

Package ‘aaSEA’

November 9, 2019

Type Package

Title Amino Acid Substitution Effect Analyser

Version 1.1.0

Author Raja Sekhara Reddy D.M

Maintainer Raja Sekhara Reddy D.M <raja.duvvuru@gmail.com>

Description

Given a protein multiple sequence alignment, it is daunting task to assess the effects of substitutions along sequence length. 'aaSEA' package is intended to help researchers to rapidly analyse property changes caused by single, multiple and correlated amino acid substitutions in proteins. Methods for identification of co-evolving positions from multiple sequence alignment are as described in : Pelé et al., (2017) <doi:10.4172/2379-1764.1000250>.

Depends R(>= 3.4.0)

Imports DT(>= 0.4), networkD3(>= 0.4), shiny(>= 1.0.5),
shinydashboard(>= 0.7.0), magrittr(>= 1.5), Bios2cor(>= 2.0),
seqinr(>= 3.4-5), plotly(>= 4.7.1), Hmisc(>= 4.1-1)

License GPL-3

Encoding UTF-8

LazyData true

RoxygenNote 6.1.1

Suggests knitr, rmarkdown

VignetteBuilder knitr

NeedsCompilation no

Repository CRAN

Date/Publication 2019-11-09 16:20:02 UTC

R topics documented:

AAindex	2
corSubFile	3
Cruciani	3

Fasgai	4
getAASub	4
getCorPropChange	5
getCorSites	5
getPropChange	6
getPropCorr	7
getTopSub	7
Kidera	8
matEncode	9
plotCorNet	9
plotCorSubChanges	10
plotMultiSubChange	11
plotSingleSubChange	11
startSeaShiny	12

Index	13
--------------	-----------

AAindex	<i>A data frame of 533 amino acid properties from AAindex.</i>
---------	--

Description

A data frame of 533 amino acid properties.

Usage

AAindex

Format

A data frame with 533 rows and 20 variables:

name amino acid property as in AAindex

A Alanine, Ala, A

C Cysteine, Cys, C

D Aspartate, Asp, D

E Glutamate, Glu, E

F Phenylalanine, Phe, F

G Glycine, Gly, G

H Histidine, His, H

I Isoleucine, Ile, I

K Lysine, Lys, K

L Leucine, Leu, L

M Methionine, Met, M

N Asparagine, Asn, N

P Proline, Pro, P
Q Glutamine, Gln, Q
R Arginine, Arg, R
S Serine, Ser, S
T Threonine, Thr, T
V Valine, Val, V
W Tryptophan, Trp, W
Y Tyrosine, Tyr, Y ...

corSubFile *A data frame of correlated sites.*

Description

This file is output of 'getTopSub' function with subset of alignment as input. This subset alignment is output of 'getCorSite' function

Usage

corSubFile

Format

A data frame with 2 Columns representing positional pairs:

Pos1 Position one prefixed with Wt amino acid and suffixed with substitution

Pos2 Position two prefixed with Wt amino acid and suffixed with substitution ...

Cruciani *A data frame of 3 Cruciani properties for 20 amino acids.*

Description

A data frame of 3 Cruciani properties for 20 amino acids.

Usage

Cruciani

Format

A data frame with 20 rows and 3 columns representing amino acids and Cruciani properties:

Names of amino acids

PP1 Polarity

PP2 Hydrophobicity

PP3 H-bonding ...

Fasgai	<i>A data frame of six Fasgai vectors for 20 amino acids.</i>
--------	---

Description

A data frame of six Fasgai vectors for 20 amino acids.

Usage

```
Fasgai
```

Format

A data frame with 20 rows and 6 columns representing amino acids and Fasgai vectors:

Names of amino acids

F1 Hydrophobicity index

F2 Alpha and turn propensities

F3 Bulky properties

F4 Compositional characteristic index

F5 Local flexibility

F6 Electronic properties ...

getAASub	<i>Get amino acid substitutions from multiple sequence alignment</i>
----------	--

Description

Get amino acid substitutions from multiple sequence alignment

Usage

```
getAASub(fileLoc)
```

Arguments

fileLoc exact location of multiple sequence alignment file in "FASTA" format

Value

Returns a list of two data frames 1.Single substitutions 2.Multiple substitutions

Examples

```
file = system.file("extdata", "linB_Prot_ali.fasta", package = "aaSEA")
getAASub(fileLoc = file)
```

getCorPropChange	<i>Get property changes associated with correlated substitutions</i>
------------------	--

Description

Get property changes associated with correlated substitutions

Usage

```
getCorPropChange(corSubFile, propertyDF = "Cruciani",
  propertyIndex = 1)
```

Arguments

corSubFile	A dataframe with two columns i.e. Pos1 and Pos2. This file is generated by getTopSub
propertyDF	Choose one of Cruciani, Fasgai, Kidera or AAindex based amino acid properties
propertyIndex	The amino acid property to analyse. It is row number in propertyDF data frame

Value

A data frame of 12 columns consisting data regarding Wild type and mutant residues and their positions along with their properties and net changes between Pos1 and Pos2

Examples

```
getCorPropChange(corSubFile = corSubFile, propertyDF = "Cruciani", propertyIndex = 1)
```

getCorSites	<i>Get correlated sites with substitutions</i>
-------------	--

Description

Get correlated sites with substitutions

Usage

```
getCorSites(fileLoc, corMethod = "mcbasc")
```

Arguments

fileLoc	exact location of multiple sequence alignment file in "FASTA" format
corMethod	One of the methods to compute correlated sites viz. 'mip', 'elsc', 'mcbasc' and 'omes'. Default is 'mcbasc'.

Value

A subset alignment matrix of original multiple sequence alignment with significant correlations.

Examples

```
file = system.file("extdata", "linB_toy_ali.fasta", package = "aaSEA")
getCorSites(fileLoc = file, corMethod="mcbasc")
```

getPropChange	<i>Get wild type and substituted amino acid properties and associated property changes</i>
---------------	--

Description

Get wild type and substituted amino acid properties and associated property changes

Usage

```
getPropChange(subFile, propertyDF = "Cruciani", propertyIndex = 1)
```

Arguments

subFile	A data frame of single or multiple substitutions obtained using 'getAASub' function
propertyDF	Choose one of Cruciani, Fasgai, Kidera or AAindex based amino acid properties
propertyIndex	The amino acid property to analyse. It is row number in propertyDF data frame

Value

A substitution data frame with three additional columns i.e. wt.Prop, mu.Prop and Delta.Prop

Examples

```
ssFileLoc <- system.file("extdata", "singleSub.rda", package = "aaSEA")
singleSubFile <- readRDS(ssFileLoc)
msFileLoc <- system.file("extdata", "multiSub.rda", package = "aaSEA")
multiSubFile <- readRDS(msFileLoc)
getPropChange(subFile = singleSubFile, propertyDF = "Cruciani", propertyIndex = 1)
getPropChange(subFile = multiSubFile, propertyDF = "Cruciani", propertyIndex = 1)
```

getPropCorr	<i>Get amino acid property wise correlations of co-evolving columns of a multiple sequence alignment</i>
-------------	--

Description

Get amino acid property wise correlations of co-evolving columns of a multiple sequence alignment

Usage

```
getPropCorr(selMat, propertyDF = "Cruciani", propertyIndex = 1)
```

Arguments

selMat	A subset matrix of original multiple sequence alignment with significant correlations identified with 'getCorSites' function
propertyDF	One of the amino acid property data frames. viz. Cruciani, Fasgai, Kidera, AAindex. Default is Cruciani properties
propertyIndex	Specific property row number from the data frame of propertyDF

Value

A data frame of four columns viz. Pos1, Pos2, Cor and p Value. Results are filtered to find position pairs with correlations above 0.8 and below -0.8

Examples

```
selMatLoc <- system.file("extdata", "selMat.rda", package = "aaSEA")
selMat <- readRDS(selMatLoc)
getPropCorr(selMat = selMat, propertyDF = "Cruciani", propertyIndex = 1)
```

getTopSub	<i>Get sites with more than one-correlated substitutions other than conserved amino acids at that position</i>
-----------	--

Description

Get sites with more than one-correlated substitutions other than conserved amino acids at that position

Usage

```
getTopSub(selMat)
```

Arguments

`selMat` A subset matrix of original multiple sequence alignment with significant correlations identified with 'getCorSites' function

Value

A data frame with two columns i.e. Pos1 and Pos2 which is a filtered subset of many correlated substitutions based on frequency of substitution after consensus of a column in multiple sequence alignment

Examples

```
selMatLoc <- system.file("extdata", "selMat.rda", package = "aaSEA")
selMat <- readRDS(selMatLoc)
getTopSub(selMat = selMat)
```

Kidera

A data frame of 10 Kidera factors for 20 amino acids

Description

A data frame of 10 Kidera factors for 20 amino acids.

Usage

Kidera

Format

A data frame with 20 rows and ten columns representing 20 amino acids and 10 kidera factors:

Names of amino acids

KF1 Helix/bend preference

KF2 Side-chain size

KF3 Extended structure preference

KF4 Hydrophobicity

KF5 Double-bend preference

KF6 Partial specific volume

KF7 Flat extended preference

KF8 Occurrence in alpha region

KF9 pK-C

KF10 Surrounding hydrophobicity ...

matEncode	<i>Function to encode Correlated columns of alignment matrix with desired properties</i>
-----------	--

Description

Function to encode Correlated columns of alignment matrix with desired properties

Usage

```
matEncode(aliMat, pIndex, propDf)
```

Arguments

aliMat	amino acid multiple sequence alignment in the form of a matrix
pIndex	Amino acid property index to be encoded. It is the row number in the property data frame
propDf	The amino acid property to analyse. It is row number in propertyDF data frame

Value

A matrix of input dimensions with amino acid alphabets replaced by amino acid properties of choice

Examples

```
aliMatLoc <- system.file("extdata", "aliMat.rda", package = "aaSEA")
aliMat <- readRDS(aliMatLoc)
matEncode(aliMat = aliMat, pIndex = 1, propDf = "Cruciani" )
```

plotCorNet	<i>simple interactive network diagram of correlated substitution pairs</i>
------------	--

Description

simple interactive network diagram of correlated substitution pairs

Usage

```
plotCorNet(corSubDF)
```

Arguments

corSubDF	A data frame with two columns i.e. Pos1 and Pos2. This file is generated by 'getTopSub' function.
----------	---

Value

An interactive network diagram of correlated substitution pairs.

Examples

```
corSub <- corSubFile
plotCorNet(corSubDF = corSub)
```

plotCorSubChanges *Plot co-evolving sites with selected property correlations*

Description

Plot co-evolving sites with selected property correlations

Usage

```
plotCorSubChanges(corSitePropChange)
```

Arguments

corSitePropChange
A data frame of coevolving sites and associated property changes obtained by "getPropCorr" function with selected MSA matrix (selMat) and desired property selected from property data frame and property index.

Value

Returns an interactive heat map of significant sites with selected property correlations

Examples

```
fileLocation <- system.file("extdata", "corSitePropChangeDF.rda", package = "aaSEA")
corSitePropChange <- readRDS(fileLocation)
plotCorSubChanges(corSitePropChange = corSitePropChange)
```

plotMultiSubChange *Plots heat map of multiple substitution associated changes per site*

Description

Plots heat map of multiple substitution associated changes per site

Usage

```
plotMultiSubChange(multiSubChangeDF)
```

Arguments

multiSubChangeDF

A data frame of multiple amino acid substitutions per site calculated with 'getAA-Sub' and associated property changes obtained by calling 'getPropChange' function

Value

An interactive heat map of multiple substitution associated changes per site

Examples

```
multiSubChangeLoc <- system.file("extdata", "multiSubChange.rda", package = "aaSEA")
multiSubChange <- readRDS(multiSubChangeLoc)
plotMultiSubChange(multiSubChangeDF = multiSubChange )
```

plotSingleSubChange *Plots single substitution change histogram*

Description

Plots single substitution change histogram

Usage

```
plotSingleSubChange(singleSubChangeDF)
```

Arguments

singleSubChangeDF

A data frame of single amino acid substitutions per site calculated with 'getAA-Sub' and associated property changes obtained by calling 'getPropChange' function

Value

An interactive histogram representing amino acid substitution associated change

Examples

```
singleSubChangeLoc <- system.file("extdata", "singleSubChange.rda", package = "aaSEA")
singleSubChange <- readRDS(singleSubChangeLoc)
plotSingleSubChange(singleSubChangeDF = singleSubChange)
```

startSeaShiny

Title Enables to start and run the app

Description

Title Enables to start and run the app

Usage

```
startSeaShiny()
```

Value

a shiny app will be launched in browser

Examples

```
if(interactive()){
  startSeaShiny()
}
```

Index

*Topic **datasets**

- AAindex, [2](#)
- corSubFile, [3](#)
- Cruciani, [3](#)
- Fasgai, [4](#)
- Kidera, [8](#)

AAindex, [2](#)

corSubFile, [3](#)
Cruciani, [3](#)

Fasgai, [4](#)

getAASub, [4](#)
getCorPropChange, [5](#)
getCorSites, [5](#)
getPropChange, [6](#)
getPropCorr, [7](#)
getTopSub, [7](#)

Kidera, [8](#)

matEncode, [9](#)

plotCorNet, [9](#)
plotCorSubChanges, [10](#)
plotMultiSubChange, [11](#)
plotSingleSubChange, [11](#)

startSeaShiny, [12](#)