

PhAL manual

PhAL (Phage Display Alignment)

1. Introduction and input files

PhAL (Phage Display Alignment software is a novel bioinformatics tool which was developed for alignment of short sequences from phage display on the reference sequence of the protein which is intended as an target for antibody which is being investigated. Software uses algorithm of local alignment and provides template-based analysis of the phage display for antibodies. A text file containing short nucleotide sequences serves as input for the software. The input file format is txt and the sequences are written as one sequence per line. The nucleotide sequences are subsequently converted to peptide sequences and subsequently analyzed using local alignment to a reference protein sequence.

2. Prerequisites

PhAL was created using Perl and is primarily designed for GNU/Linux but can run on other operating systems such as MS Windows or Mac OSX fulfilling requirements for launching. The software prerequisites are Perl v5.18.2 or higher. The program further requires the following Perl modules, which are available on CPAN (Perl Archiving Perl Network, <https://www.cpan.org>):

```
use strict;
use Term::ANSIColor;      # Module for colors and errors
use Getopt::Long qw(:config no_ignore_case bundling);
use Statistics::Descriptive::Smoother;
```

3. Obtaining and installation

The latest source release is available on the web pages of MOU (<https://www.mou.cz/>). To install the software from the source package unpack the directory by the command:
unzip PhAL_v1.1.zip

4. Using PhAL

Usage: Alignment_analysis_Orders.sh [options]

The software can be launched directly from installation directory by the command:
./Alignment_analysis_Orders.sh --in input_file --out output_file.txt --reference reference_file.fa

All parameters are required and the software uses only the long switches (--) instead of the short ones (-). Options for software usage are defined as follows:

```
--reference <string> Path to input file which containing the sequence of a protein that is the
target of an antibody in fasta format.
--out <string> Path to output file which will be provided in txt format.
--in <string> Path to input file with nucleotide sequences from phage display (one per line).
```

5. PhAL output files

The PhAL produces one output file in the directory in which it was launched. The file contains a table with a tab as a column separator. The file is in txt format. The columns have the following meaning:

column 1 - amino acid order in the protein sequence

column 2 - amino acid in the protein sequence

column 3 - number of sequences aligned to the given amino acid, where an amino acid match is found

The amino acids with the highest score in column 3 are the likely site of antibody binding to the protein.