

PhagePred: a package for predicting the lifestyles of phages using alignment-free measures d_2^S and d_2^*

Version: 1.1

Authors: Kai Song

Maintainer: Kai Song ksong@qdu.edu.cn

Description

The package provides functions to count the number of occurrences of k-tuples for phage contigs and calculate the distance of contigs to temperate and lytic phages using alignment-free measures d_2^S and d_2^* . Several orders of Markov models can be used for estimating the expectation of the number of occurrence of k-tuples.

Thus, the package provides functions for

- (1) Counting the number of occurrences of k-tuples for each contig;
- (2) Compute d_2^S and d_2^* based on the contigs.

Dependencies

The python (≥ 2.7) and R (≥ 3.5) are needed for the PhagePred execution.

Installation

To quick start, first download the package file PhagePred.tar.bz2 according to your operating system.

For Linux user, you can install the package from the command line. Simply type the following to the command line,

```
tar jxvf PhagePred.tar.bz2
```

Usage

(1) Calculation of the distance of phage contigs to the temperate and lytic phages

using d_2^S and d_2^* .

First, one can download the k-tuples for temperate and lytic phages from

“/trained_model/”, then, put the four files of trained Markov models in the directory:

```
<path_to_the_PhagePred>/PhagePred/trained_model/
```

As an example, the package provides one testing data set containing a phage contigs in the directory of “/test_data/”. Then, the test data should be put in the main directory of the software. The usage is:

```
Python PhagePred_1.1.py test_sample.fa
```

There is only one input in the command, it is the path of the file of phage contig.

The above package only supports the input data sets with format of “fa”.

The result will be something like the following.

```
Markov order = 0, k = 6, The distance to Temperate Phages: d2* = 0.335663163518678, d2S = 0.434388875467936
Markov order = 0, k = 6, The distance to Lytic Phages: d2* = 0.346787246796094, d2S = 0.43760237527879
Markov order = 1, k = 6, The distance to Temperate Phages: d2* = 0.38409701488084, d2S = 0.447489347345293
Markov order = 1, k = 6, The distance to Lytic Phages: d2* = 0.383869670861765, d2S = 0.462234630928492
Markov order = 2, k = 6, The distance to Temperate Phages: d2* = 0.409922486265928, d2S = 0.460538231821917
Markov order = 2, k = 6, The distance to Lytic Phages: d2* = 0.414011216843687, d2S = 0.46719595047331
Markov order = 0, k = 7, The distance to Temperate Phages: d2* = 0.393621735369509, d2S = 0.453840677910701
Markov order = 0, k = 7, The distance to Lytic Phages: d2* = 0.399729747280432, d2S = 0.462358574418976
Markov order = 1, k = 7, The distance to Temperate Phages: d2* = 0.424690999669482, d2S = 0.469156797720656
Markov order = 1, k = 7, The distance to Lytic Phages: d2* = 0.423665204018548, d2S = 0.475232032970493
Markov order = 2, k = 7, The distance to Temperate Phages: d2* = 0.44181897732879, d2S = 0.477270566350555
Markov order = 2, k = 7, The distance to Lytic Phages: d2* = 0.443633818863224, d2S = 0.481034004608862
Markov order = 0, k = 8, The distance to Temperate Phages: d2* = 0.434885778955234, d2S = 0.471964415062921
Markov order = 0, k = 8, The distance to Lytic Phages: d2* = 0.43990535984144, d2S = 0.477956654939124
Markov order = 1, k = 8, The distance to Temperate Phages: d2* = 0.452596508380064, d2S = 0.479930945532997
Markov order = 1, k = 8, The distance to Lytic Phages: d2* = 0.454032666958743, d2S = 0.483365054063144
Markov order = 2, k = 8, The distance to Temperate Phages: d2* = 0.462160790339302, d2S = 0.485309346332405
Markov order = 2, k = 8, The distance to Lytic Phages: d2* = 0.465385264985957, d2S = 0.488031529109127
Markov order = 0, k = 9, The distance to Temperate Phages: d2* = 0.461021257530101, d2S = 0.480960011348131
Markov order = 0, k = 9, The distance to Lytic Phages: d2* = 0.465890827515619, d2S = 0.486015710439769
Markov order = 1, k = 9, The distance to Temperate Phages: d2* = 0.469893542588303, d2S = 0.486534145369869
Markov order = 1, k = 9, The distance to Lytic Phages: d2* = 0.473560810269777, d2S = 0.489344263079449
Markov order = 2, k = 9, The distance to Temperate Phages: d2* = 0.474881605026805, d2S = 0.488558455941141
Markov order = 2, k = 9, The distance to Lytic Phages: d2* = 0.479761268116475, d2S = 0.492049439479449
```

The first column is the Markov order used for background sequences, the second column is the length of k-tuple, the third and fourth columns are d2* and d2S values.