New analysis: Mothur version 1.43.0

Mothur pipeline for processing 16S 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=NIJARFSP16S.fasta, type=fastq, prefix=NIJARFSP16S)**

Output File Names:

NIJARFSP16S.fasta\NIJARFSP16S.files

# Merge paired reads:

**mothur> make.contigs(file=NIJARFSP16S.files, processors=8)**

Output File Names:

NIJARFSP16S.trim.contigs.fasta

NIJARFSP16S.scrap.contigs.fasta

NIJARFSP16S.contigs.report

NIJARFSP16S.contigs.groups

# A file containing primer sequences was inserted into the NIJARFSP16S.fasta file. The oligo file is named oligos16.txt and the contents are as follows:

forward GTGYCAGCMGCCGCGGTAA

reverse GGACTACNVGGGTWTCTAAT

# Trim primers:

**mothur> pcr.seqs(fasta=NIJARFSP16S.trim.contigs.fasta, group=NIJARFSP16S.contigs.groups, oligos=oligos16.txt, pdiffs=2, rdiffs=2)**

Output File Names:

NIJARFSP16S.contigs.pick.groups

Output File Names:

NIJARFSP16S.trim.contigs.pcr.fasta

NIJARFSP16S.trim.contigs.bad.accnos

NIJARFSP16S.trim.contigs.scrap.pcr.fasta

NIJARFSP16S.contigs.pcr.groups

# The next command removes sequences with ambiguous bases, sequences less than 50, and greater than 275.

**screen.seqs(fasta=NIJARFSP16S.trim.contigs.pcr.fasta, group=NIJARFSP16S.contigs.pcr.groups,** **maxambig=0, minlength=50, maxlength=275)**

Output File Names:

NIJARFSP16S.contigs.pcr.pick.groups

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/

Output File Names:

NIJARFSP16S.trim.contigs.pcr.good.fasta

NIJARFSP16S.trim.contigs.pcr.bad.accnos

NIJARFSP16S.contigs.pcr.good.groups

# Keep only the unique sequences:

**unique.seqs(fasta=NIJARFSP16S.trim.contigs.pcr.good.fasta)**

Output File Names:

NIJARFSP16S.trim.contigs.pcr.good.names

NIJARFSP16S.trim.contigs.pcr.good.unique.fasta

# Create a table with unique sequences (rows) and group names (columns)

count.seqs(name=NIJARFSP16S.trim.contigs.pcr.good.names, group=NIJARFSP16S.contigs.pcr.good.groups)

Output File Names:

NIJARFSP16S.trim.contigs.pcr.good.count\_table

#\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

# make reduced SILVA database:

#\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

# Download SILVA database release 132 from the MOTHUR website.

**mothur > pcr.seqs(fasta=ecoli16SrRNA.fasta, oligos=oligos16.txt)**

Output File Names:

ecoli16SrRNA.pcr.fasta

**mothur > align.seqs(fasta=ecoli16SrRNA.pcr.fasta, reference=silva.nr\_v132.align)**

Output File Names:

ecoli16SrRNA.pcr.align

ecoli16SrRNA.pcr.align.report

**mothur > pcr.seqs(fasta=silva.nr\_v132.align, start=13862, end=23444, keepdots=F, processors=8)**

Output File Names:

silva.nr\_v132.pcr.align

**mothur > rename.file(input=silva.nr\_v132.pcr.align, new=silva.nr\_v132.v4.fasta)**

Current files saved by mothur:

accnos=NIJARFSP16S.trim.contigs.pcr.bad.accnos

fasta=silva.nr\_v132.pcr.align

group=NIJARFSP16S.contigs.pcr.good.groups

name=NIJARFSP16S.trim.contigs.pcr.good.names

oligos=oligos16.txt

count=NIJARFSP16S.trim.contigs.pcr.good.count\_table

processors=8

summary=ecoli16SrRNA.pcr.summary

#\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

# Done with manipulating the SILVA database….now we align our sequences….

#\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

# align unique sequences to the reduced reference:

**mothur > align.seqs(fasta=NIJARFSP16S.trim.contigs.pcr.good.unique.fasta, reference=silva.nr\_v132.v4.fasta)**

Output File Names:

NIJARFSP16S.trim.contigs.pcr.good.unique.align

NIJARFSP16S.trim.contigs.pcr.good.unique.align.report

NIJARFSP16S.trim.contigs.pcr.good.unique.flip.accnos

**mothur > screen.seqs(fasta= NIJARFSP16S.trim.contigs.pcr.good.unique.align, count= NIJARFSP16S.trim.contigs.pcr.good.count\_table, summary= NIJARFSP16S.trim.contigs.pcr.good.unique.summary, start=8, end=9582, maxhomop=8)**

Running command: remove.seqs(accnos=NIJARFSP16S.trim.contigs.pcr.good.unique.bad.accnos.temp, count=NIJARFSP16S.trim.contigs.pcr.good.count\_table)

Removed 1467473 sequences from your count file.

Output File Names:

NIJARFSP16S.trim.contigs.pcr.good.pick.count\_table

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/

Output File Names:

NIJARFSP16S.trim.contigs.pcr.good.unique.good.summary

NIJARFSP16S.trim.contigs.pcr.good.unique.good.align

NIJARFSP16S.trim.contigs.pcr.good.unique.bad.accnos

NIJARFSP16S.trim.contigs.pcr.good.good.count\_table

# Remove the overhangs

**mothur > filter.seqs(fasta=NIJARFSP16S.trim.contigs.pcr.good.unique.good.align, vertical=T, trump=.)**

Output File Names:

NIJARFSP16S.filter

NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.fasta

# Identify the unique sequences

**mothur > unique.seqs(fasta= NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.fasta, count= NIJARFSP16S.trim.contigs.pcr.good.good.count\_table)**

Output File Names:

NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.count\_table

NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.unique.fasta

# Pre-cluster sequences, splitting by group and sorting by abundance; sequences within a difference=2 will be merged

**mothur > pre.cluster(fasta=NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.unique.fasta, count= NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.count\_table, diffs=2)**

Output File Names:

NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.unique.pick.fasta

# Remove chimeras using vsearch (dereplicate=t, defaults are dereplicate=t, dups=t)

**mothur > chimera.vsearch(fasta=NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.unique.precluster.fasta, count= NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.unique.precluster.count\_table, dereplicate=T)**

Output File Names:

NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.unique.precluster.denovo.vsearch.pick.count\_table

NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.unique.precluster.denovo.vsearch.chimeras

NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.unique.precluster.denovo.vsearch.accnos

# Now remove chimeras

**mothur > remove.seqs(fasta=NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.unique.precluster.fasta, accnos=NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.unique.precluster.denovo.vsearch.accnos)**

Output File Names:

NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.unique.precluster.pick.fasta

# Classify the sequences

**mothur > classify.seqs(fasta=NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.unique.precluster.pick.fasta, count=NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.unique.precluster.denovo.vsearch.pick.count\_table, reference=silva.nr\_v132.align, taxonomy=silva.nr\_v132.tax, cutoff=80)**

Output File Names:

NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.unique.precluster.pick.nr\_v132.wang.taxonomy

NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.unique.precluster.pick.nr\_v132.wang.tax.summary

# Remove unwanted lineages

**mothur > remove.lineage(fasta=NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.unique.precluster.pick.fasta, count=NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.unique.precluster.denovo.vsearch.pick.count\_table, taxonomy=NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.unique.precluster.pick.nr\_v132.wang.taxonomy, taxon=Chloroplast=Mitochondria-unknown-Archaea-Eukaryota)**

Output File Names:

NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.unique.precluster.pick.pick.fasta

NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.unique.precluster.denovo.vsearch.pick.pick.count\_table

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/

Output File Names:

**NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.unique.precluster.pick.nr\_v132.wang.pick.taxonomy**

NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.unique.precluster.pick.nr\_v132.wang.accnos

**NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.unique.precluster.denovo.vsearch.pick.pick.count\_table**

**NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.unique.precluster.pick.pick.fasta**

**mothur > summary.tax(taxonomy=current, count=current)**

Output File Names:

NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.unique.precluster.pick.nr\_v132.wang.pick.tax.summary

# Cluster sequences into OTUs based on 97% sequence similarity

**mothur > dist.seqs(fasta=current, cutoff=0.03)**

Output File Names:

NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.unique.precluster.pick.pick.dist

**mothur > cluster(column=NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.unique.precluster.pick.pick.dist, count=NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.unique.precluster.denovo.vsearch.pick.pick.count\_table)**

Output File Names:

NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.unique.precluster.pick.pick.opti\_mcc.list

NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.unique.precluster.pick.pick.opti\_mcc.steps

NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.unique.precluster.pick.pick.opti\_mcc.sensspec

# Make the OTU table

**mothur > make.shared(list=NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.unique.precluster.pick.pick.opti\_mcc.list, count=NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.unique.precluster.denovo.vsearch.pick.pick.count\_table)**

Output File Names:

NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.unique.precluster.pick.pick.opti\_mcc.shared

# Classify the OTUs based on SILVA database taxonomy

**mothur > classify.otu(list=NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.unique.precluster.pick.pick.opti\_mcc.list, count=NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.unique.precluster.denovo.vsearch.pick.pick.count\_table, taxonomy=NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.unique.precluster.pick.nr\_v132.wang.pick.taxonomy, label=0.03)**

Output File Names:

NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.unique.precluster.pick.pick.opti\_mcc.0.03.cons.taxonomy

NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.unique.precluster.pick.pick.opti\_mcc.0.03.cons.tax.summary

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

NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.unique.precluster.pick.pick.opti\_mcc.shared

NIJARFSP16S.trim.contigs.pcr.good.unique.good.filter.unique.precluster.pick.pick.opti\_mcc.0.03.cons.taxonomy

# For ease of use, these files will copied to (and renamed in): SP\_16S\_donor series data files and pipelines. File names are as follows:

NIJARFSP16S.shared

NIJARFSP16S.taxonomy

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

#For the WIN data:

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

NIJARFWIN16S.trim.contigs.pcr.good.unique.good.filter.unique.precluster.pick.pick.opti\_mcc.shared

NIJARFWIN16S.trim.contigs.pcr.good.unique.good.filter.unique.precluster.pick.pick.opti\_mcc.0.03.cons.taxonomy

# For ease of use, these files will copied to (and renamed in): WIN\_16S\_donor series data files and pipelines. File names are as follows:

NIJARFWIN16S.shared

NIJARFWIN16S.taxonomy