mothur tutorial
STAMPS 2017

Tracy K. Teal @tracykteal

Data Carpentry

Michigan State University, UC Davis

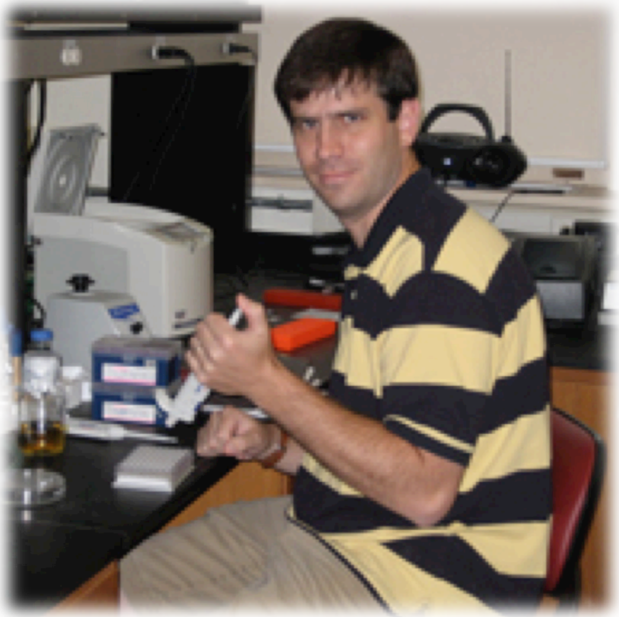
mothur

Pat Schloss

Microbiology & Immunology

University of Michigan

Sarah Westcott



http://www.mothur.org

Outline

- Case study – microbial communities in soil

- mothur workflow

- mothur MiSeq SOP

http://www.mothur.org/wiki/MiSeq_SOP

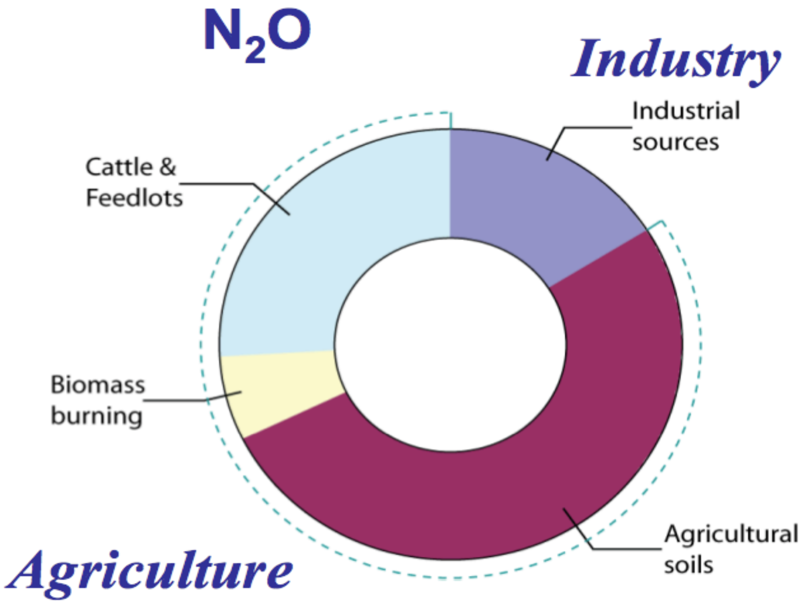
How to get your data in, get your data out (an OTU table), and get help

- Some preliminary statistics

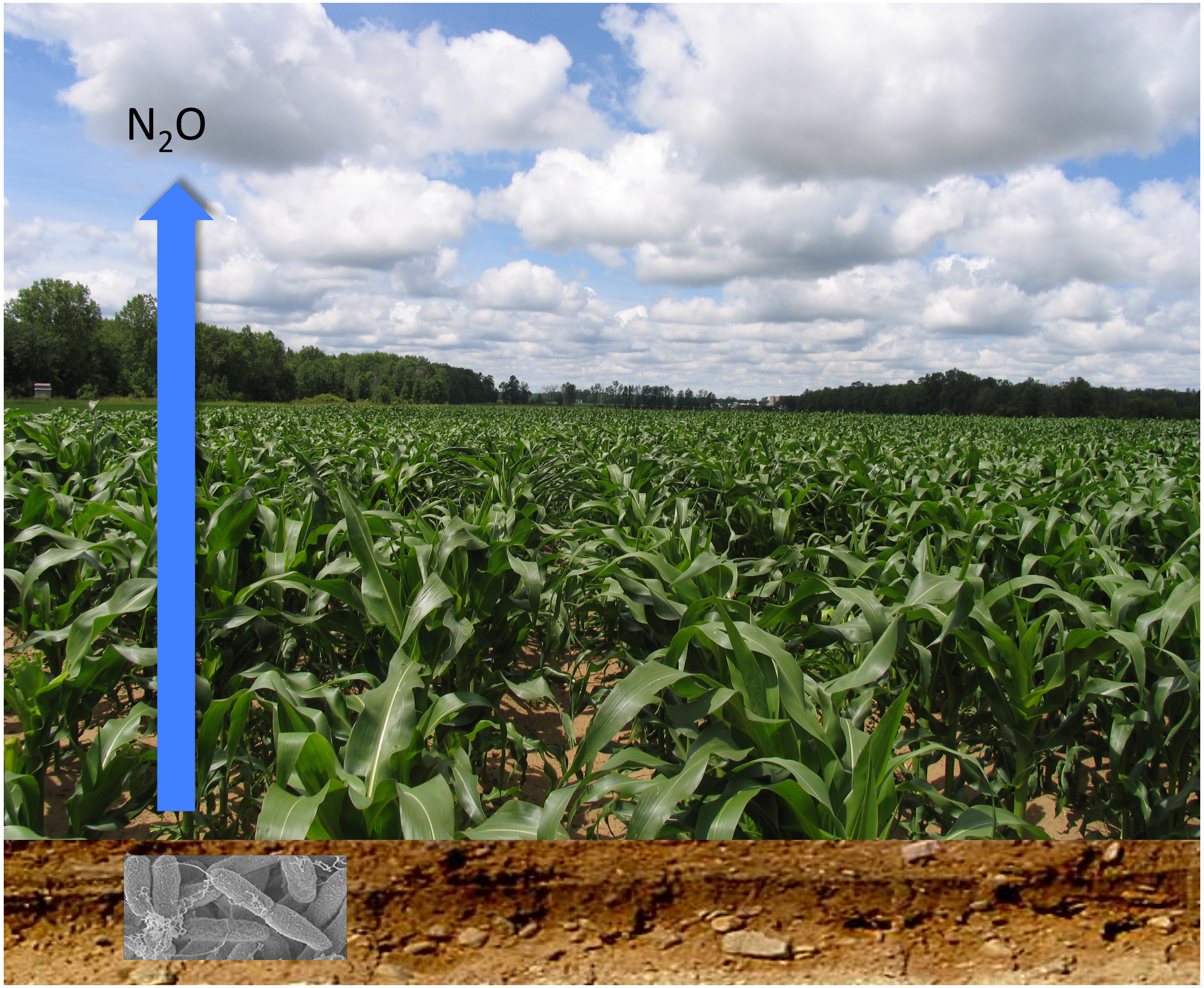
The functional potential of soil microbial communities shifts with agricultural management, increasing their capacity to produce greenhouse gases

Tracy Teal, Michigan State University
Vicente Gomez-Alvarez, Environmental Protection Agency
Tom Schmidt, University of Michigan

Land use change and intensive agriculture increase greenhouse gas fluxes



Microbes are the primary mediators of nitrous oxide



Land management, sustainability and microbial communities

- How do microbial communities change with land management?
- What is the relationship between denitrifying bacteria and N_2O in agricultural and native soils?

Experimental sites

Kellogg Biological Station LTER

How do microbial communities change with land management?

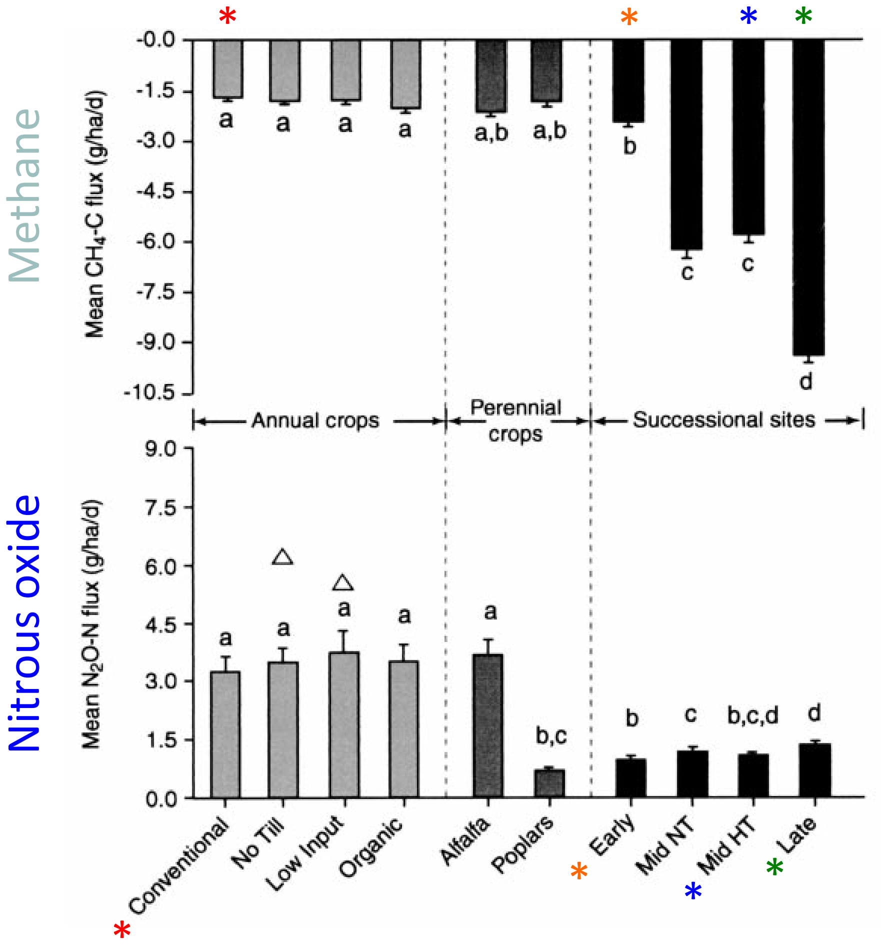
Kellogg Biological Station LTER

- Vicente Gomez-Alvarez

AG 🟥 Conventional Agriculture (3 crop rotation)
ES 🟧 Early Successional (20 years abandoned)
SF 🟦 Successional Forest (40 years abandoned)
DF 🟩 Deciduous Forest (native forest, never tilled)

How do microbial communities change with land management?

Kellogg Biological Station LTER

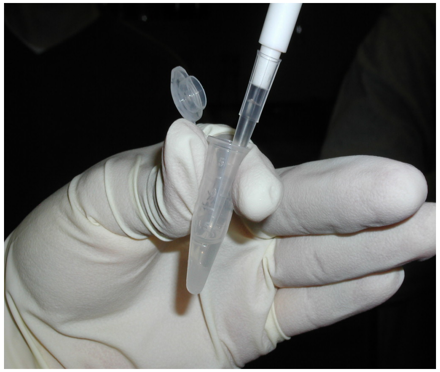


🟥	AG	Conventional Agriculture
🟧	ES	Early Successional
🟦	SF	Successional Forest
🟩	DF	Deciduous Forest

Metagenomics approach

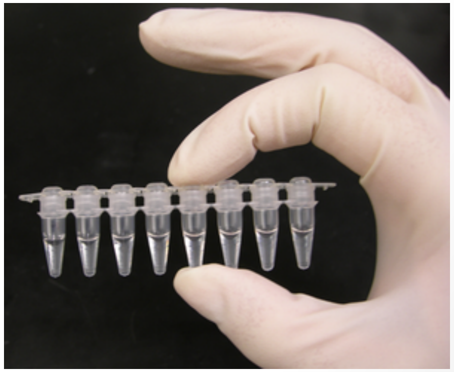
Collect, composite
and sieve

Extract
DNA

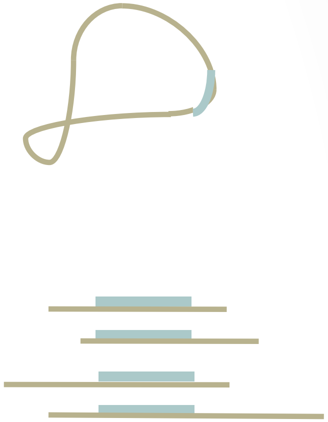


Shotgun
metagenomes

PCR

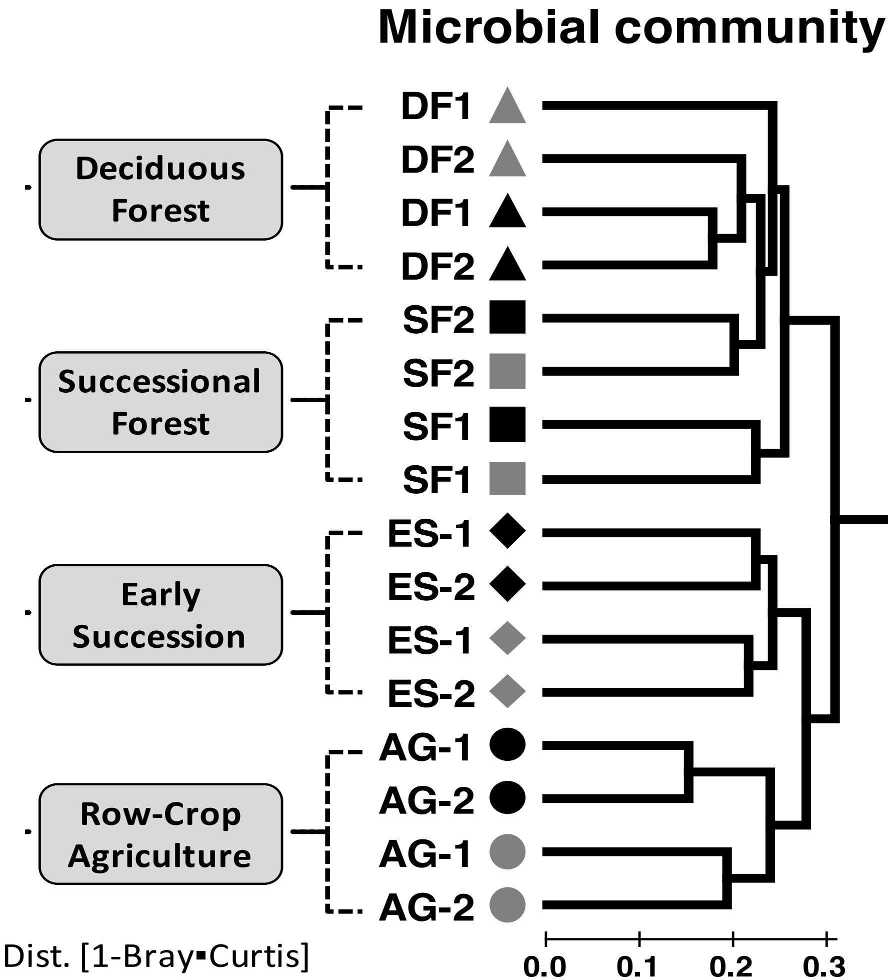


Amplicon

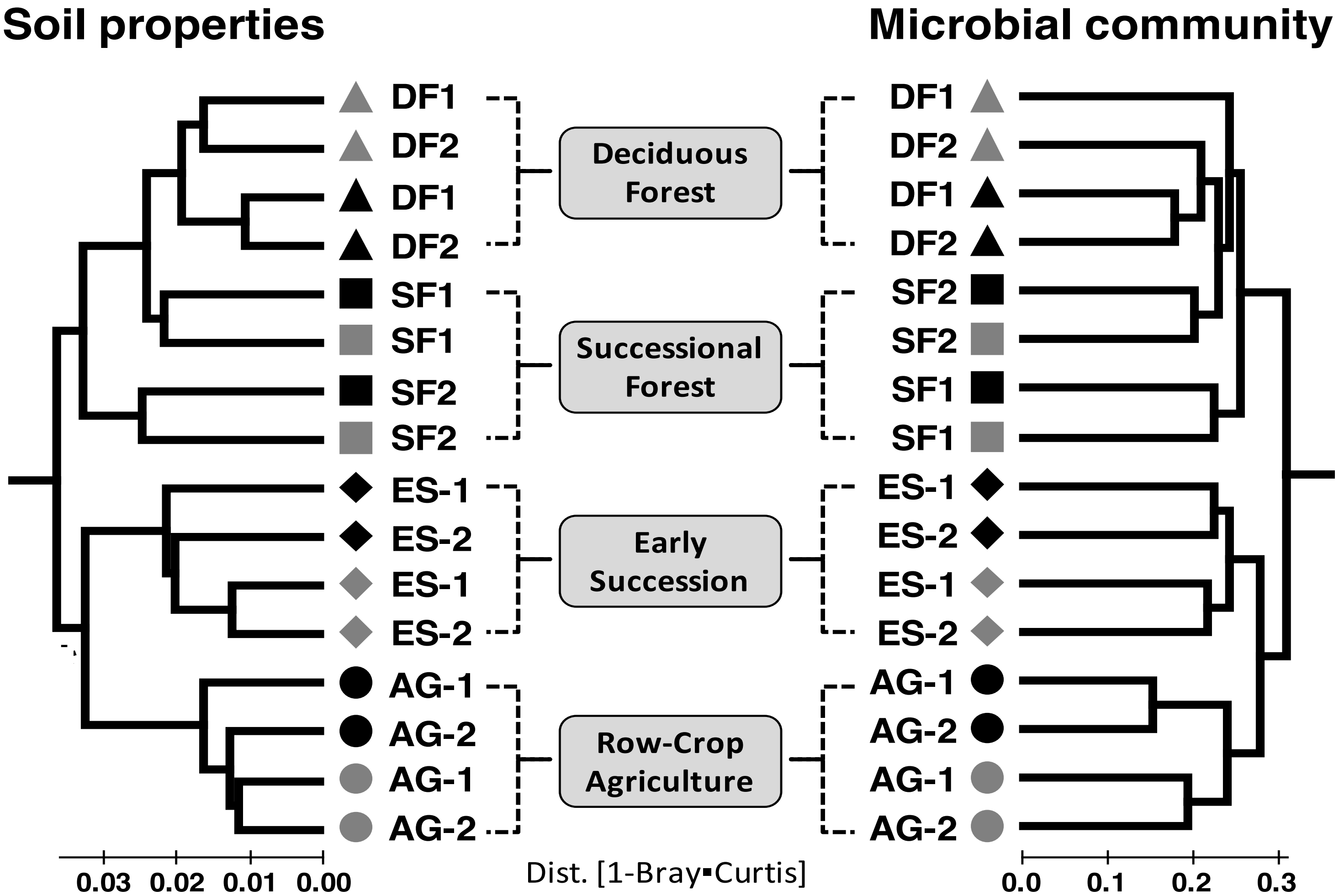


2 replicates of two treatments in each of two years

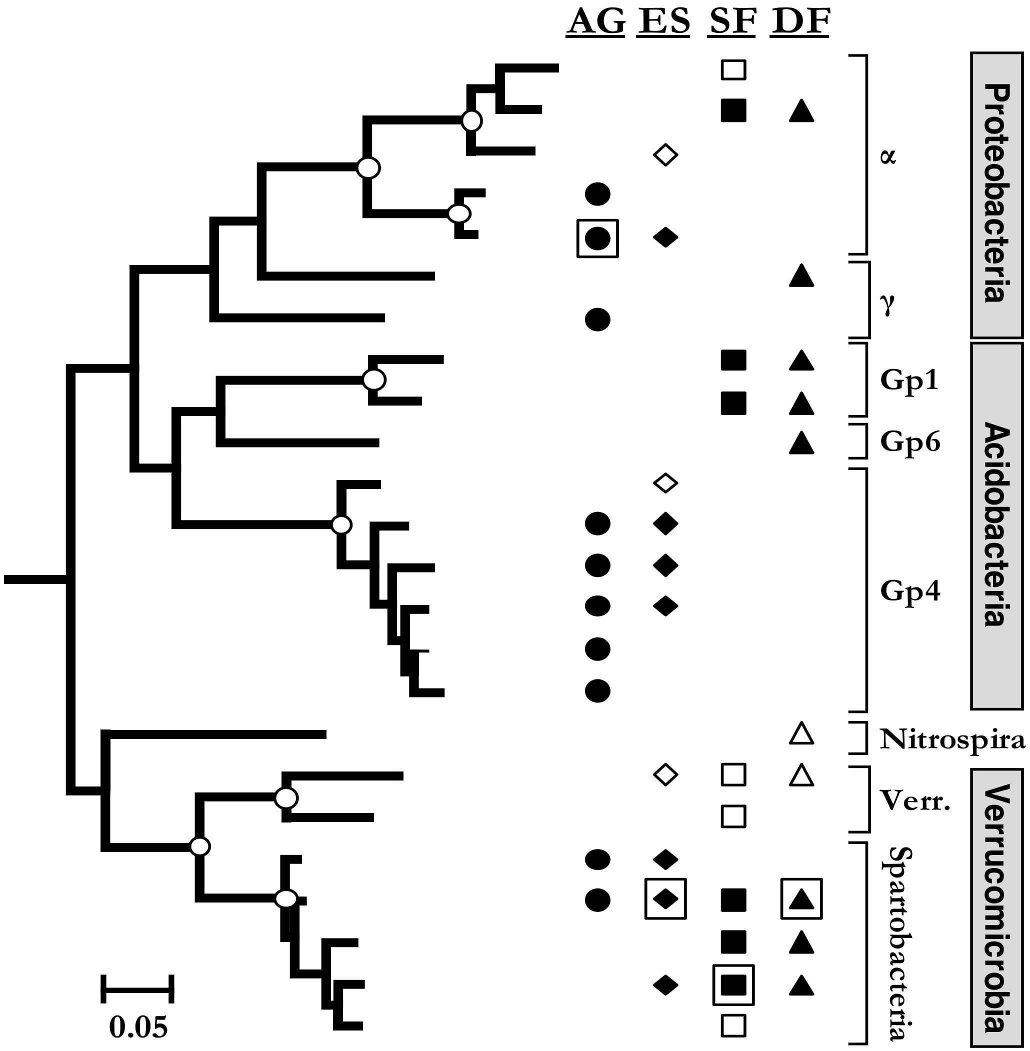
Taxonomic composition of bacterial communities changes with treatment



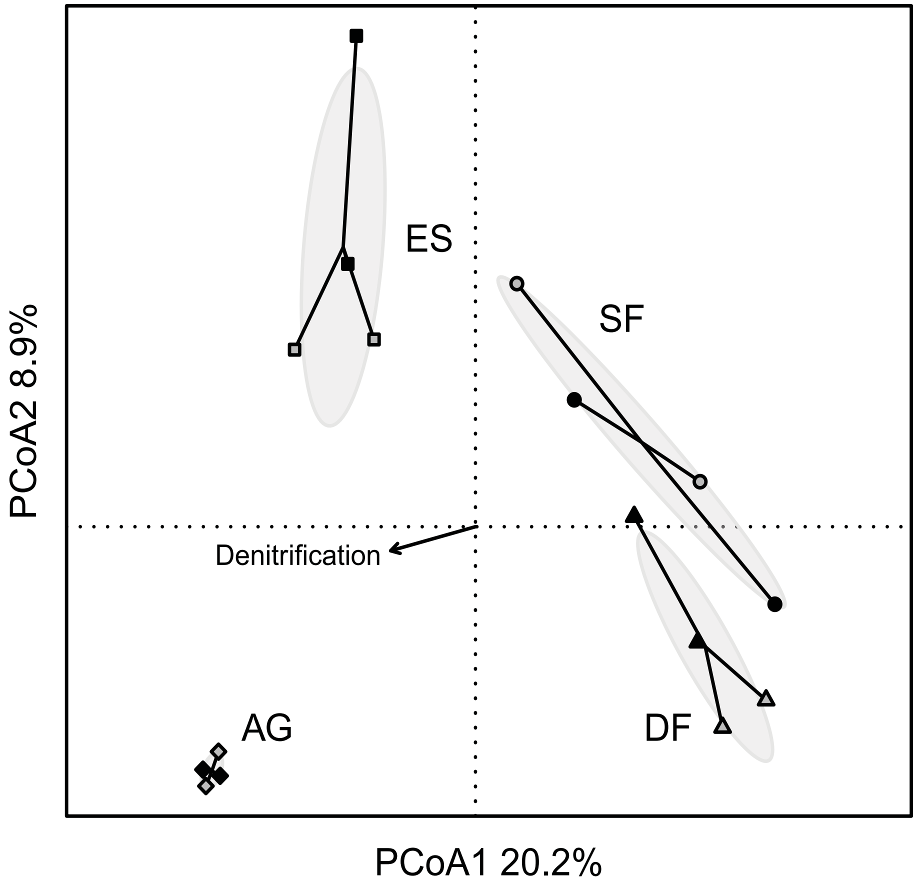
Biogeochemistry and bacterial community change concomitantly



Shifts within phyla differentiate communities



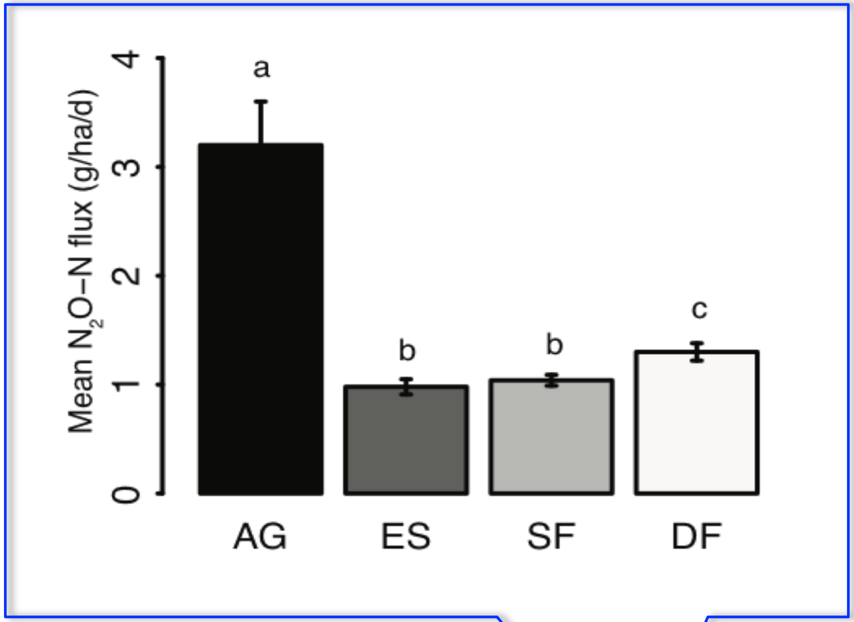
Functional potential changes with land management



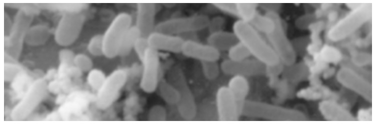
Analysis of 7058 genes annotated by MG-RAST against the SEED database

What is the relationship between denitrifying bacteria and N_2O in agricultural and native soils?

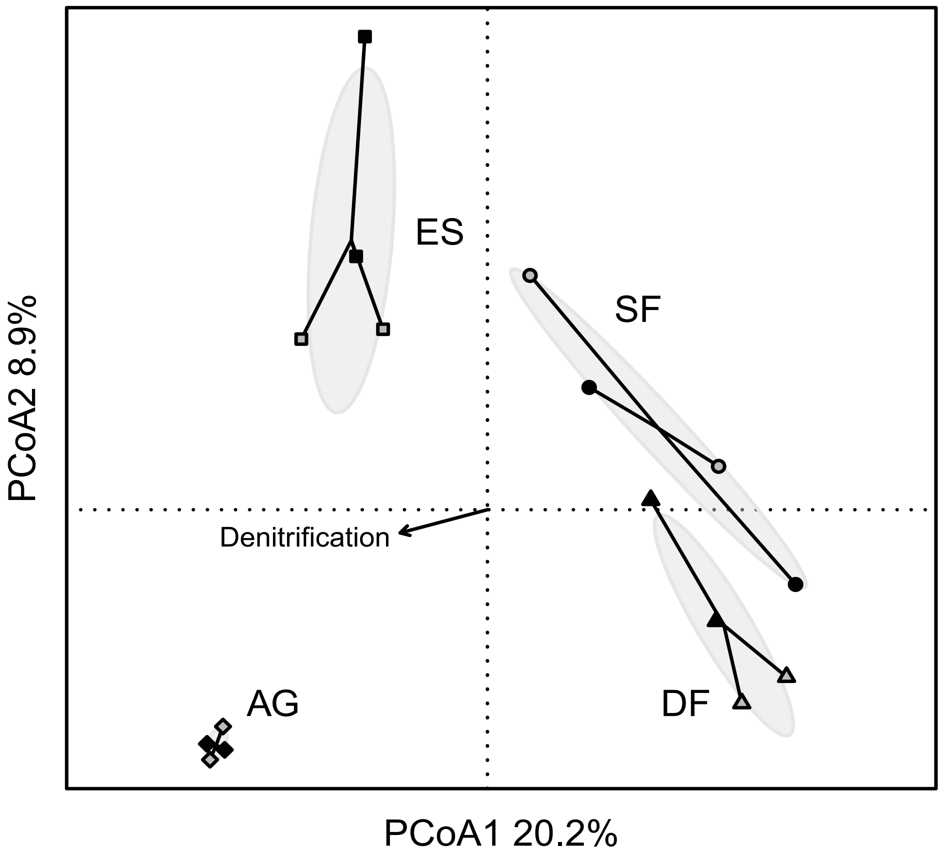
Denitrifying microbes



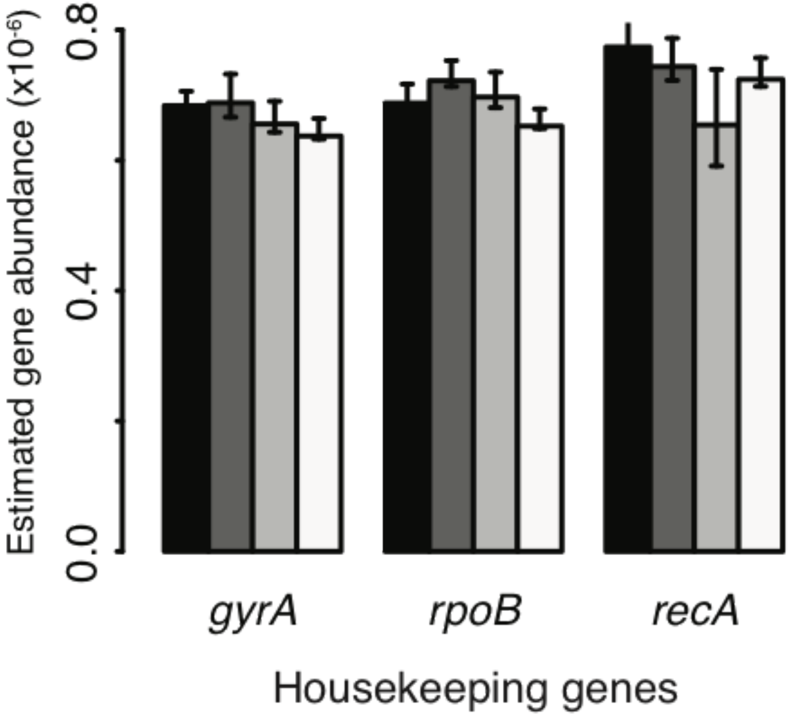
$$NO_2^- \xrightarrow{nirK} NO \xrightarrow{nor} N_2O \xrightarrow{nosZ} N_2$$



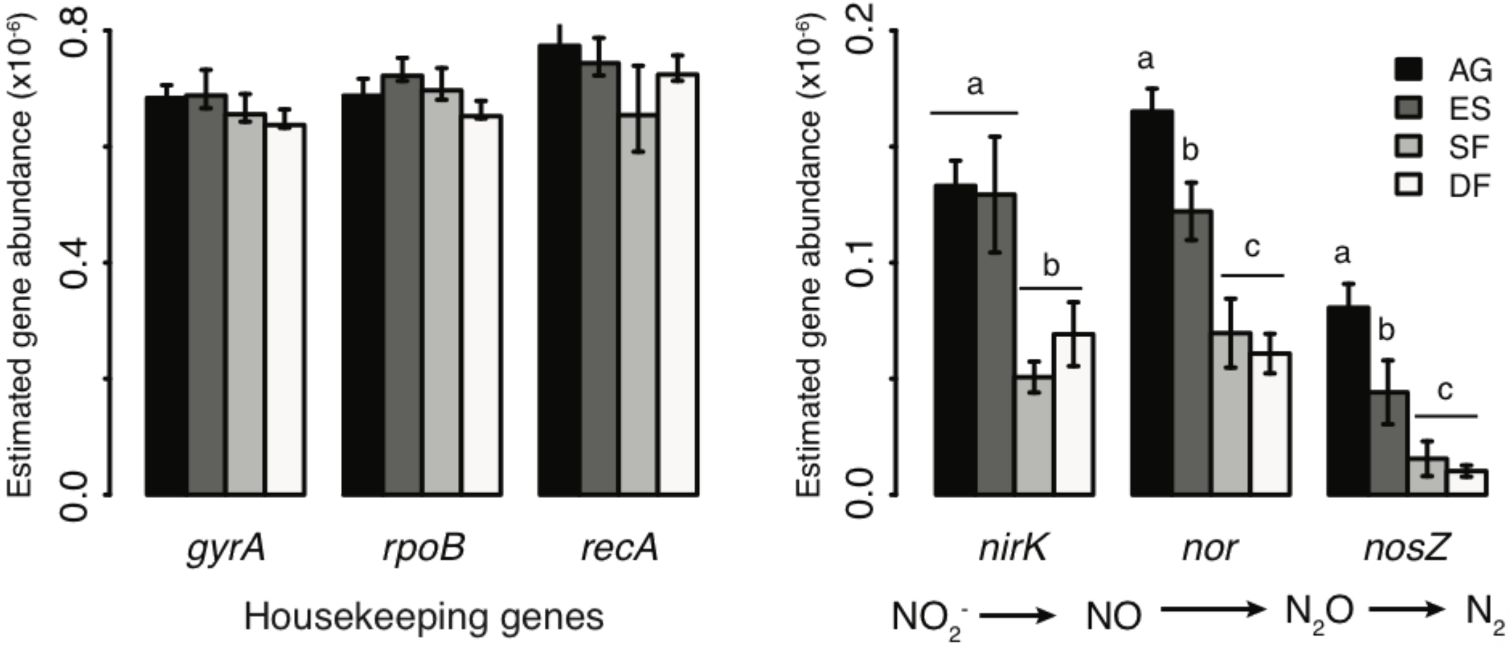
Nitrogen metabolism contributes to the differentiation of communities



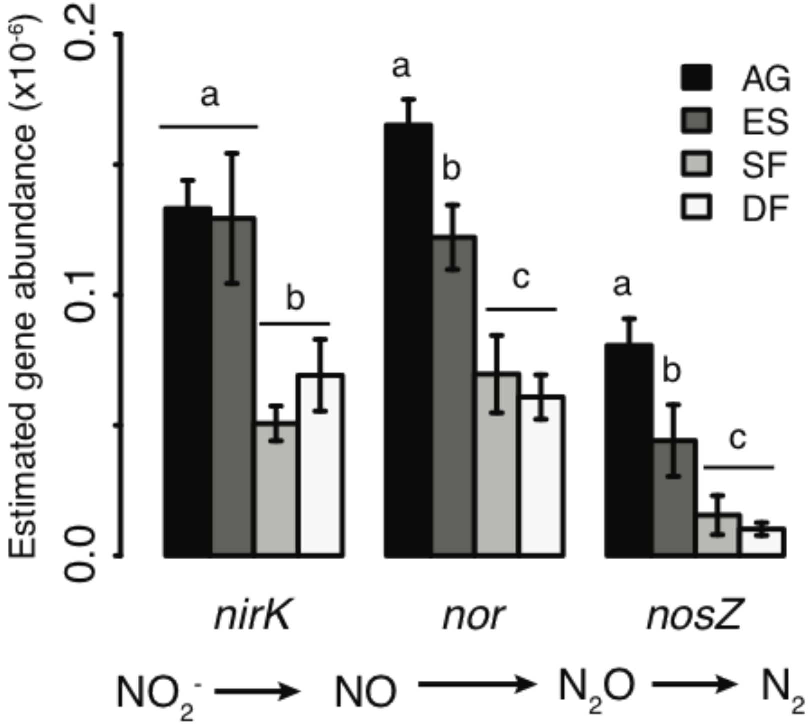
Gene abundances



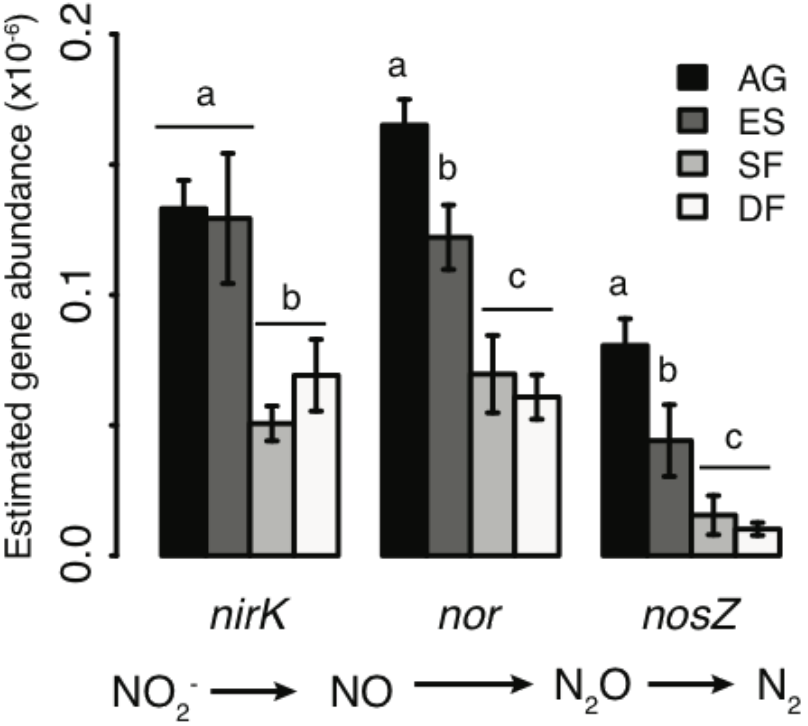
More denitrification potential in Ag soils

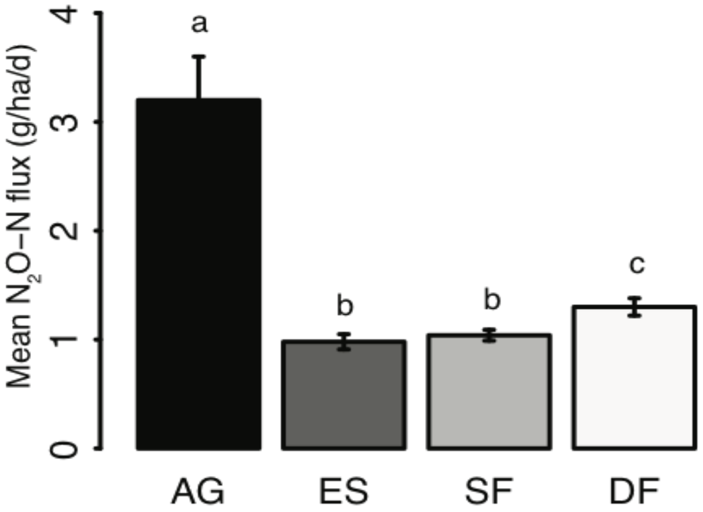


Not only abundance contributes to N_2O flux

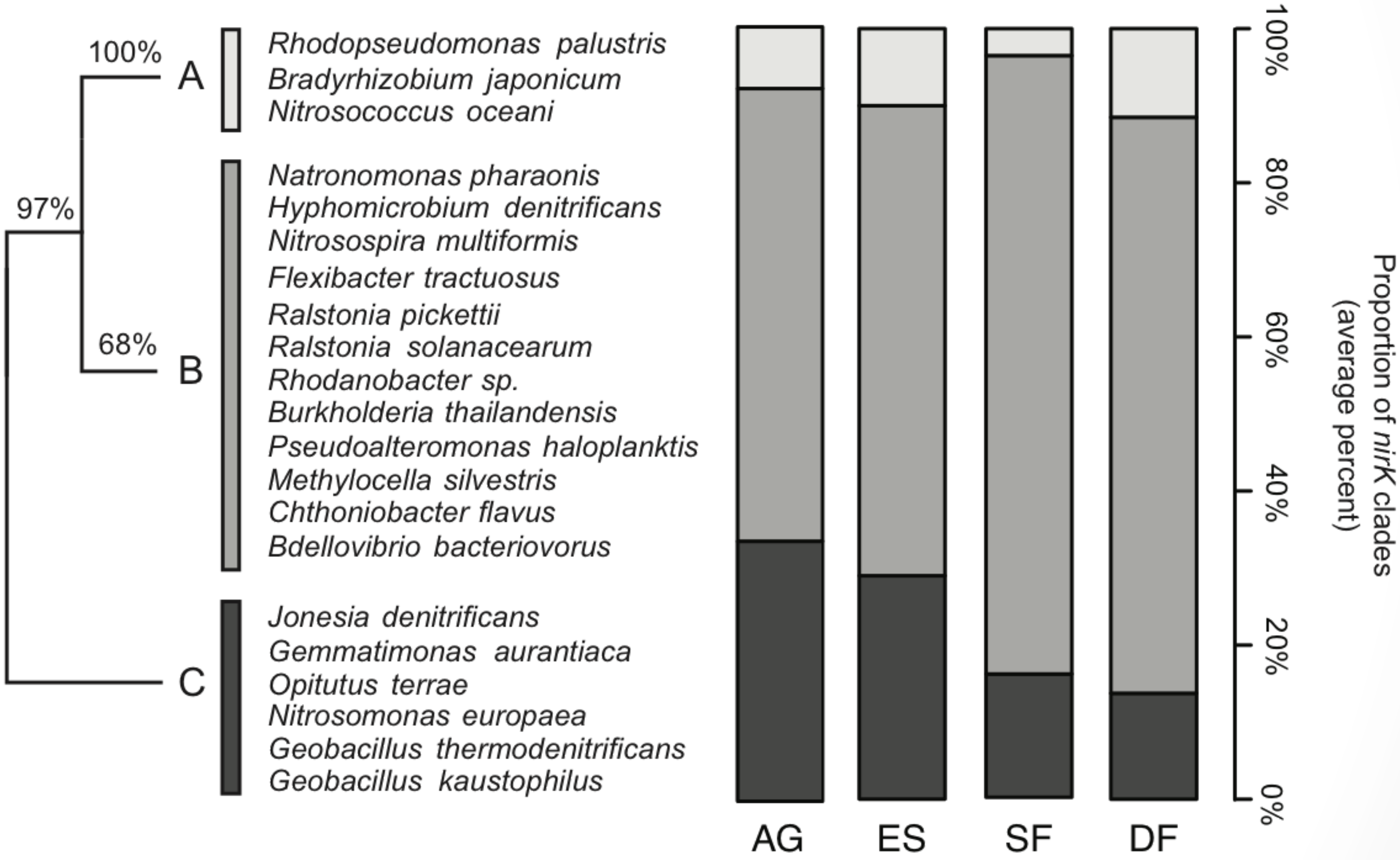


Not only abundance contributes to N_2O flux

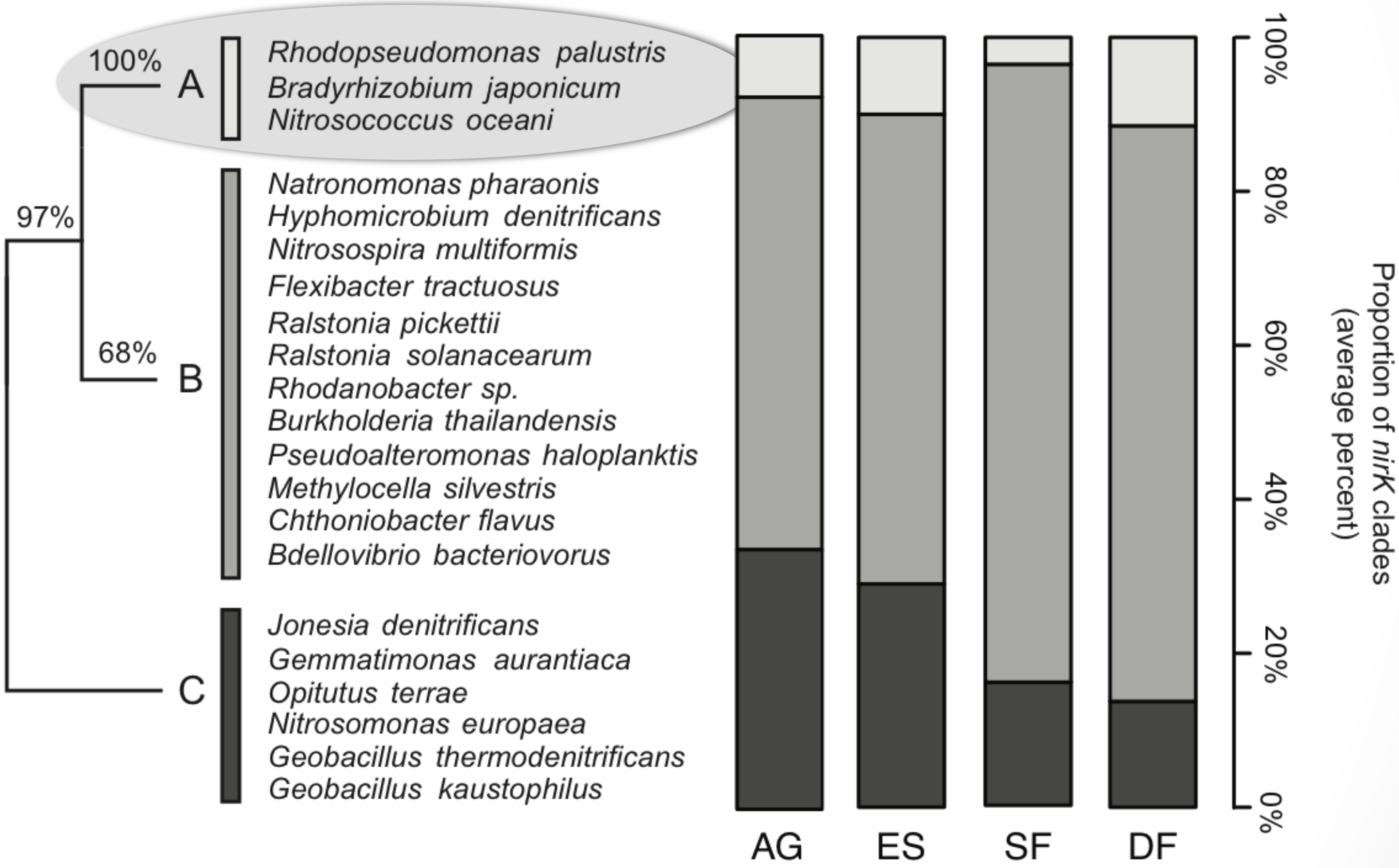




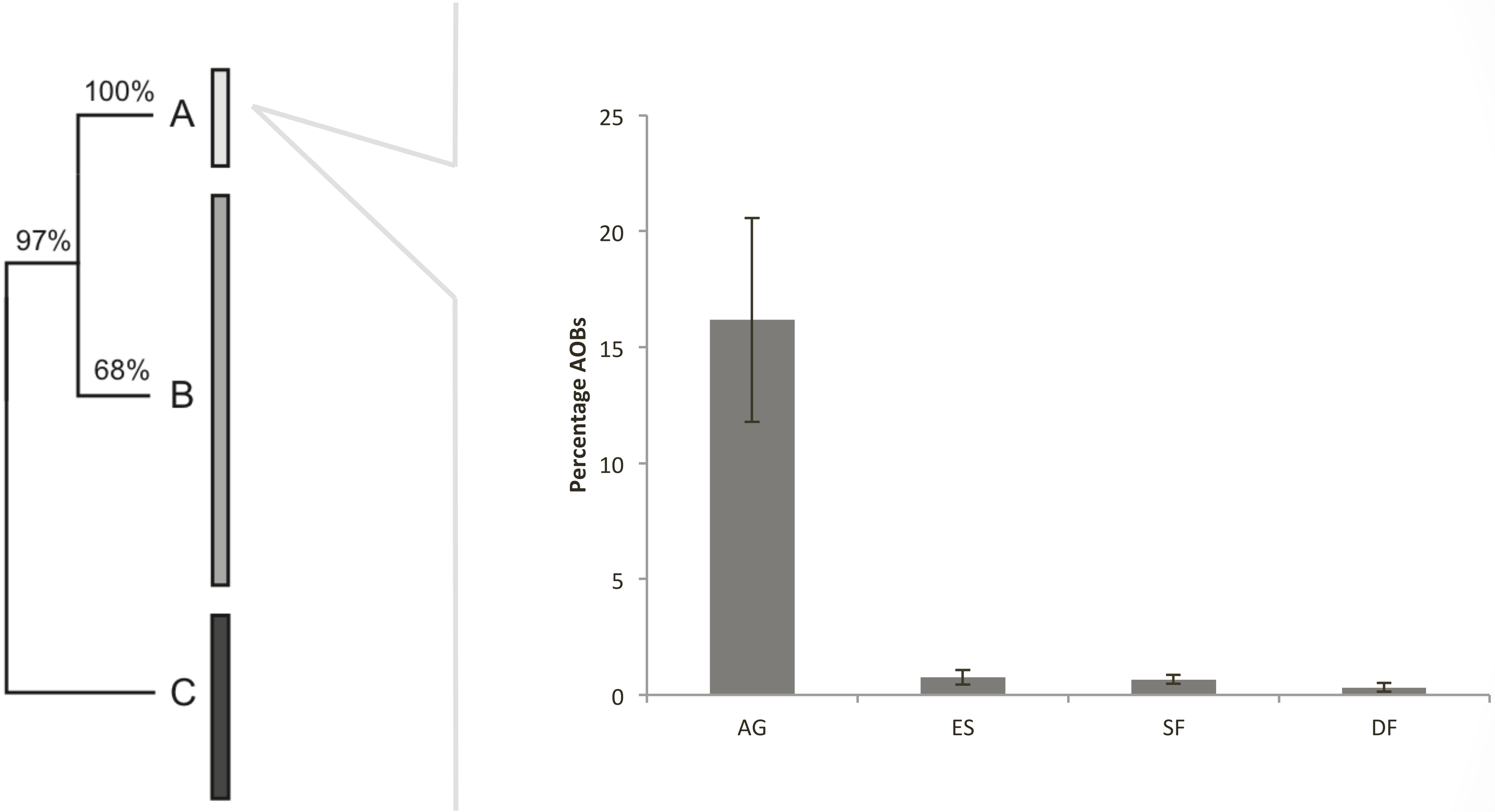
Denitrifier composition also changes



High denitrifer diversity



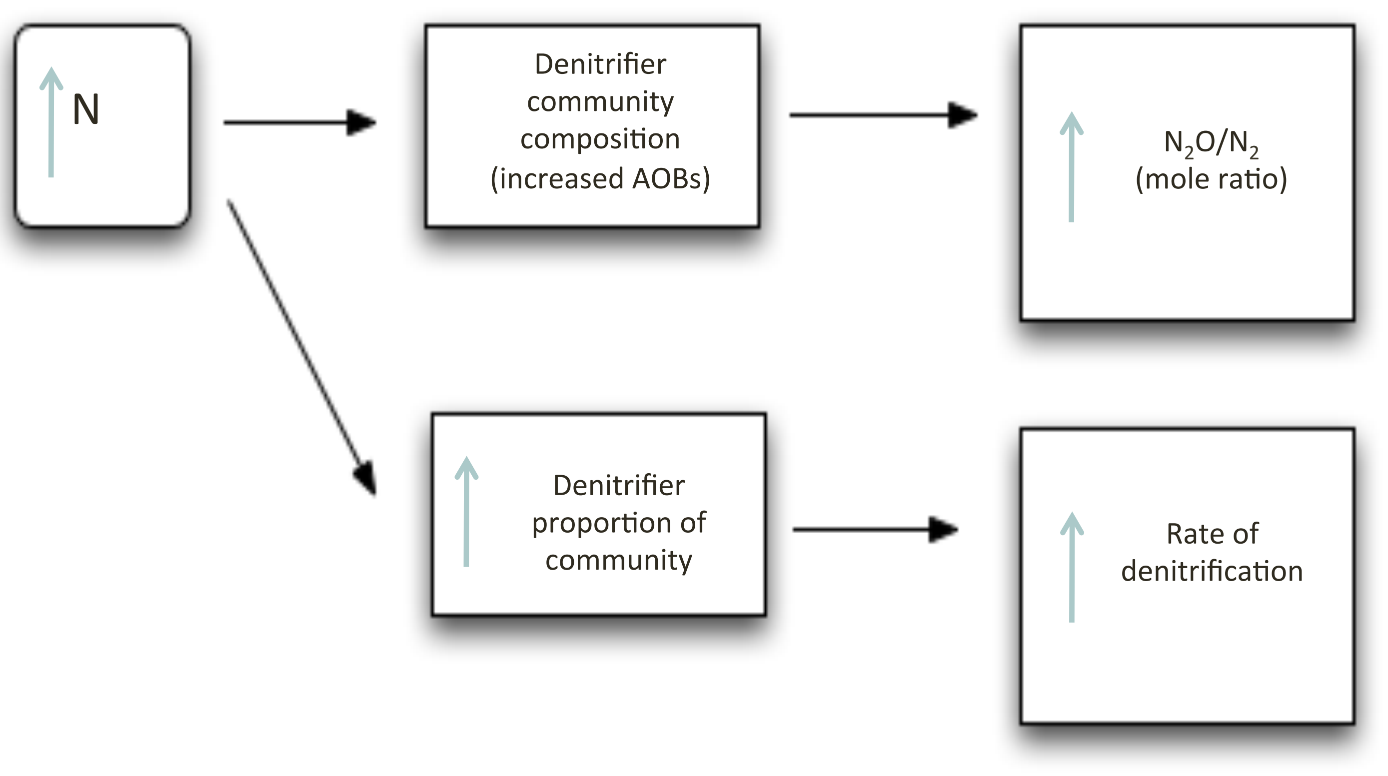
AOB proportion increased significantly in AG



Types of denitrification

	Conditions	Function
Heterotrophic denitrification $NO_3^- \longrightarrow NO_2^- \xrightarrow{\textit{nirK}} N_2O \dashrightarrow N_2$	Anoxic	Energetics
Autotrophic denitrification (AOBs) $NH_4^+ \xrightarrow{\overset{O_2}{H_2O}} NH_2OH \longrightarrow NO_2^- \xrightarrow{\textit{nirK}} N_2O$	Oxic	Nitrite detoxification

Microbial communities in agriculture are poised for denitrification



What is the relationship between denitrifying bacteria and N_2O in agricultural and native soils?

- Denitrifier abundance and composition changes with agricultural management. Communities more diverse than previously thought and poised for denitrification.
- May be potential for microbial mediation

Tom Schmidt
University of Michigan

Other studies on:
Methanotrophs
Biofuel crops with the GLBRC
Longitudinal study of KBS LTER sites
Effects of fertilization
HGBF studies of soil health in sub-Saharan Africa

mothur

- Updated versions are released every few months
  - First release, v1.1.0, March 2009
  - Last release, Version 1.39.5, March 2017
- Approaches

  Taxonomy OTUs Phylogeny
- Sequencing systems

  Sanger 454 Illumina
- Tutorials for OTU-based approach
  - 454 http://www.mothur.org/wiki/454_SOP
  - Illumina http://www.mothur.org/wiki/MiSeq_SOP

mothur and QIIME

- Both are open source and on github
- Both aim to enable advances in microbial ecology and are actively maintained and developed
- Both require alignment
- mothur is not an acronym
- While QIIME connects multiple tools, mothur reimplements algorithms, so that it is all one program

Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities

Schloss et al, AEM, 2009

- mothur in C++, QIIME connections in Python
- QIIME has mothur and mothur has Unifrac, but the default behavior for mothur is to do clustering based on sequence distance
- mothur’s clustering can be very memory intensive, Uparse as used in QIIME requires less memory

Outline

- Case study – microbial communities in soil

- mothur workflow

- mothur MiSeq SOP

http://www.mothur.org/wiki/MiSeq_SOP

How to get your data in, get your data out (an OTU table), and get help

- Some preliminary statistics

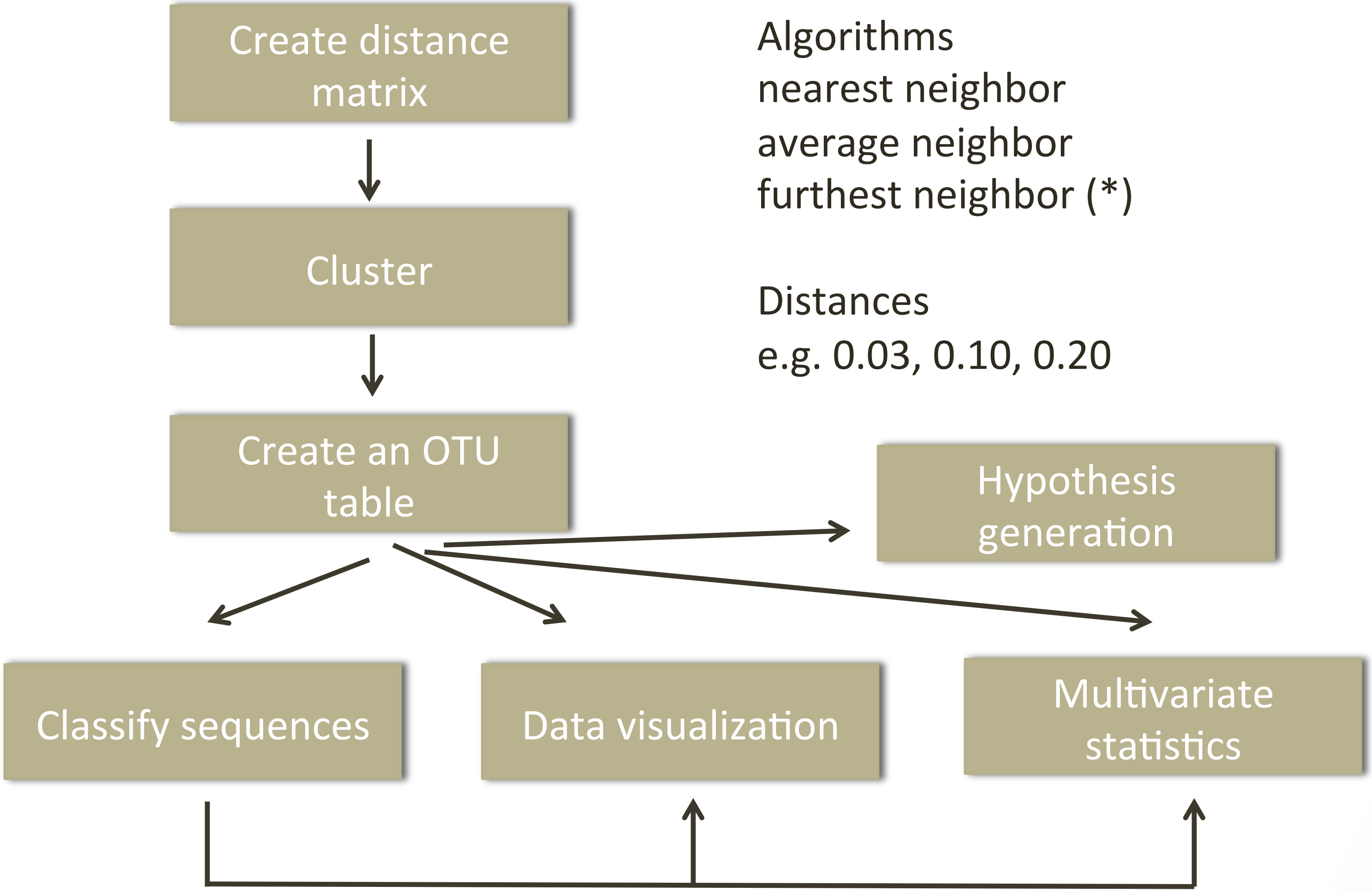
mothur workflow

(Schloss 2009; Schloss 2010; Schloss 2013; Pruesse et al. 2007, doi:10.1093/nar/gkm864; Pruesse et al. 2012, doi:10.1093/bioinformatics/bts252)

Prepare the data ↓	FASTQ files, information file, assemble paired ends
Quality filter ↓	Remove ambiguous base pairs & any sequences longer than expected Algorithms detailed in Kozich et al. 2013
Reduce data set ↓	Create a file with just unique sequences & track what samples they're in
Align sequences & quality filter ↓	Align to a reference, reassess quality & reduce data size again
Remove chimeras & non-bacterial ↓	Several chimera removal options & remove sequences not classified as bacterial

mothur workflow

(Schloss 2009; Schloss 2010; Schloss 2013; Pruesse et al. 2007, doi:10.1093/nar/gkm864; Pruesse et al. 2012, doi:10.1093/bioinformatics/bts252)



Create distance matrix

Cluster

Create an OTU table

Hypothesis generation

Classify sequences

Data visualization

Multivariate statistics

Algorithms
nearest neighbor
average neighbor
furthest neighbor (*)

Distances
e.g. 0.03, 0.10, 0.20

Working in mothur

Learn how to:

- Get in (get data in)
- Get out (Come home with an OTU table or on it)
- Get help

https://github.com/tracykteal/tutorials/tree/master/mothur

http://www.mothur.org/wiki/MiSeq_SOP

Options for running mothur

- http://www.mothur.org/wiki/Download_mothur
- Windows, Mac, Linux
    - Interactive
    - Batch
    - GUI
- However, for a real analysis, you'll need significant computation, so likely cloud or HPC resources. Mothur needs memory in particular.

mothur SOP data

Looking at the effect of normal variation on the gut microbiome and host health

- Collected fresh feces from mice on a daily basis for 365 days post weaning.
- During the first 150 days post weaning (dpw), nothing was done to the mice except allow them to eat, get fat, and be merry. They were curious whether the rapid change in weight observed during the first 10 dpw affected the stability of the microbiome compared to the microbiome observed between days 140 and 150.
- To make this tutorial easier to execute, they are providing only part of the data - you are given the flow files for one animal at 10 time points (5 early and 5 late). In addition, to sequencing samples from mice fecal material, they resequenced a mock community composed of genomic DNA from 21 bacterial strains. We will use the 10 fecal samples to look at how to analyze microbial communities and the mock community to measure the error rate and its effect on other analyses.

mothur tutorial

- Is your sample coverage sufficient for meaningful analyses?

- Do mouse fecal microbiota differ between weanling and adult mice?
  - alpha, beta diversity
  - OTUs responsible?
- Is variation in fecal microbiota greater among weanlings than adults?

Use both graphical and statistical tools to answer each question.
Ignore the obvious pseudoreplication and simultaneous lack of replication.