Package 'kscons'

March 31, 2017

Type Package

Title A Bayesian Approach for Protein Residue Contact Prediction using the Knob-Socket Model of Protein Tertiary Structure

Version 0.7.0

Author Qiwei Li, David B. Dahl, Marina Vannucci, Hyun Joo, Jerry W. Tsai

Maintainer Qiwei Li <liqiwei2000@gmail.com>

Description Predicts a protein's residue contact map, based on the estimation of the corresponding knob-socket list. For more details, please refer to our paper: Q. Li, D. B. Dahl, M. Vannucci, H. Joo, J. W. Tsai (2016), KScons: A Bayesian Approach for Protein Residue Contact Prediction using the Knob-socket Model of Protein Tertiary Structure, Bioinformatics, 32(24): 3774-3781 <doi:10.1093/bioinformatics/btw553>.

License GPL (>= 2)

Encoding UTF-8

LazyData true

NeedsCompilation no

Depends R (>= 2.10)

Repository CRAN

Date/Publication 2017-03-31 06:12:53 UTC

R topics documented:

allmodel_local_false *allmodel_local_false*

Description

Training data

Usage

```
allmodel_local_false
```

allmodel_local_true_free *allmodel_local_true_free*

Description

Training data

Usage

```
allmodel_local_true_free
```

allmodel_local_true_knob *allmodel_local_true_knob*

Description

Training data

Usage

```
allmodel_local_true_knob
```

allmodel_nonlocal_false *allmodel_nonlocal_false*

Description

Training data

Usage

```
allmodel_nonlocal_false
```

allmodel_nonlocal_true_free *allmodel_nonlocal_true_free*

Description

Training data

Usage

```
allmodel_nonlocal_true_free
```

allmodel_nonlocal_true_knob *allmodel_nonlocal_true_knob*

Description

Training data

Usage

```
allmodel_nonlocal_true_knob
```

estimate *estimate*

Description

Predict the knob-socket model and the corresponding residual contact map based on the amino acid sequence, denoted by a, and the secondary structure, denoted by rho. The users also need to give the iteration and burn-in numbers for the MCMC chain. The users could also input the multiple structure alignment information, by incorporating it into a matrix, where 1) the number of rows and columns of the SYMMETRIC and NON-NEGATIVE matrix should be the length of amino acid sequence; 2) each element give the number of times that each pair of residuals is contacted in the multiple structure alignment; 3) elements are separated by white space. The output is a list, where 1) a is the target primary structure; 2) rho is the target secondary structure; 3) socket is the a 3-by-K matrix, where each column is the position of the triplets; 4) gamma_map is a K-dimension binary vector, it is the MAP estimates of gamma; 5) mpv is a K-diemension vector, it is the estimated probability vector of gamma; 6) $contact_map is a L-by-L binary matrix, it is the contact map corresponding to $gamma_map; 7) $mpm is a L-by-L matrix, it is the estimated probability matrix of contact map.

Usage

```
estimate(a, rho, iter, burn, msa)
```

Arguments

a	amino acid sequence
rho	secondary structure
iter	number of MCMC iterations
burn	number of burn-in
msa	multiple structure alingment matrix

Examples

```
a <- "GRIAFTDADDVAILTYVKENARSPSSVTGNALWKAMEKSSLTQHSWQSLKDRYLKHLRG";
rho <- "CCCCCCHHHHHHHHHHHHHCCCTTTTTTTTHHHHHHHHCCTTTCCCHHHHHHHHHHCCC";
iter <- 1000;
burn <- 500;
data(msa_example);
msa <- as.matrix(msa_example);
y <- estimate(a, rho, iter, burn);
```

---

msa_example *msa_example*

Description

The multiple structure alignment information, by incorporating it into a matrix, where 1) the number of rows and columns of the SYMMETRIC and NON-NEGATIVE matrix should be the length of amino acid sequence; 2) each element give the number of times that each pair of residuals is contacted in the multiple structure alignment; 3) elements are separated by white space.

Usage

```
msa_example
```

---

validation_local *validation_local*

Description

Training data

Usage

```
validation_local
```

validation_nonlocal *validation_nonlocal*

Description

Training data

Usage

```
validation_nonlocal
```

Index