

Package: neopeptides (via r-universe)

July 21, 2024

Title Calculate and Explore Property Indices of Neopeptides

Version 0.1.0

Description Includes functions to calculate and explore several property indices for neopeptides, which are abnormal peptides generated from genome.

License MIT + file LICENSE

URL <https://github.com/ShixiangWang/neopeptides>

BugReports <https://github.com/ShixiangWang/neopeptides/issues>

Depends R (>= 3.5)

Imports Biostrings, data.table, magrittr, parallel, Peptides, psych, purrr, R.utils, stringr, yaml

biocViews SomaticMutation,ImmunoOncology,Alignment

Encoding UTF-8

LazyData true

Roxygen list(markdown = TRUE)

RoxygenNote 7.1.2

Repository <https://shixiangwang.r-universe.dev>

RemoteUrl <https://github.com/ShixiangWang/neopeptides>

RemoteRef HEAD

RemoteSha 6aff6c735d0ab599733d95248fe54ac3baee5b6a

Contents

calc_dissimilarity	2
calc_geometric_mean	3
calc_harmonic_mean	4
calc_iedb_score	4
find_path	6
install_database	6
set_blast_path	7

Index	8
--------------	----------

calc_dissimilarity	<i>Calculate Dissimilarity Value to Reference Proteome for Peptides</i>
--------------------	---

Description

Calculate Dissimilarity Value to Reference Proteome for Peptides

Usage

```
calc_dissimilarity(
  pep,
  db = "human",
  k_val = 4.86936,
  a_val = 32,
  fill = NA_real_,
  use_blastp_short = TRUE,
  threads = parallel::detectCores(),
  tmp_dir = file.path(tempdir(), "neopeptides"),
  clean_tmp = TRUE
)
```

Arguments

pep	a vector of peptides.
db	an available database, can be one of "human" and "mouse". Or you can set this to specify the (fasta) database file to be searched instead of using standard database. A blast database will be created if it does not exist.
k_val	numeric. Steepness of sigmoidal curve at k. Default 4.86936, the value used in the analysis of Van Allen, Snyder, Rizvi, Riaz, and Hellmann datasets. See reference.
a_val	numeric. Optionally can be "mean" to use mean alignment for nmers passed. Horizontal displacement of partition function. Default is 32, based on max_SW of 75 million 8-15mers from the five clinical datasets against human, if using max_SW, use 52. This value may not be meaningful for murine alignment so use with care. See reference.
fill	a numeric value for filling default NA value when no blast result.
use_blastp_short	default is TRUE, optimize blast for <30 AA by setting blastp flag -task as 'blastp-short'. If FALSE, use default task mode, i.e. 'blastp'.
threads	number of threads to run.
tmp_dir	path for storing temp files.
clean_tmp	if TRUE, remove temp directory.

Value

Data table of dissimilarity values (to the non-mutated proteome).

- peptide - input peptide
- dissimilarity - dissimilarity value

References

Richman LP, Vonderheide RH, and Rech AJ. "Neoantigen dissimilarity to the self-proteome predicts immunogenicity and response to immune checkpoint blockade." *Cell Systems* 9, 375-382.E4, (2019).

See Also

[set_blast_path](#) for setting blast path, [install_database](#) for installing database.

Examples

```
## Not run:
calc_dissimilarity("AAAAAAAA")
calc_dissimilarity("MRLVDRRWA")
calc_dissimilarity("VRLVDRRWA")
calc_dissimilarity("MTEYKLVVVGAGDVGKSALTI")
calc_dissimilarity("MTEYKLVVVGAGDVGKSALTI", db = "mouse")
calc_dissimilarity(c("MTEYKLVVVGAGDVGKSALTIQLIQNHVFDEYDP", "MTEYKLVVVG"))

## End(Not run)
```

calc_geometric_mean	<i>Calculate the Geometric Mean</i>
---------------------	-------------------------------------

Description

Calculate the Geometric Mean

Usage

```
calc_geometric_mean(x, na.rm = TRUE)
```

Arguments

x	a vector or data.frame
na.rm	remove NA values before processing

Value

The geometric mean(s)

Examples

```
x <- 1:5  
calc_geometric_mean(x)
```

calc_harmonic_mean	<i>Calculate the Harmonic Mean</i>
--------------------	------------------------------------

Description

Calculate the Harmonic Mean

Usage

```
calc_harmonic_mean(x, na.rm = TRUE, zero = TRUE)
```

Arguments

x	a vector, matrix, or data.frame
na.rm	na.rm=TRUE remove NA values before processing
zero	If TRUE, then if there are any zeros, return 0, else, return the harmonic mean of the non-zero elements

Value

The harmonic mean(s)

Examples

```
x <- 1:5  
calc_harmonic_mean(x)
```

calc_iedb_score	<i>Calculate IEDB Score for Peptides</i>
-----------------	--

Description

Calculate IEDB Score for Peptides

Usage

```
calc_iedb_score(
  pep,
  db = "human",
  fill = NA_real_,
  use_blastp_short = TRUE,
  threads = parallel::detectCores(),
  tmp_dir = file.path(tempdir(), "neopeptides"),
  clean_tmp = TRUE
)
```

Arguments

pep	a vector of peptides.
db	an available database, can be one of "human" and "mouse". Or you can set this to specify the (fasta) database file to be searched instead of using standard database. A blast database will be created if it does not exist.
fill	a numeric value for filling default NA value when no blast result.
use_blastp_short	default is TRUE, optimize blast for <30 AA by setting blastp flag -task as 'blastp-short'. If FALSE, use default task mode, i.e. 'blastp'.
threads	number of threads to run.
tmp_dir	path for storing temp files.
clean_tmp	if TRUE, remove temp directory.

Value

Data table of IEDB scores.

- peptide - input peptide
- iedb_score - IEDB score
- annotation - IEDB annotation info

See Also

[set_blast_path](#) for setting blast path, [install_database](#) for installing database.

Examples

```
## Not run:
calc_iedb_score("AAAAAAAA")
calc_iedb_score("MTEYKLVVVGAGDVGKSALTIQLIQNHFVDEYDP")
calc_iedb_score("MTEYKLVVVGAGDVGKSALTIQLIQNHFVDEYDP", db = "mouse")
calc_iedb_score(c("MTEYKLVVVGAGDVGKSALTIQLIQNHFVDEYDP", "MTEYKLVVVG"))

## End(Not run)
```

find_path	<i>Find Entry Path from Environment Variable or Config file</i>
-----------	---

Description

Firstly environment variable PATH will be searched, if not found, configure file then will be used to search the entry.

Usage

```
find_path(entry)
```

Arguments

entry	a program or blast database.
-------	------------------------------

Value

a path.

See Also

[set_blast_path](#) for setting blast path, [install_database](#) for installing database.

install_database	<i>Install IEDB and Proteome Reference Database</i>
------------------	---

Description

At default, configure file is '~/.neopeptide/config.yml', you can set by options(neopeptides.config="new_path.yml"). Please remember, if you change the configure file, you have to set it per R session.

Usage

```
install_database(
  db_path = normalizePath("~/neopeptide/db", mustWork = FALSE),
  force = FALSE,
  species = c("human", "mouse"),
  data_type = c("IEDB", "Proteome")
)
```

Arguments

db_path	a path to store the databases.
force	if TRUE, force to re-download and re-install.
species	can be 'human' and 'mouse'. Default select them all.
data_type	can be 'IEDB' and 'Proteome'. Default select them all.

Value

Nothing.

See Also

[set_blast_path](#), [calc_iedb_score](#), [calc_dissimilarity](#)

Examples

```
## Not run:  
# Default all databases will be downloaded and installed.  
install_database()  
  
## End(Not run)
```

set_blast_path

Set Path to Blast

Description

At default, configure file is '~/.neopeptide/config.yml', you can set by `options(neopeptides.config="new_path.yml")`. Please remember, if you change the configure file, you have to set it per R session.

Usage

```
set_blast_path(path)
```

Arguments

path	a directory containing executable Blast program. Installing it by conda is recommended.
------	---

Value

Nothing

See Also

[find_path](#), [install_database](#)

Examples

```
## Not run:  
set_blast_path("/Users/wsx/anaconda3/bin/")  
  
## End(Not run)
```

Index

`calc_dissimilarity`, [2](#), [7](#)
`calc_geometric_mean`, [3](#)
`calc_harmonic_mean`, [4](#)
`calc_iedb_score`, [4](#), [7](#)

`find_path`, [6](#), [7](#)

`install_database`, [3](#), [5](#), [6](#), [6](#), [7](#)

`set_blast_path`, [3](#), [5–7](#), [7](#)