

# Package ‘neutralitytestr’

May 21, 2018

**Title** Test for a Neutral Evolutionary Model in Cancer Sequencing Data

**Version** 0.0.2

**Description** Package takes frequencies of mutations as reported by high throughput sequencing data from cancer and fits a theoretical neutral model of tumour evolution. Package outputs summary statistics and contains code for plotting the data and model fits. See Williams et al 2016 <doi:10.1038/ng.3489> and Williams et al 2017 <doi:10.1101/096305> for the details of the method.

**Depends** R (>= 3.4)

**License** MIT + file LICENSE

**Encoding** UTF-8

**LazyData** true

**Imports** dplyr, ggplot2, scales, pracma, ggpmisc, cowplot

**Suggests** knitr, rmarkdown, testthat

**VignetteBuilder** knitr

**URL** <https://github.com/marcjwilliams1/neutralitytestr>

**BugReports** <https://github.com/marcjwilliams1/neutralitytestr/issues>

**RoxygenNote** 6.0.1

**NeedsCompilation** no

**Author** Marc Williams [aut, cre]

**Maintainer** Marc Williams <marcjwilliams1@gmail.com>

**Repository** CRAN

**Date/Publication** 2018-05-21 15:11:24 UTC

## R topics documented:

lsq_plot . . . . .	2
neutralitytest . . . . .	2
neutralitytestr . . . . .	3
normalized_plot . . . . .	4

plot_all . . . . .	4
VAFneutral . . . . .	5
VAFselection . . . . .	5
vaf_histogram . . . . .	6

<b>Index</b>	<b>7</b>
--------------	----------

---

lsq_plot	<i>Plot cumulative distribution lsq_plot Plots the cumulative distribution of the data as well as the best fit linear model line.</i>
----------	---

---

### Description

Plot cumulative distribution lsq\_plot Plots the cumulative distribution of the data as well as the best fit linear model line.

### Usage

```
lsq_plot(object)
```

### Arguments

object            neutrality test object

### Value

ggplot object.

### Examples

```
lsq_plot(neutralitytest(VAFselection, fmin = 0.1, fmax = 0.25))
```

---

neutralitytest	<i>Testing for neutrality on cancer sequencing data</i>
----------------	---

---

### Description

neutralitytest returns a neutralitytest object which contains the result of various test statistics to test for neutrality as described in Williams et al. Nature Genetics 2018.

### Usage

```
neutralitytest(VAF, fmin = 0.1, fmax = 0.25, read_depth = NULL, rho = 0,
  cellularity = 1, ploidy = 2)
```

## Arguments

VAF	Vector of variant allele frequencies (VAFs) from a deep sequencing experiment, numbers should be between 0 and 1
fmin	Minimum VAF of integration range, default is 0.1
fmax	Maximum VAF of integration range, default is 0.25
read_depth	Read depth of sample, if this is specified it will be used to calculate an appropriate integration range. default is NULL in which case the default or inputted fmin and fmax will be used.
rho	Overdispersion of sample if known, default is 0.0. Will be used to calculate integration range if read_depth != NULL
cellularity	Cellularity of sample, default is 1.0. Will be used to calculate integration range if read_depth != NULL
ploidy	Ploidy of the genome, default is 2. Ideally mutations should be filtered for this ploidy before running the test. Will be used to calculate integration range if read_depth != NULL

## Value

neutralitytest object which contains test statistics which tests if the sequencing data is consistent a neutral evolutionary model. Test statistics are area between theoretical and empirical curves, kolmogorov distance, mean distance and R<sup>2</sup> statistics from linear model fit. Also returns an estimate of the mutation rate per tumour tumour doubling, the raw VAFs and cumulative distribution

## Examples

```
neutralitytest(runif(100))
neutralitytest(VAFselection, fmin = 0.1, fmax = 0.25)
neutralitytest(VAFneutral, read_depth = 100.0, cellularity = 0.8)
```

---

neutralitytestr      neutralitytestr *package*

---

## Description

Package to test a neutral evolutionary model on deep sequencing data.

## Details

See the README on [GitHub](#)

---

normalized_plot	<i>Plot normalized cumulative distribution (normalized) cumulative distribution of the data as well as the theoretical expectation from a neutral evolutionary model.</i>	normalized_plot	<i>Plots the (normalized) cumulative distribution of the data as well as the theoretical expectation from a neutral evolutionary model.</i>
-----------------	---	-----------------	---

---

**Description**

Plot normalized cumulative distribution normalized\_plot Plots the (normalized) cumulative distribution of the data as well as the theoretical expectation from a neutral evolutionary model.

**Usage**

```
normalized_plot(object)
```

**Arguments**

object            neutrality test object

**Value**

ggplot object.

**Examples**

```
normalized_plot(neutralitytest(VAFselection, fmin = 0.1, fmax = 0.25))
```

---

plot_all	<i>Plot all plots in the package and make composite figure. Plots histogram, linear model best fit plot and normalized plot and plot and makes composite figure.</i>	plot_all	<i>Plots histogram, linear model best fit plot and normalized plot and plot and makes composite figure.</i>
----------	--	----------	---

---

**Description**

Plot all plots in the package and make composite figure. plot\_all Plots histogram, linear model best fit plot and normalized plot and plot and makes composite figure.

**Usage**

```
plot_all(object)
```

**Arguments**

object            neutrality test object

**Value**

ggplot object.

**Examples**

```
plot_all(neutralitytest(VAFselection, fmin = 0.1, fmax = 0.25))
```

---

VAFneutral	<i>Synthetic sequencing data generated from a evolutionary based cancer simulation.</i>
------------	---

---

**Description**

This data is generated from a neutral evolutionary model where all subclonal mutations are neutral.

**Usage**

```
VAFneutral
```

**Format**

A vector with variant allele frequencies (VAFs) ranging from 0 to 1

**Source**

Generated using cancer sequencing simulation <https://github.com/marcjwilliams1/CancerSeqSim.jl>

---

VAFselection	<i>Synthetic sequencing data generated from a evolutionary based cancer simulation.</i>
--------------	---

---

**Description**

This data is generated from an evolutionary model where there is one subclonal population and all other mutations are neutral passengers.

**Usage**

```
VAFselection
```

**Format**

A vector with variant allele frequencies (VAFs) ranging from 0 to 1

**Source**

Generated using cancer sequencing simulation <https://github.com/marcjwilliams1/CancerSeqSim.jl>

---

vaf_histogram	<i>Plot VAF histogram vaf_histogram Plots a histogram of the variant allele frequencies.</i>
---------------	--

---

**Description**

Plot VAF histogram vaf\_histogram Plots a histogram of the variant allele frequencies.

**Usage**

```
vaf_histogram(object)
```

**Arguments**

object            neutrality test object

**Value**

ggplot object.

**Examples**

```
vaf_histogram(neutralitytest(VAFselection, fmin = 0.1, fmax = 0.25))
```

# Index

## \*Topic **datasets**

VAFneutral, [5](#)

VAFselection, [5](#)

lsq\_plot, [2](#)

neutralitytest, [2](#)

neutralitytestr, [3](#)

neutralitytestr-package

(neutralitytestr), [3](#)

normalized\_plot, [4](#)

plot\_all, [4](#)

vaf\_histogram, [6](#)

VAFneutral, [5](#)

VAFselection, [5](#)