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_package_2_03.tar.gz`.

Step 2

Enter the directory `MLTreeMap_package_2_03/install/`.

Type `make`.

The make program now creates the data structure of MLTreeMap (to be found in `MLTreeMap_package_2_03/mltreemap_2_03/`), and compiles most needed sub-programs (hmmalign, Genewise and RAxML).

Step 3

BLAST and Gblocks have to be added manually. For this enter the directory `MLTreeMap_package_2_03/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_package_2_03/Mltreemap_2_03/data/sub_binaries/`.

Repeat this for Gblocks (to be found in
`MLTreeMap_package_2_03/install/sources/Gblocks/`).

Now MLTreeMap is ready to use. You can copy the directory `Mltreemap_2_03` to any place you like (as well as renaming it), the only dependency is that you have Perl installed on your system.

B) Usage of MLTreeMap.

MLTreeMap has to be accessed on the command line.

An example for a valid input command is: `./mltreemap.pl —i example_input/rubisco.txt`

This will analyze the sequence in the file rubisco.txt and write the result to the standard output directory `mltreemap_2_03/output/`.

Further (optional) input options are:
-b   activates non-parametric Bootstrapping. Note: usually bootstrapping is not necessary.
-c   usage a computer cluster (0 = no cluster (default), s = sun grid). Note: This has only been tested on our cluster, thus further code-adaptations might be necessary to get it running on another cluster.
-f   v (Maximum likelihood analysis, default) or p (maximum parsimony analyis).
-g   minimal sequence length after Gblocks (default = 50).

-o   use this option if you want to write the output to another directory than the default.

-s   minimum bitscore of the blast hits (default = 60).

-t   phylogenetic reference tree (p = MLTreeMap reference tree (default), g = GEBA reference tree).

Note: the mandatory "-q" option of earlier MLTreeMap versions has been abandoned. Instead the handling of output directories has been improved.