

Package ‘tcr1’

November 4, 2021

Type Package

Title Association analysis of T-cell receptor repertoire and clinical phenotypes

Version 0.1.0

Author Meiling Liu

Maintainer Meiling Liu <mliu@fredhutch.org>

Description This package provides a sequence of functions for analyzing the link between the T-cell receptor repertoire and clinical phenotypes.

Depends R (>= 3.5.0), Biostrings

License LGPL(>=2.0)

Encoding UTF-8

LazyData true

RoxygenNote 7.1.2

R topics documented:

AAfreq	1
example.data	2
seqhom	3
tcr1	3
TCRL_bin	4
TCRseq_bin	5
Index	6

AAfreq	<i>Compute amino acid frequency in TCR repertoire.</i>
--------	--

Description

Compute amino acid frequency in TCR repertoire.

Usage

AAfreq(Subject.ID, AAs, Abundance)

Arguments

<code>Subject.ID</code>	a vector of subject IDs.
<code>AAseq</code>	A vector of amino acid sequences. The order of sequences should be consistent with <code>Subject.ID</code> .
<code>Abundance</code>	A numeric vector of abundance. The order of abundances should be consistent with <code>Subject.ID</code> and <code>AAseq</code> .

Value

out an amino acid matrix with each row represents a subject and each column represents an amino acid.

Examples

```
data("Example.data")

## extract features
fR <- AAfreq(TCRdat[,1],TCRdat[,2],as.numeric(TCRdat[,3]))
```

<code>example.data</code>	<i>A list containing TCR data, continuous reponse and covariate variables.</i>
---------------------------	--

Description

A simulated dataset list containing the TCR information, covariate and response variables of 20 patients.

Usage

```
example.data
```

Format

A dataset containing the TCR information, covariate and response variables of 20 patients.

TCRdat The TCR information with three columns, `Subject.ID`, `AAseq`, and `Abundance`, indicating the patient ID, amino acid sequence, and corresponding abundance.

Y A vector of continous response.

X A matrix of covariate variables.

W A vector of Kyte and Doolittle hydrophobicity information.

seqhom	<i>Compute TCR repertoire homology between subjects.</i>
--------	--

Description

Compute TCR repertoire homology between subjects.

Usage

```
seqhom(Subject.ID, AAseq, Abundance, substitutionMatrix)
```

Arguments

Subject.ID	a vector of subject IDs.
AAseq	A vector of amino acid sequences. The order of sequences should be consistent with Subject.ID.
Abundance	A numeric vector of abundance. The order of abundances should be consistent with Subject.ID and AAseq.
substitutionMatrix	A character, representing the fixe substitution scores for an alignment. It can take value from following options: BLOSUM45, BLOSUM50, BLOSUM62, BLOSUM80, BLOSUM100, PAM30, PAM40, PAM70, PAM120, and PAM250.

Value

S a homology matrix.

Examples

```
data("Example.data")

## homology matrix
S <- seqhom(TCRdat[,1],TCRdat[,2],as.numeric(TCRdat[,3]),'BLOSUM62')
```

tcrl	<i>tcrl: Association analysis of T-cell receptor repertoire and clinical phenotypes</i>
------	---

Description

This package provides a sequence of functions for analyzing the link between the T-cell receptor repertoire and clinical phenotypes.

TCRL_bin

Score test for TCR repertoirre and phenotypes

Description

Score test for TCR repertoirre and phenotypes

Usage

```
TCRL_bin(Y, X, fR, W, S)
```

```
TCRL_cont(Y, X, fR, W, S)
```

Arguments

Y	a vector, Y should be continous or binary variable.
X	a covariate matrix, each row represents a subject.
fR	a variant matrix, each row represents a subject, each column represents an extracted feature of TCR repertoirre.
W	a feature character matrix, each row represents an extracted feature, each column represents a variant. The row number of W should be consistent with the column number of fR.
S	a homology matrix, which represents the correlation between subjects. It can be caucalated by using function seqhom.

Value

fix.effect p-value for testing the fixed variant effect.

random.effect p-value for testing the random variant effect.

Overall.pval overall p-value for testing the association between TCR repertoirre and phenotype. It combines fix.effect and rand.effect by using Fisher's procedure.

Examples

```
data("Example.data")
## extract features
fR <- AAFreq(TCRdat[,1],TCRdat[,2],as.numeric(TCRdat[,3]))

## homology matrix
S <- seqhom(TCRdat[,1],TCRdat[,2],as.numeric(TCRdat[,3]),'BLOSUM62')

## include both fixed and random effect
TCRL.cont(Y,X,fR,W,S)
```

TCRseq_bin

Score test for TCR repertoirre and phenotypes

Description

Score test for TCR repertoirre and phenotypes

Usage

```
TCRseq_bin(Y, X, S)
```

```
TCRseq_cont(Y, X, S)
```

Arguments

Y	a vector, Y should be continous variable.
X	a covariate matrix, each row represents a subject.
S	a homology matrix, which represents the correlation between subjects. It can be caculated by using function seqhom.

Value

Overall.pval p-value for testing the association between TCR repertoirre and phenotype.

Examples

```
data("Example.data")

## extract features
fR <- AAfreq(TCRdat[,1],TCRdat[,2],as.numeric(TCRdat[,3]))

## homology matrix
S <- seqhom(TCRdat[,1],TCRdat[,2],as.numeric(TCRdat[,3]),'BLOSUM62')

## include random effect only
TCRseq.cont(Y,X,S)
```

Index

* **datasets**

example.data, [2](#)

AAfreq, [1](#)

example.data, [2](#)

seqhom, [3](#)

tcrl, [3](#)

TCRL_bin, [4](#)

TCRL_cont (TCRL_bin), [4](#)

TCRseq_bin, [5](#)

TCRseq_cont (TCRseq_bin), [5](#)