareil plots

Usage

```
Nonpareil.legend(np, x, y = 0.3, ...)
```

Arguments

np	A 'Nonpareil.Set' object or a list of 'Nonpareil.Curve' objects
x	X coordinate, or any character string accepted by legend (e.g., 'bottomright').
y	Y coordinate.
...	Any other parameters supported by legend().

Value

Returns invisibly a list, same as 'legend'

Author(s)

Luis M. Rodriguez-R [aut, cre]

Nonpareil.read_data *Nonpareil read data*

Description

Read the data tables and extract direct estimates

Usage

```
Nonpareil.read_data(x, correction.factor)
```

Arguments

x	'Nonpareil.Curve' object
correction.factor	Logical; see 'Nonpareil.curve' for details

Author(s)

Luis M. Rodriguez-R [aut, cre]

Nonpareil.read_metadata
Nonpareil read metadata

Description

Read the metadata headers

Usage

```
Nonpareil.read_metadata(x)
```

Arguments

x ‘Nonpareil.Curve‘ object

Author(s)

Luis M. Rodriguez-R [aut, cre]

Nonpareil.set *Nonpareil set*

Description

Generates a collection of Nonpareil curves (a ‘Nonpareil.Set‘ object) and (optionally) plots all of them in a single canvas

Usage

```
Nonpareil.set(files, col = NA, labels = NA, plot = TRUE, plot.opts = list(),
  ...)
```

Arguments

files	Vector with the paths to the .npo files
col	Color of the curves (vector). If not passed, values are randomly assigned. Values are recycled
labels	Labels of the curves (vector). If not passed, values are determined by the file-name. Values are recycled
plot	If TRUE, it generates the Nonpareil curve plots
plot.opts	Any parameters accepted by ‘plot.Nonpareil.Set‘ as a list
...	Any additional parameters accepted by ‘Nonpareil.curve‘

Value

Returns invisibly a ‘Nonpareil.Set’ object

Author(s)

Luis M. Rodriguez-R [aut, cre]

Examples

```
# Generate a Nonpareil plot with multiple curves
files <- system.file("extdata",
  c("HumanGut.npo", "LakeLanier.npo", "IowaSoil.npo"), package="Nonpareil")
col <- c("orange", "darkcyan", "firebrick4")
nps <- Nonpareil.set(files, col=col,
  plot.opts=list(plot.observed=FALSE, model.lwd=2))

# Show the estimated values
print(nps)

# Show current coverage (as %)
summary(nps)[,"C"]*100

# Extract Nd diversity index
summary(nps)[,"diversity"]

# Extract sequencing effort for nearly complete coverage (in Gbp)
summary(nps)[,"LRstar"]/1e9

# Predict coverage for a sequencing effort of 10Gbp
sapply(nps$np.curves, predict, 10e9)
```

Nonpareil.Set-class *Nonpareil.Set S4 class*

Description

Collection of ‘Nonpareil.Curve’ objects. This object can be produced by ‘Nonpareil.curve.batch’ and supports S4 methods ‘plot’, ‘summary’, and ‘print’

Objects from the Class

Objects can be created by calls of the form `new(Nonpareil.Set ...)`

Slots

np.curves: (list) List of ‘Nonpareil.Curve’ objects
call: (call) Call producing this object

Methods

```
+ signature(e1 = "Nonpareil.Set", e2 = "ANY"): ...
$ signature(x = "Nonpareil.Set"): ...
$<- signature(x = "Nonpareil.Set"): ...
```

Author(s)

Luis M. Rodriguez-R [aut, cre]

plot.Nonpareil.Curve *plot Nonpareil Curve*

Description

Plot a 'Nonpareil.Curve' object

Usage

```
## S3 method for class 'Nonpareil.Curve'
plot(x, col = NA, add = FALSE, new = !add, plot.observed = TRUE,
     plot.model = TRUE, plot.dispersion = FALSE, plot.diversity = TRUE,
     xlim = c(1000, 1e+13), ylim = c(1e-06, 1), main = paste("Nonpareil Curve for",
     x$label), xlab = "Sequencing effort (bp)", ylab = "Estimated Average Coverage",
     curve.lwd = 2, curve.alpha = 0.4, model.lwd = 1, model.alpha = 1,
     log = "x", arrow.length = 0.05, arrow.head = arrow.length,
     ...)
```

Arguments

x	'Nonpareil.Curve' object to plot
col	Color of the curve. If passed, it overrides the colors set in the 'Nonpareil.Curve' object
add	If TRUE, it attempts to use a previous (active) canvas to plot the curve
new	Inverse of 'add'
plot.observed	Indicates if the observed (rarefied) coverage is to be plotted
plot.model	Indicates if the fitted model is to be plotted
plot.dispersion	Indicates if (and how) dispersion of the replicates should be plotted. Supported values are: FALSE: no dispersion is plotted (default), 'sd': one standard deviation around the mean, 'ci95': 95% confidence interval, 'ci90': 90% confidence interval, 'ci50': 50% confidence interval, 'iq': Inter-quartile range
plot.diversity	If TRUE, the diversity estimate is plotted as a small arrow below the Nonpareil curve
xlim	Limits of the sequencing effort (X-axis)

<code>ylim</code>	Limits of the coverage (Y-axis)
<code>main</code>	Title of the plot
<code>xlab</code>	Label of the X-axis
<code>ylab</code>	Label of the Y-axis
<code>curve.lwd</code>	Line width of the rarefied coverage
<code>curve.alpha</code>	Alpha value (from 0 to 1) of the rarefied coverage
<code>model.lwd</code>	Line width of the model
<code>model.alpha</code>	Alpha value (from 0 to 1) of the model
<code>log</code>	Axis to plot in logarithmic scale. Supported values are: 'x': sequencing effort (default), 'y': coverage, 'xy': both logarithmic, or '': both linear
<code>arrow.length</code>	If <code>plot.diversity=TRUE</code> , it determines the length of the arrow to display the diversity (as a fraction of the <code>ylim</code> range).
<code>arrow.head</code>	If <code>plot.diversity=TRUE</code> , it determines the length of the arrow head to display the diversity index (in inches).
<code>...</code>	Additional graphical parameters

Value

Returns invisibly a 'Nonpareil.Curve' object (same as 'x' input). For additional details see help for 'summary.Nonpareil.Curve'

Author(s)

Luis M. Rodriguez-R [aut, cre]

`plot.Nonpareil.Set` *plot Nonpareil Set*

Description

Plot a 'Nonpareil.Set' object

Usage

```
## S3 method for class 'Nonpareil.Set'
plot(x, col = NA, labels = NA, main = "Nonpareil Curves",
     legend.opts = list(), ...)
```

Arguments

x	'Nonpareil.Set' object to plot
col	Color of the curves (vector). If passed, it overrides the colors set in the 'Nonpareil.Curve' objects. Values are recycled
labels	Labels of the curves (vector). If passed, it overrides the labels set in the 'Nonpareil.Curve' objects. Values are recycled
main	Title of the plot
legend.opts	Any additional parameters passed to 'Nonpareil.legend'. If FALSE, the legend is not displayed
...	Any additional parameters passed to 'plot.Nonpareil.Curve'

Value

Returns invisibly a 'Nonpareil.Set' object (same as 'x' input).

Author(s)

Luis M. Rodriguez-R [aut, cre]

predict.Nonpareil.Curve
predict Nonpareil Curve

Description

Predict the coverage for a given sequencing effort

Usage

```
## S3 method for class 'Nonpareil.Curve'
predict(object, lr = object$LR, ...)
```

Arguments

object	'Nonpareil.Curve' object
lr	Sequencing effort for the prediction (in bp)
...	Additional parameters ignored

Value

Returns the expected coverage at the given sequencing effort.

Author(s)

Luis M. Rodriguez-R [aut, cre]

`print.Nonpareil.Curve` *print Nonpareil Curve*

Description

Prints and returns invisibly a summary of the ‘Nonpareil.Curve’ results

Usage

```
## S3 method for class 'Nonpareil.Curve'  
print(x, ...)
```

Arguments

<code>x</code>	‘Nonpareil.Set’ object
<code>...</code>	Additional parameters ignored

Value

Returns the summary invisibly. See help for ‘summary.Nonpareil.Curve’ for additional information.

Author(s)

Luis M. Rodriguez-R [aut, cre]

`print.Nonpareil.Set` *print Nonpareil Set*

Description

Prints and returns invisibly a summary of the ‘Nonpareil.Set’ results

Usage

```
## S3 method for class 'Nonpareil.Set'  
print(x, ...)
```

Arguments

<code>x</code>	‘Nonpareil.Set’ object
<code>...</code>	Additional parameters ignored

Value

Returns the summary invisibly. See help for ‘summary.Nonpareil.Curve’ and ‘summary.Nonpareil.Set’ for additional information

Author(s)

Luis M. Rodriguez-R [aut, cre]

summary.Nonpareil.Curve
summary Nonpareil Curve

Description

Returns a summary of the Nonpareil.Curve results

Usage

```
## S3 method for class 'Nonpareil.Curve'  
summary(object, ...)
```

Arguments

object	'Nonpareil.Curve' object
...	Additional parameters ignored

Value

Returns a matrix with the following values for the dataset:

kappa: "Redundancy" value of the entire dataset.

C: Average coverage of the entire dataset.

LRstar: Estimated sequencing effort required to reach the objective average coverage (star, 95% by default).

LR: Actual sequencing effort of the dataset.

modelR: Pearson's R coefficient between the rarefied data and the projected model.

diversity: Nonpareil sequence-diversity index (Nd). This value's units are the natural logarithm of the units of sequencing effort (log-bp), and indicates the inflection point of the fitted model for the Nonpareil curve. If the fit doesn't converge, or the model is not estimated, the value is zero (0).

Author(s)

Luis M. Rodriguez-R [aut, cre]

```
summary.Nonpareil.Set summary Nonpareil Set
```

Description

Returns a summary of the Nonpareil.Set results

Usage

```
## S3 method for class 'Nonpareil.Set'
summary(object, ...)
```

Arguments

```
object      'Nonpareil.Set' object
...         Additional parameters ignored
```

Value

Returns a matrix with different values for each dataset. For additional details on the values returned, see help for 'summary.Nonpareil.Curve'

Author(s)

Luis M. Rodriguez-R [aut, cre]

```
$-methods      ~~ Methods for Function $ ~~
```

Description

~~ Methods for function \$ ~~

Methods

```
signature(x = "Nonpareil.Curve")
signature(x = "Nonpareil.Set")
```

`$<--methods` *~~ Methods for Function \$<- ~~*

Description

~~ Methods for function \$<- ~~

Methods

`signature(x = "Nonpareil.Curve")`

`signature(x = "Nonpareil.Set")`

Index

- *Topic **\textasciitilde\textasciitilde**
other possible keyword(s)
\textasciitilde\textasciitilde
 - +~methods, [2](#)
 - \$~methods, [18](#)
 - \$<--methods, [19](#)
- *Topic **methods**
 - +~methods, [2](#)
 - \$~methods, [18](#)
 - \$<--methods, [19](#)
- *Topic **package**
 - Nonpareil-package, [2](#)
 - +, Nonpareil.Set (Nonpareil.Set-class), [12](#)
 - +, Nonpareil.Set, ANY-method (+~methods), [2](#)
 - +~methods, [2](#)
 - \$, Nonpareil.Curve-method (\$~methods), [18](#)
 - \$, Nonpareil.Set-method (\$~methods), [18](#)
 - \$~methods, [18](#)
 - \$<--methods, [19](#)
 - \$<- , Nonpareil.Curve-method ([\\$<--methods](#)), [19](#)
 - \$<- , Nonpareil.Set-method ([\\$<--methods](#)), [19](#)
- Nonpareil (Nonpareil-package), [2](#)
- Nonpareil-package, [2](#)
- Nonpareil.add.curve, [3](#)
- Nonpareil.antif, [3](#)
- Nonpareil.col, [4](#)
- Nonpareil.coverage_factor, [4](#)
- Nonpareil.Curve (Nonpareil.Curve-class), [6](#)
- Nonpareil.curve, [5](#)
- Nonpareil.Curve-class, [6](#)
- Nonpareil.curve.batch, [7](#)
- Nonpareil.f, [9](#)
- Nonpareil.fit_model, [9](#)
- Nonpareil.legend, [10](#)
- Nonpareil.read_data, [10](#)
- Nonpareil.read_metadata, [11](#)
- Nonpareil.Set (Nonpareil.Set-class), [12](#)
- Nonpareil.set, [11](#)
- Nonpareil.Set-class, [12](#)
- plot.Nonpareil.Curve, [13](#)
- plot.Nonpareil.Set, [14](#)
- predict.Nonpareil.Curve, [15](#)
- print.Nonpareil.Curve, [16](#)
- print.Nonpareil.Set, [16](#)
- summary.Nonpareil.Curve, [17](#)
- summary.Nonpareil.Set, [18](#)