How to create a custom reference database for BLAST in SEED2

https://www.mothur.org/wiki/Silva_reference_files

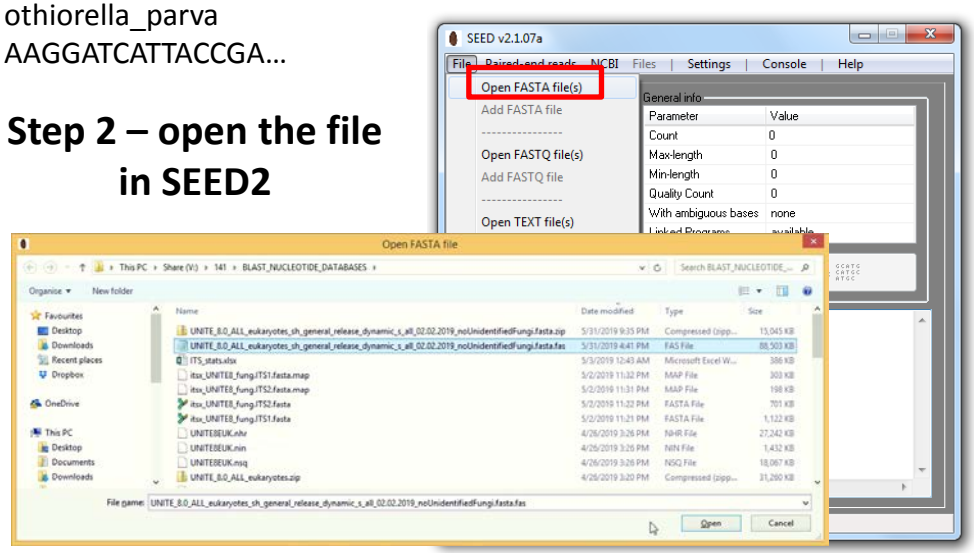
Step 1 – dowload reference sequences as FASTA or create FASTA file from your edited sequences and put desired information into the titles of the sequences, e.g.:

<Species name>|<ACCESSION#>|<full taxonomy>

Here is example of FASTA:

>Thelephora_sp|UDB014120|k__Fungi;p__Basidiomycota;c__Agaricomycetes;o__Thelephorales;f__Thelephoraceae;g__Thelephora;s__Thelephora_sp
GGAAGGATCATTACT…
>Dothiorella_parva|KC898234|k__Fungi;p__Ascomycota;c__Dothideomycetes;o__Botryosphaeriales;f__Botryosphaeriaceae;g__Dothiorella;s__Dothiorella_parva
AAGGATCATTACCGA…

Step 2 – open the file in SEED2



https://unite.ut.ee/repository.php

QIIME release (download)

Three sets of QIIME files are released, corresponding to the SHs resulting from clustering at the 97% and 99% threshold levels. The third set of files is the result of a dynamic use of clustering thresholds, such that some SHs are delimited at the 97% level, some at the 97.5% level, some at the 98% level, and so on; these choices were made manually by experts of those particular lineages of fungi. The syntax is the same throughout the three sets of files.

Each SH is given a stable name of the accession number type, here shown in the FASTA file of the dynamic set:

```
>SH099456.05FU_FJ357315_refs
CACAATATGAAGGCGGGCTGGCACTCCTTGAGAGGACCGGC…

SH099456 = accession number of the SH
05FU = global key release 5, organism group FUngi
FJ357315 = GenBank/UNITE accession number of sequence chosen to
refs = this is a manually designated RefS
(reps = this is an automatically chosen RepS)
```

Step 3 – save FASTA as local database

This step will generate 3 binary files in database "folder":

MYREFERENCE.nhr	5.2.2020 11:36	Soubor NHR	26 329 kB
MYREFERENCE.nin	5.2.2020 11:36	Soubor NIN	1 381 kB
MYREFERENCE.nsq	5.2.2020 11:36	Soubor NSQ	17 509 kB

Step 4 – use the reference database for BLAST identification