

A Guide to the stand alone version of MLTreeMap:

MLTreeMap is a software framework, designed for phylogenetic and functional analysis of metagenomics data. It searches for instances of marker genes on nucleotide sequences and deduces their most likely origin in a set of reference phylogenies. The current version of MLTreeMap can be downloaded and installed individually. MLTreeMap runs on Mac and Linux systems.

A) Installation of MLTreeMap

Step 1

Unzip and unpack the file `MLTreeMap_packages.tar.gz`.

Step 2

Enter the directory `MLTreeMap_packages/install/`.

Type `make`.

The `make` program now creates the data structure of MLTreeMap (to be found in `MLTreeMap_packages/MLTreeMap_ansi_C/`), compiles MLTreeMap itself and most needed sub-programs (`hmmalign`, `Genewise` and `RAXML`).

Step 3

BLAST and Gblocks have to be added manually. For this enter the directory `MLTreeMap_packages/install/sources/BLAST/`.

Here you will find a collection of binaries for BLAST. Choose the one appropriate for your system and copy it to the directory `MLTreeMap_packages/MLTreeMap_ansi_C/data/sub_binaries/`.

Repeat this for Gblocks (to be found in `MLTreeMap_packages/install/sources/Gblocks/`).

Now MLTreeMap is ready to use. You can copy the directory `MLTreeMap_ansi_C` to any place you like (as well as renaming it), there are no further dependencies.

B) Usage of MLTreeMap.

MLTreeMap has to be accessed on the command line.

An example for a valid input command is: `./MLTreeMap -i example_input/rubisco.txt -q test`

This will analyze the sequence in the file `rubisco.txt` and write the result to the standard output directory `MLTreeMap_ansi_C/outputs/`.

The `-q` option defines the denominator, which you can choose individually. It will be added to the beginning of each output file.

Note that the `-i` and the `-q` option are both mandatory.

Further (optional) input options are:

- o use this option if you want to write the output to another directory than the default.
- b activates non-parametric Bootstrapping (max = 99). Note, usually bootstrapping is not necessary.
- f v (Maximum likelihood analysis, default) or p (maximum parsimony analysis).
- t 1 (use the GEBA reference phylogeny). If you want to use the default reference phylogeny of MLTreeMap, omit this option.