# Taxonomy

* Made file of representative OTU sequences at 97% sequence similarity for UNITE taxonomic analysis
  + dereplicated\_otu97.csv in ‘data\_output/’
  + made a fasta file with the otu name and sequence called dereplicated\_otu97.fasta
* Downloaded the UNITE USEARCH database version 8
  + <https://plutof.ut.ee/#/doi/10.15156/BIO/786345>
  + **When using this resource, please cite it as follows:**  
    UNITE Community (2019): UNITE USEARCH/UTAX release for Fungi. Version 18.11.2018. UNITE Community. [**https://doi.org/10.15156/BIO/786345**](https://doi.org/10.15156/BIO/786345)
* usearch -sintax data\_output/dereplicated\_otu97.fasta -db ~/../../Volumes/Cenococcum/UNITE/utax\_reference\_dataset\_02.02.2019.fasta -tabbedout unite/reads.sintax -strand both -sintax\_cutoff 0.8

00:00 1.5Mb 0.1% Reading /Users/eabowman/../../Volumes/Cenococcum/UNITE/utax\_reference\_dataset\_0200:01 52Mb 19.5% Reading /Users/eabowman/../../Volumes/Cenococcum/UNITE/utax\_reference\_dataset\_0200:01 119Mb 100.0% Reading /Users/eabowman/../../Volumes/Cenococcum/UNITE/utax\_reference\_dataset\_02.02.2019.fasta

00:01 86Mb 0.1% Masking (fastnucleo)

00:03 86Mb 100.0% Masking (fastnucleo)

00:05 87Mb 100.0% Word stats

00:05 87Mb 100.0% Alloc rows

00:08 264Mb 100.0% Build index

00:10 275Mb 100.0% Initialize taxonomy data

00:10 277Mb 100.0% Building name table

00:10 277Mb 39004 names, tax levels min 1, avg 5.8, max 7

WARNING: 3 taxonomy nodes have >1 parent

00:13 279Mb 100.0% Processing

* Created Funguild.R script to add taxonomic data to site by species matrix
  + Saved as OTU97\_withTax.csv
* Ran above file on Funguild online site
  + Output file FUNguild/OTU97\_withTax.guilds.txt