New analysis: Mothur version 1.44.0

Mothur pipeline for processing ITS sequences accompanying the manuscript:

Soil microbial communities and biogeochemistry during human decomposition differs between seasons: evidence from year-long trials

Authors: Lois S. Taylor, Allison R. Mason, Hannah L. Noel, Michael E. Essington, Mary C. Davis, Veronica A. Brown, Dawnie W. Steadman, Jennifer M. DeBruyn

This is the full pipeline for the Spring dataset, and includes the filenames accordingly. Winter data was processed via the same pipeline, and final file names are found at the end of this document.

**mothur > make.file(inputdir=NIJARFSP.fasta, type=fastq, prefix=NIJARFSP)**

Output File Names:

NIJARFSP.fasta\NIJARFSP.files

# Now we merge paired reads with the following command:

**mothur> make.contigs(file=NIJARFSP.files, processors=8)**

Output File Names:

NIJARFSP.fasta\NIJARFSP.trim.contigs.fasta

NIJARFSP.fasta\NIJARFSP.scrap.contigs.fasta

NIJARFSP.fasta\NIJARFSP.contigs.report

NIJARFSP.fasta\NIJARFSP.contigs.groupssummary.seqs(fasta=TOX1.trim.contigs.fasta)

# A file containing primer sequences was inserted into the NIJARFSP.fasta file. In this case the oligo file is named Creggerprimers2.txt and the contents are as follows:

forward CATCGATGAAGAACGCAG

forward CAACGATGAAGAACGCAG

forward CACCGATGAAGAACGCAG

forward CATCGATGAAGAACGTAG

forward CATCGATGAAGAACGTGG

forward CATCGATGAAGAACGCTG

reverse TCCTSCGCTTATTGATATGC

reverse TCCTCGCCTTATTGATATGC

# The following command was run in order to trim the various primers off :

**mothur> pcr.seqs(fasta= NIJARFSP.fasta\NIJARFSP.trim.contigs.fasta, group= NIJARFSP.fasta\NIJARFSP.contigs.groups, oligos= NIJARFSP.fasta\Creggerprimers2.txt, pdiffs=2, rdiffs=2)**

Output File Names:

NIJARFSP.fasta\NIJARFSP.contigs.pick.groups

Output File Names:

NIJARFSP.fasta\NIJARFSP.trim.contigs.pcr.fasta

NIJARFSP.fasta\NIJARFSP.trim.contigs.bad.accnos

NIJARFSP.fasta\NIJARFSP.trim.contigs.scrap.pcr.fasta

NIJARFSP.fasta\NIJARFSP.contigs.pcr.groups

#copied NIJARFSP.trim.contigs.pcr.fasta and NIJARFSP.contigs.pcr.groups to a QC file entitled NIJARFSPqc. A copy of the mothur and vsearch executable (mothur.exe, vsearch.exe) files were also included in this same file.

# The next command would ordinarily be used to remove sequences with ANY ambiguous bases and sequences less than 50 and greater than 275 in the case of 16S processing. According to Cregger (2018), only reads below 200 bp were removed. The ITS2 region that we are amplifying here is about 300-400 bp, and fungal ITS regions can vary considerably in size. While it is possible that there may be long sequences with minimal overlap, it is worth considering not removing long sequences in order not to lose information at the top end. So while the general command may read as follows:

screen.seqs(fasta=FILE.trim.contigs.pcr.fasta, group=FILE.contigs.pcr.groups, maxambig=0, maxlength=275)….we instead run an edited version (immediately below).

# The following command was run:

**screen.seqs(fasta=NIJARFSP.trim.contigs.pcr.fasta, group=NIJARFSP.contigs.pcr.groups, maxambig=0, minlength=200)**

Output File Names:

NIJARFSP.contigs.pcr.pick.groups

NIJARFSP.trim.contigs.pcr.good.fasta

NIJARFSP.trim.contigs.pcr.bad.accnos

NIJARFSP.contigs.pcr.good.groups

# The following command was run to keep only the unique sequences:

**unique.seqs(fasta=NIJARFSP.trim.contigs.pcr.good.fasta)**

Output File Names:

NIJARFSP.trim.contigs.pcr.good.names

NIJARFSP.trim.contigs.pcr.good.unique.fasta

# The following command creates a table with rows of unique sequences and columns having the names of the groups

**count.seqs(name=NIJARFSP.trim.contigs.pcr.good.names, group=NIJARFSP.contigs.pcr.good.groups)**

Output File Names:

NIJARFSP.trim.contigs.pcr.good.count\_table

We ran the following command:

**pre.cluster(fasta=NIJARFSP.trim.contigs.pcr.good.unique.fasta,count=NIJARFSP.trim.contigs.pcr.good.count\_table,diffs=3,processors=12)**

Output File Names:

NIJARFSP.trim.contigs.pcr.good.unique.pick.fasta

Output File Names:

NIJARFSP.trim.contigs.pcr.good.unique.precluster.fasta

NIJARFSP.trim.contigs.pcr.good.unique.precluster.count\_table

We ran the following command to remove chimeras:

**mothur> chimera.vsearch(fasta=current, count=current,dereplicate=t)**

Output File Names:

NIJARFSP.trim.contigs.pcr.good.unique.precluster.denovo.vsearch.pick.count\_table

NIJARFSP.trim.contigs.pcr.good.unique.precluster.denovo.vsearch.chimeras

NIJARFSP.trim.contigs.pcr.good.unique.precluster.denovo.vsearch.accnos

We ran the following command:

**mothur > remove.seqs(fasta=current, accnos=current)**

Output File Names:

NIJARFSP.trim.contigs.pcr.good.unique.precluster.pick.fasta

\_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_

# Download UNITE database release 8.2 (2020-02-04) for fungi.

#We ran the following command to classify sequences:

**mothur > classify.seqs(fasta= NIJARFSP.trim.contigs.pcr.good.unique.precluster.pick.fasta, count= NIJARFSP.trim.contigs.pcr.good.unique.precluster.denovo.vsearch.pick.count\_table, reference=UNITEv8\_sh\_dynamic.fasta, taxonomy= UNITEv8\_sh\_dynamic.tax)**

Output File Names:

NIJARFSP.trim.contigs.pcr.good.unique.precluster.pick.UNITEv8\_sh\_dynamic.wang.taxonomy

NIJARFSP.trim.contigs.pcr.good.unique.precluster.pick.UNITEv8\_sh\_dynamic.wang.tax.summary

NIJARFSP.trim.contigs.pcr.good.unique.precluster.pick.UNITEv8\_sh\_dynamic.wang.flip.accnos

#We ran the following command to remove unwanted lineages:

**mothur > remove.lineage(fasta=current, count=current, taxonomy= NIJARFSP.trim.contigs.pcr.good.unique.precluster.pick.UNITEv8\_sh\_dynamic.wang.taxonomy,taxon=unknown-Protista-Protozoa-Plantae-Chromista-Bacteria-Animalia)**

Output File Names:

NIJARFSP.trim.contigs.pcr.good.unique.precluster.pick.UNITEv8\_sh\_dynamic.wang.pick.taxonomy

# We cluster sequences with greedy clustering, using a cutoff of 0.05 and method=agc. Agc is a VSEARCH clustering option because here contigs are of differing lengths, and thus we cannot calculate a distance matrix.

#We ran the following command:

**cluster(fasta= NIJARFSP.trim.contigs.pcr.good.unique.precluster.pick.fasta, count= NIJARFSP.trim.contigs.pcr.good.unique.precluster.denovo.vsearch.pick.count\_table, method=agc, cutoff=0.05)**

Output File Names:

NIJARFSP.trim.contigs.pcr.good.unique.precluster.pick.agc.list

# Make the OTU table.

#We ran the following command: **make.shared(list=NIJARFSP.trim.contigs.pcr.good.unique.precluster.pick.agc.list, count= NIJARFSP.trim.contigs.pcr.good.unique.precluster.denovo.vsearch.pick.count\_table)**

Output File Names:

NIJARFSP.trim.contigs.pcr.good.unique.precluster.pick.agc.shared

# And now classify the OTUs based UNITE database taxonomy

#We ran the following command:

**classify.otu(list=NIJARFSP.trim.contigs.pcr.good.unique.precluster.pick.agc.list, count= NIJARFSP.trim.contigs.pcr.good.unique.precluster.denovo.vsearch.pick.count\_table, taxonomy= NIJARFSP.trim.contigs.pcr.good.unique.precluster.pick.UNITEv8\_sh\_dynamic.wang.pick.taxonomy)**

Output File Names:

NIJARFSP.trim.contigs.pcr.good.unique.precluster.pick.agc.0.05.cons.taxonomy

NIJARFSP.trim.contigs.pcr.good.unique.precluster.pick.agc.0.05.cons.tax.summary

# Shared (from make.shared) and Taxonomy (from classify.otu) are now used for downstream processing in Phyloseq.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

For the winter data:

\*\*\*\*\*NOTE ON SAMPLE NUMBERING: THERE WAS NO SAMPLE TAKEN ON DAY 5, THIS WAS THE TEMPERATURE EQUILLIBRATION DATE. THE NUMBERING OUTPUTS FROM MOTHUR WILL BE USED FOR THE DNA SETUP, BUT STUDY DATES WILL BE LISTED IN THE METADATA CORRECTLY. ESSENTIALLY, NUMBER 5 IS DELETED , NUMBERS 21, 38, 55, 75, 94, 110, 126 PROCEED IN ORDER, 140 IS ADDED IN BETWEEN, AND THEN ALL NUMBERS PROCEED CORRECTLY FROM 158 WHICH APPEARS HERE AS 1589!

Final outputs:

Output File Names:

NIJARFWIN.trim.contigs.pcr.good.unique.precluster.pick.agc.0.05.cons.taxonomy

NIJARFWIN.trim.contigs.pcr.good.unique.precluster.pick.agc.0.05.cons.tax.summary

# Shared (from make.shared) and Taxonomy (from classify.otu) are now used for downstream processing in Phyloseq.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_