Rhizosphere Microbiome: Manipulation and Investigation

Kevin Panke-Buisse
PhD Candidate

Kao-Kniffin Lab
School of Integrative Plant Science
Horticulture Section

Presentation Summary

- **Part I: Introduction to the Plant-Microbe Holobiont**
  - The soil microbiome
  - Plant-microbe interactions
- **Part II: Microbiome Manipulation**
  - Group selection
  - Reproducibility of effects on plant function
- **Part III: Investigating the Microbial Component**

Part I: Introduction to the Plant-Microbe Holobiont

- Holobiont – The functional “whole” or community formed by an organism and its associated microbiota[2]

References Slide is at the end of the presentation

The Plant-Microbe Holobiont

Soil microbial diversity

- Estimates range from ~1000 – several million bacterial species per gram of soil.

- Genetic diversity is the marble to evolution’s Michelangelo

Plant-microbe interactions

- Best characterized:
  - Disease
  - Direct symbiosis
    - Rhizobia
  - Plant Growth Promoting Bacteria
- Poorly Understood:
  - Complex communities
  - Temporal shifts
  - Role in, and response to, stress

Part II: Microbiome Manipulation

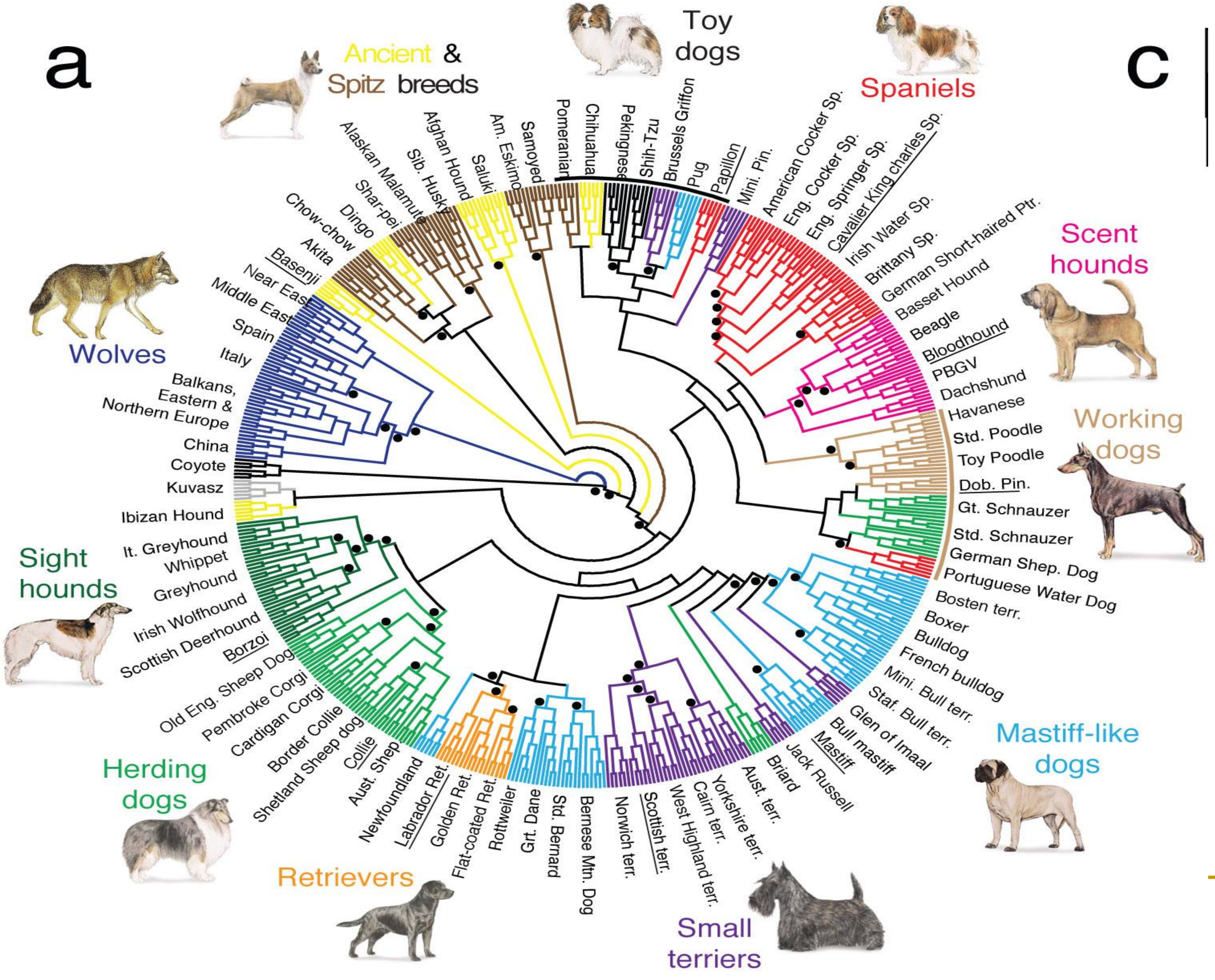
Can a microbiome shape host-plant development?

Artificial selection

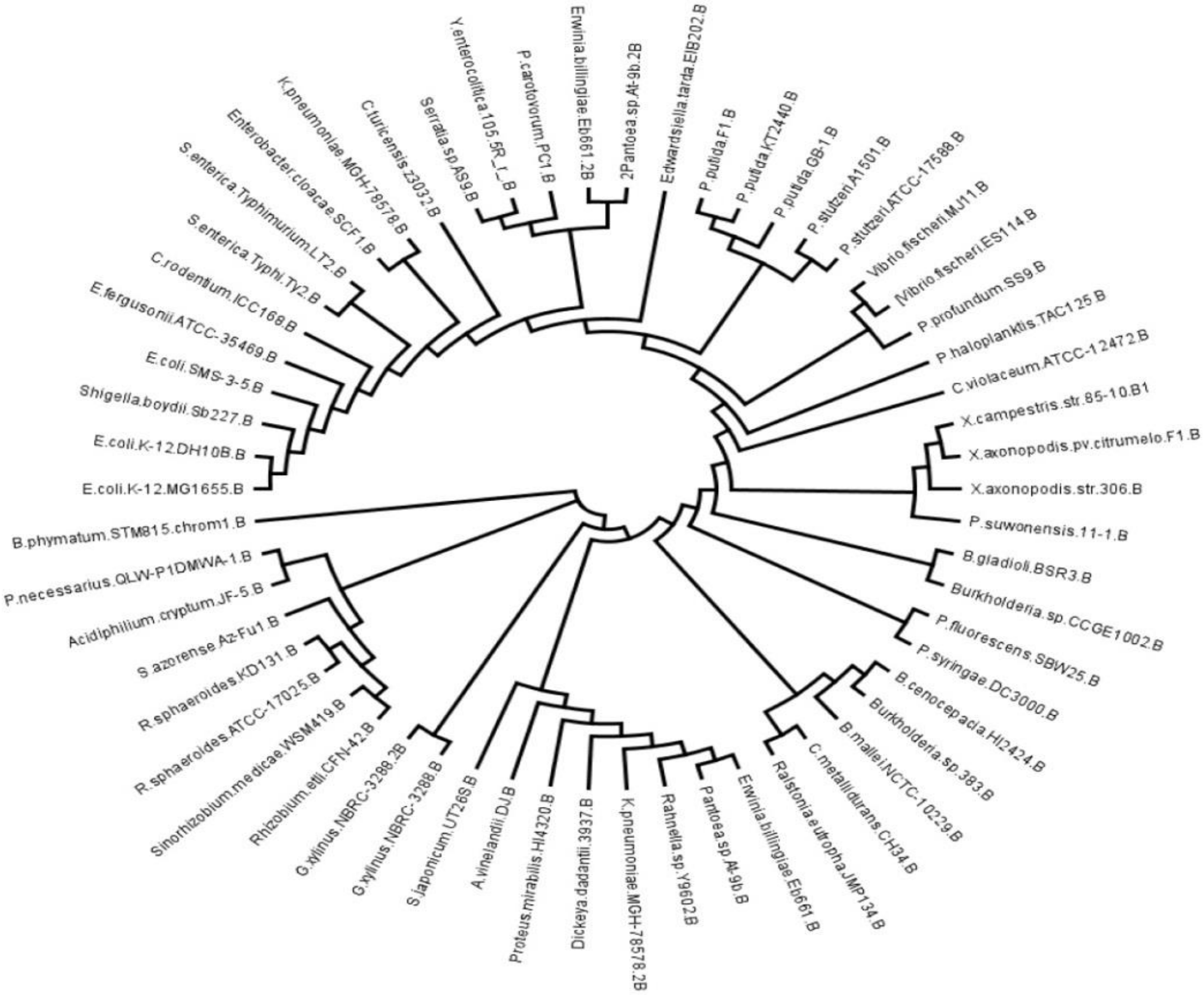
- “Genetic diversity is the marble to evolution’s Michelangelo”

- Can we pick up the sculptors tools?

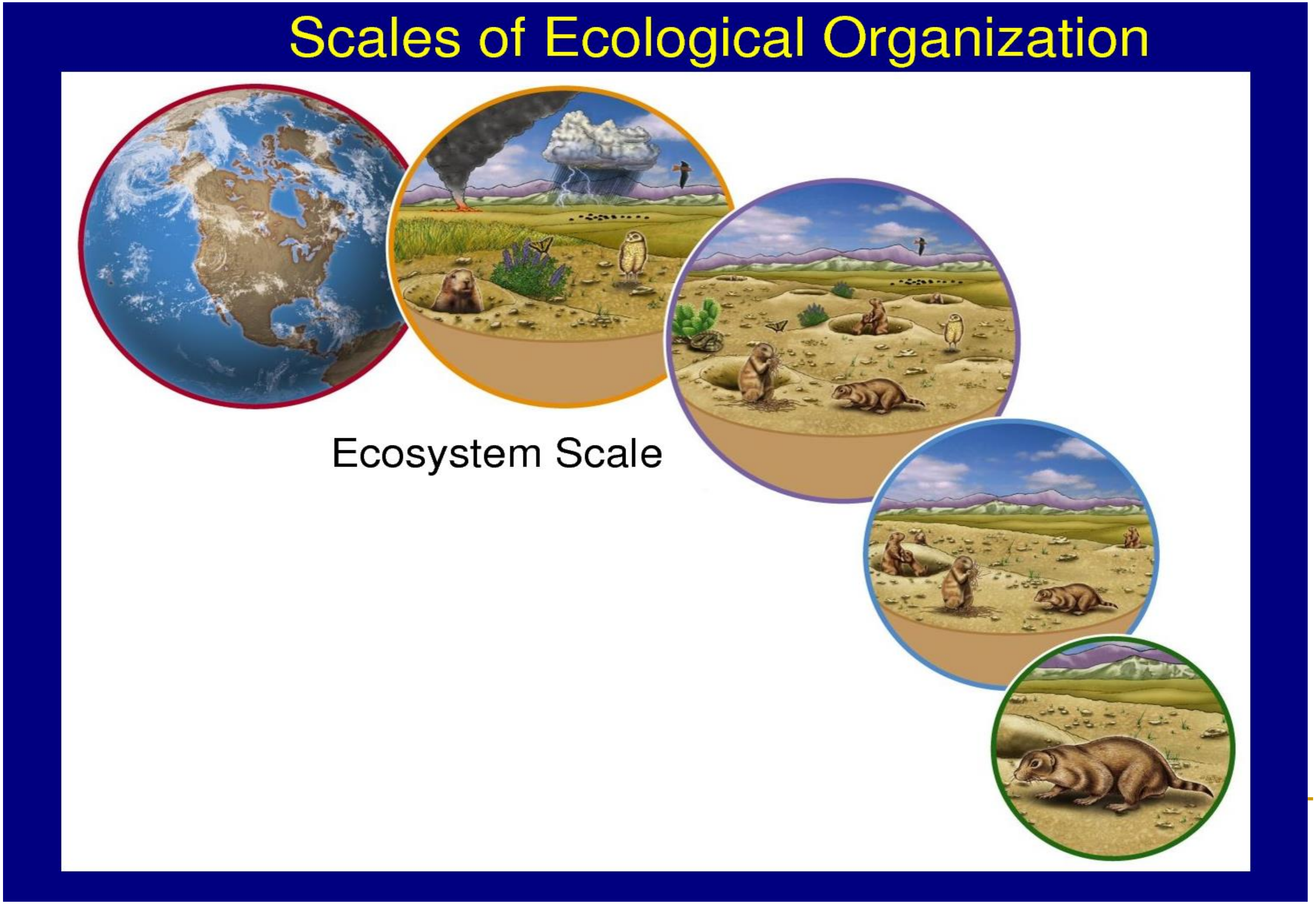
Yes, we can. However,



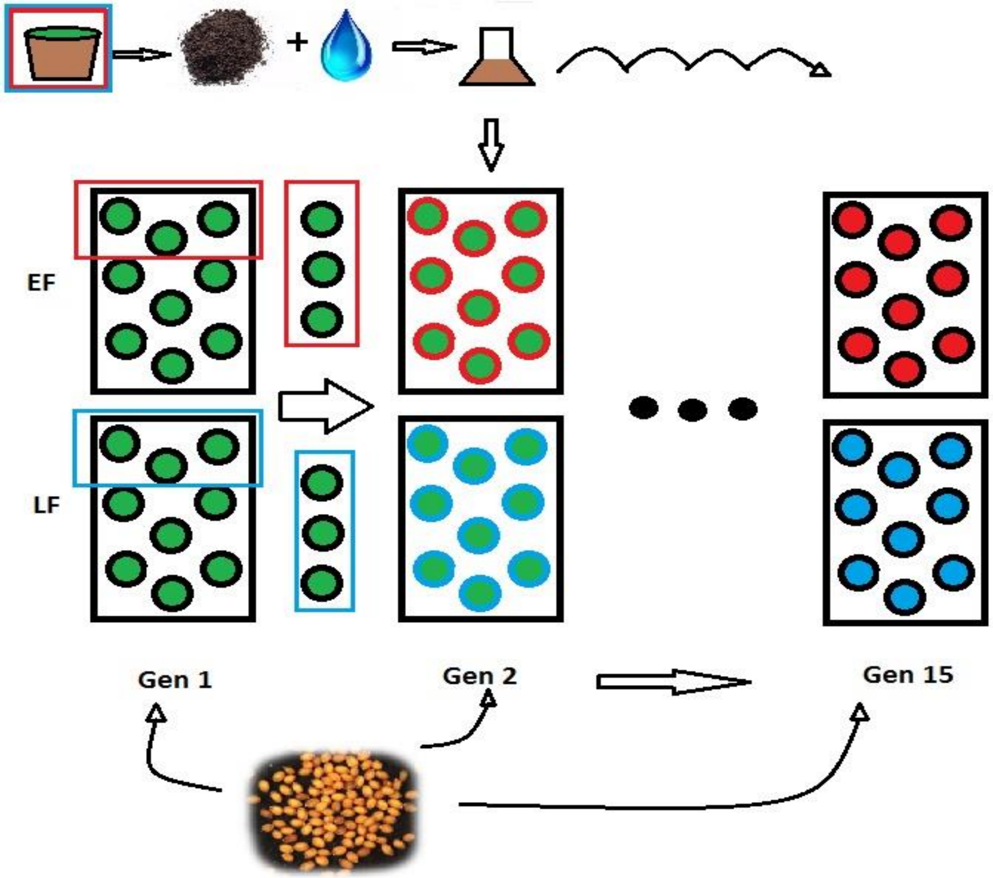
Microbiomes require another approach

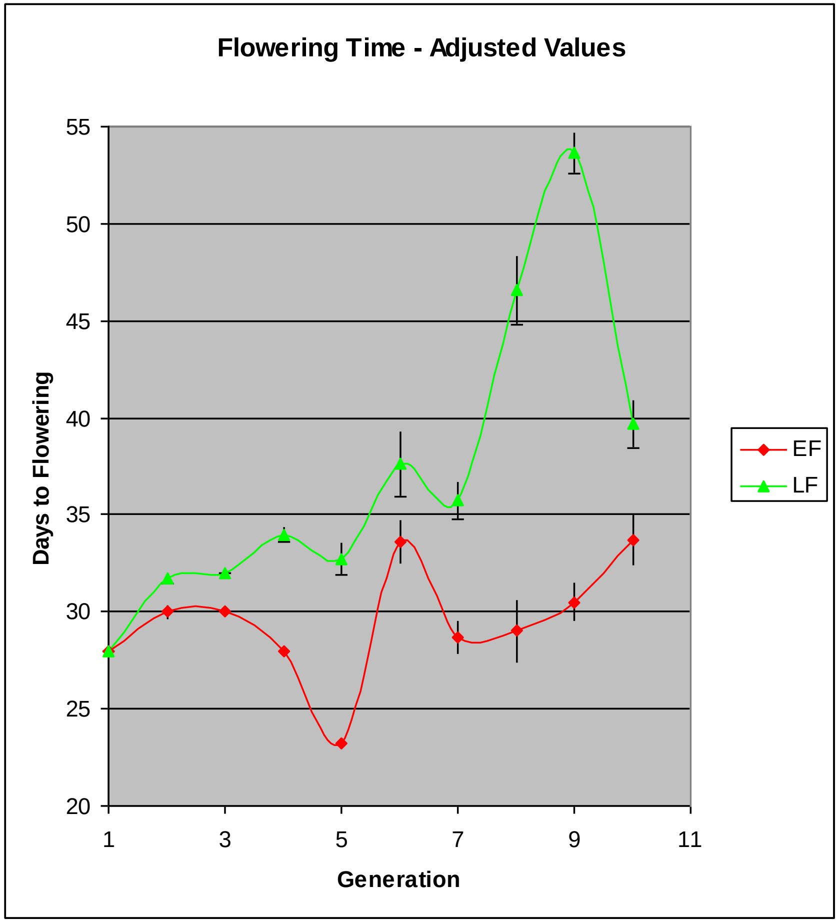


Resolution/Scale



Iterative Selection Process



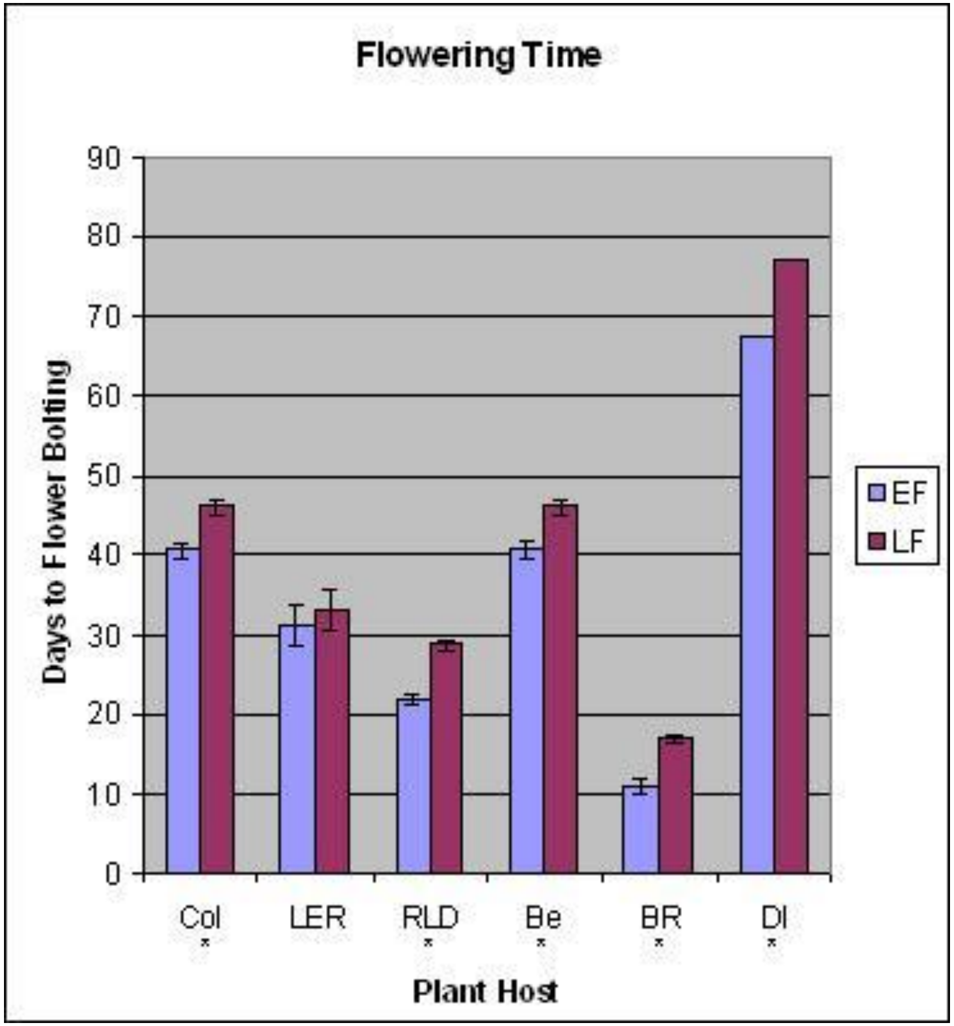


Selection process adapted from Swenson et al., 2000 [9]

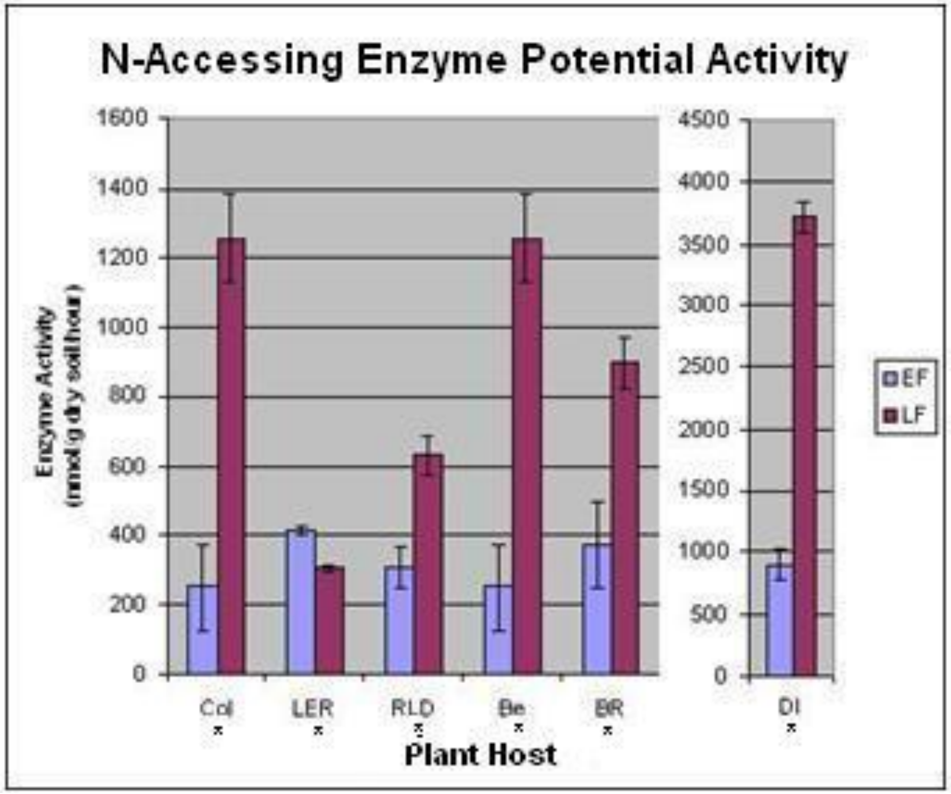
Reproduction of function

- Early- and late-flowering microbiomes were inoculated into novel plant hosts[4]
  - *Brassica rapa*
  - *Arabidopsis thaliana* genotypes
    - LER
    - BE
    - RLD
    - Col-0
  - *Digitaria ischaemum*

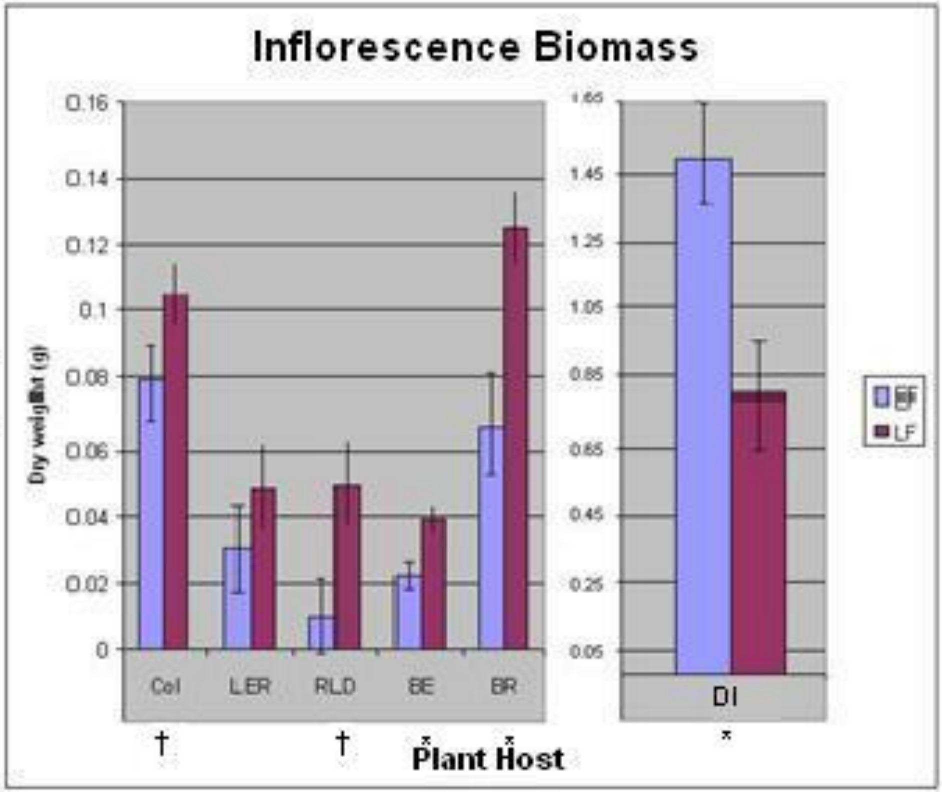
Flowering Time



Exoenzymes and Reproductive Biomass



N-Accessing Enzyme Potential Activity



Inflorescence Biomass

Part III: Investigating the Microbial Component

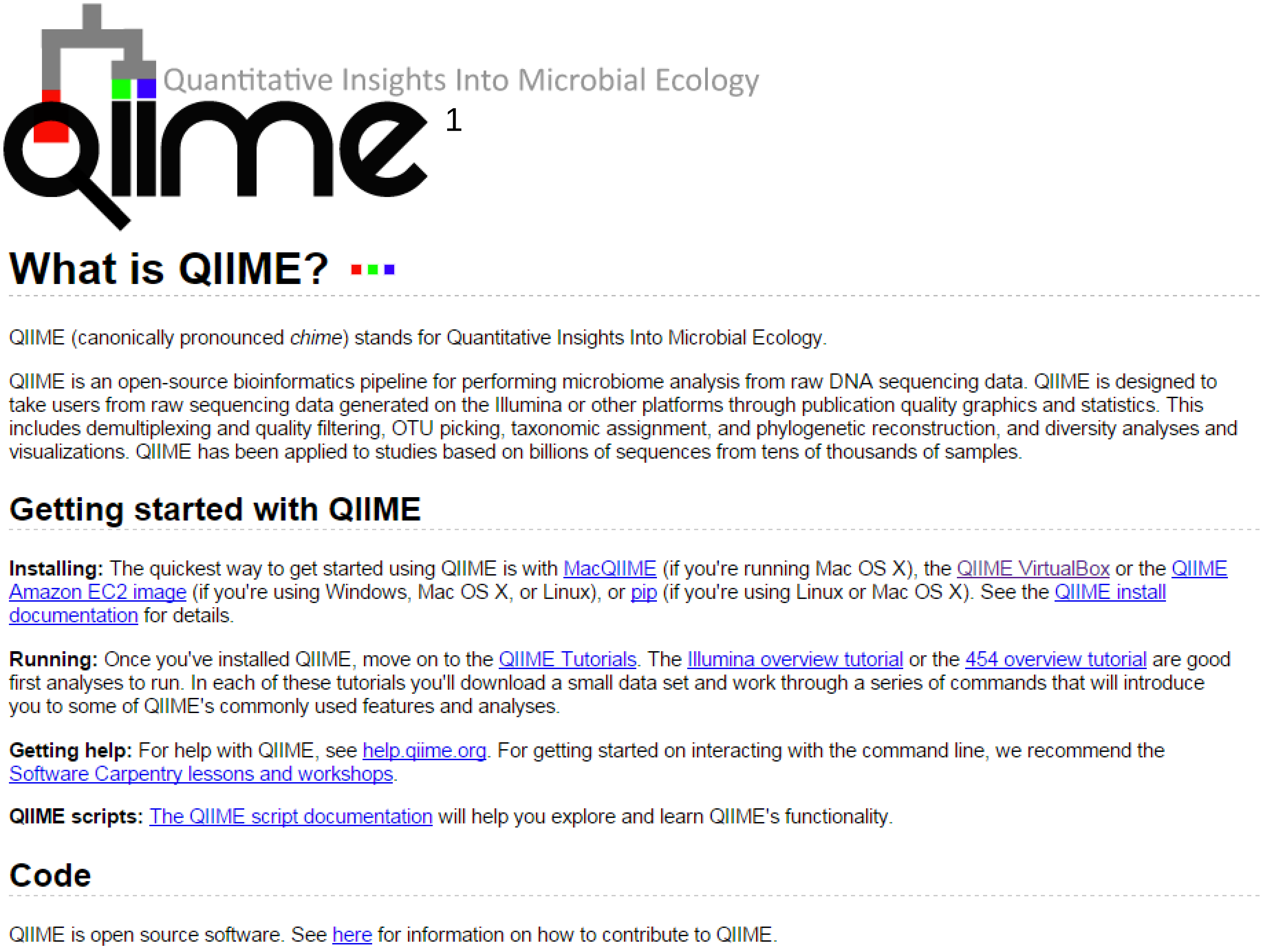
- Full Disclosure:
  - I am not a computational biologist
  - My informatics skills are not fluent
  - I can code *lightly* in perl
  - My R skills are intermediate
  - I spend A LOT of time trying to figure out what I’m doing wrong

Tools

- Quantitative Insights Into Microbial Ecology (QIIME)
- R
- Cytoscape
- Photoshop

QIIME[1]

- Heavy Lifting
    - From raw fastq to OTU biom table
    - Scripts for 2D and 3D visualization
    - Alpha/Beta diversity calculation
- Continuously under development
    - Support for Illumina, paired-end reads
- Python-based
- Highly active dev community



Quantitative Insights Into Microbial Ecology

1

What is QIIME?

QIIME (canonically pronounced *chime*) stands for Quantitative Insights Into Microbial Ecology.

QIIME is an open-source bioinformatics pipeline for performing microbiome analysis from raw DNA sequencing data. QIIME is designed to take users from raw sequencing data generated on the Illumina or other platforms through publication quality graphics and statistics. This includes demultiplexing and quality filtering, OTU picking, taxonomic assignment, and phylogenetic reconstruction, and diversity analyses and visualizations. QIIME has been applied to studies based on billions of sequences from tens of thousands of samples.

Getting started with QIIME

Installing: The quickest way to get started using QIIME is with MacQIIME (if you're running Mac OS X), the QIIME VirtualBox or the QIIME Amazon EC2 image (if you're using Windows, Mac OS X, or Linux), or pip (if you're using Linux or Mac OS X). See the QIIME install documentation for details.

Running: Once you've installed QIIME, move on to the QIIME Tutorials. The Illumina overview tutorial or the 454 overview tutorial are good first analyses to run. In each of these tutorials you'll download a small data set and work through a series of commands that will introduce you to some of QIIME's commonly used features and analyses.

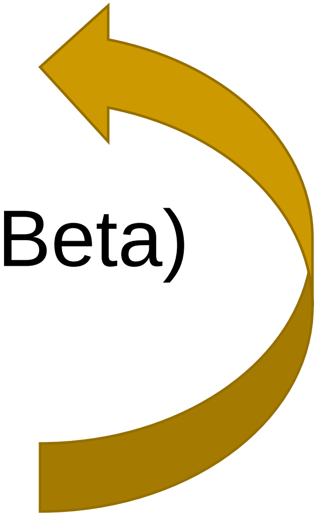
Getting help: For help with QIIME, see help.qiime.org. For getting started on interacting with the command line, we recommend the Software Carpentry lessons and workshops.

QIIME scripts: The QIIME script documentation will help you explore and learn QIIME's functionality.

Code

QIIME is open source software. See here for information on how to contribute to QIIME.

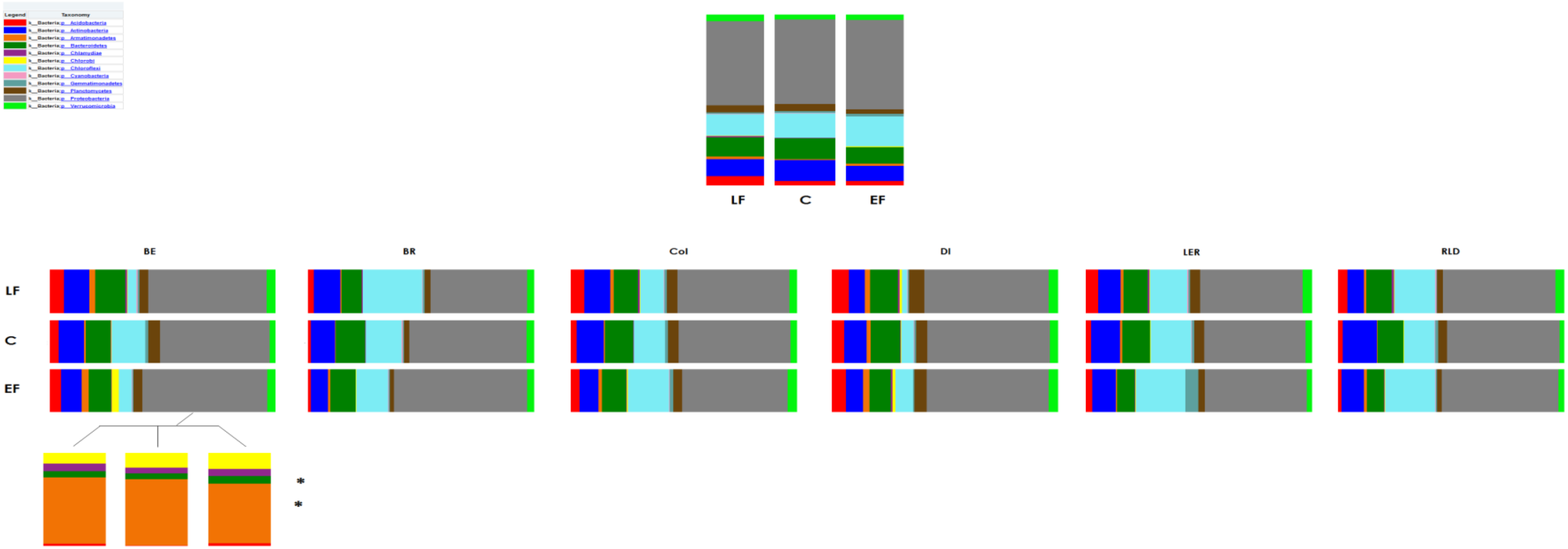
QIIME Workflow



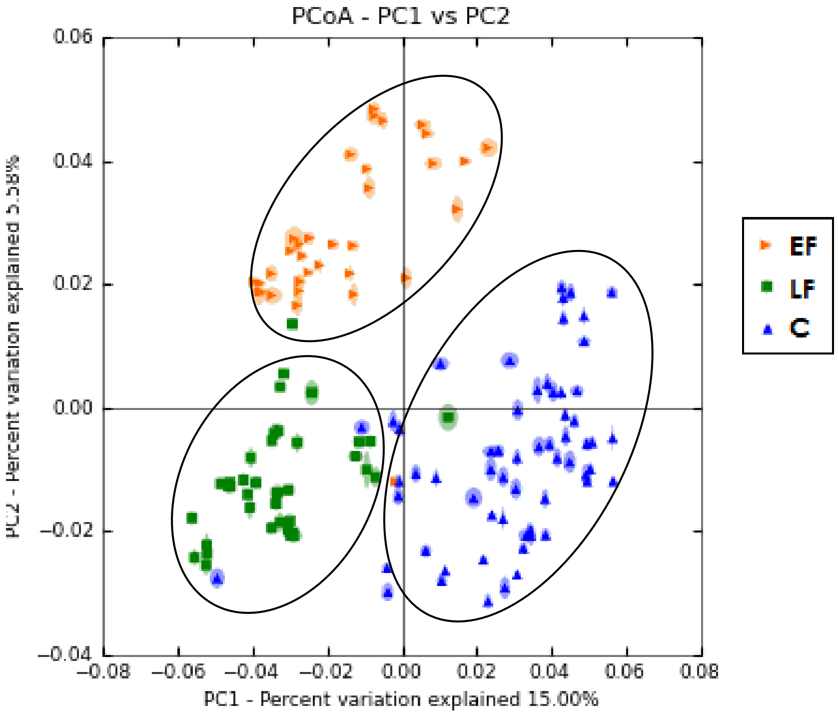
- Demultiplex
- Join reads
- Pick OTUs (Many options)
- Calculate Diversity (Alpha/Beta)
- Visualize Data
- Statistics
- Export OTU tables for external uses
  - R, Cytoscape, etc.

Visualizing community composition

- Relative abundance bar charts (QIIME)



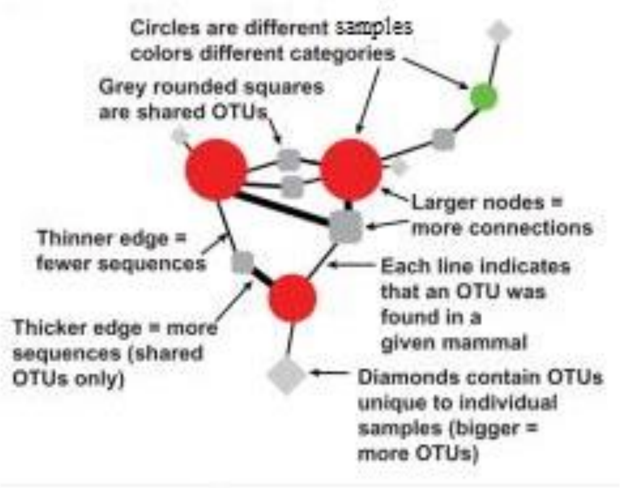
Community Structure



- Major Phyla (Classes)
  - Late Flowering
    - Acidobacteria
    - Cyanobacteria
    - Verrucomicrobia
    - Gamma-Proteobacteria
  - Early Flowering
    - Beta, Delta-Proteobacteria
    - Bacteroidetes
    - Planctomycetes
    - Cyanobacteria

Community Structure

- Cytoscape[5, 6, 8]
  - Network mapping



Circles are different samples
colors different categories

Grey rounded squares
are shared OTUs

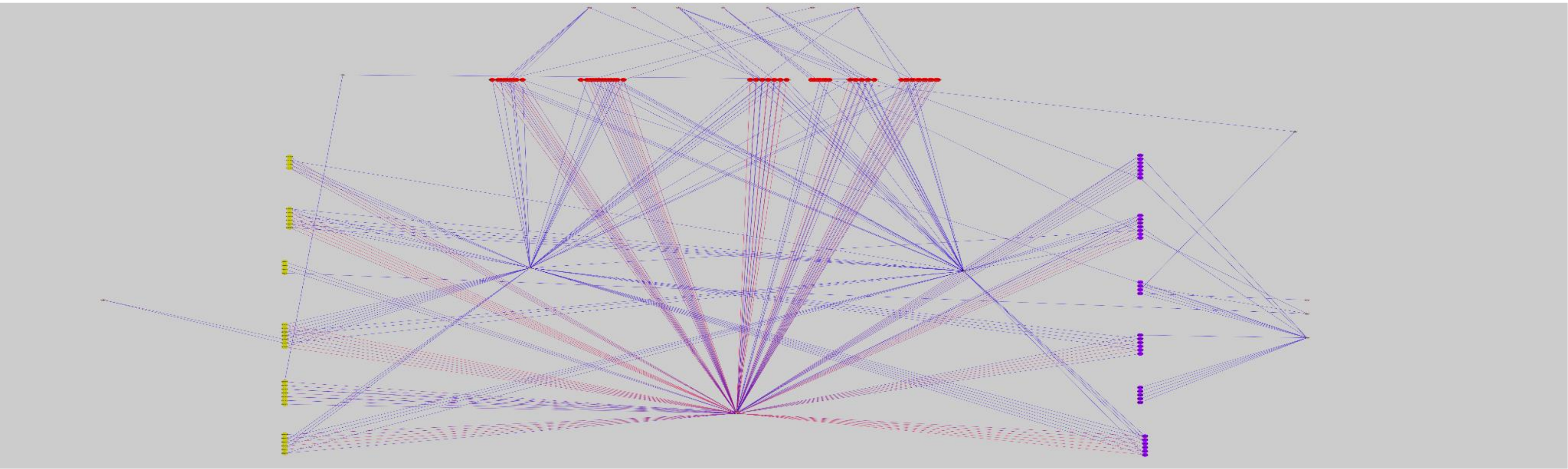
Larger nodes =
more connections

Thinner edge =
fewer sequences

Each line indicates
that an OTU was
found in a
given mammal

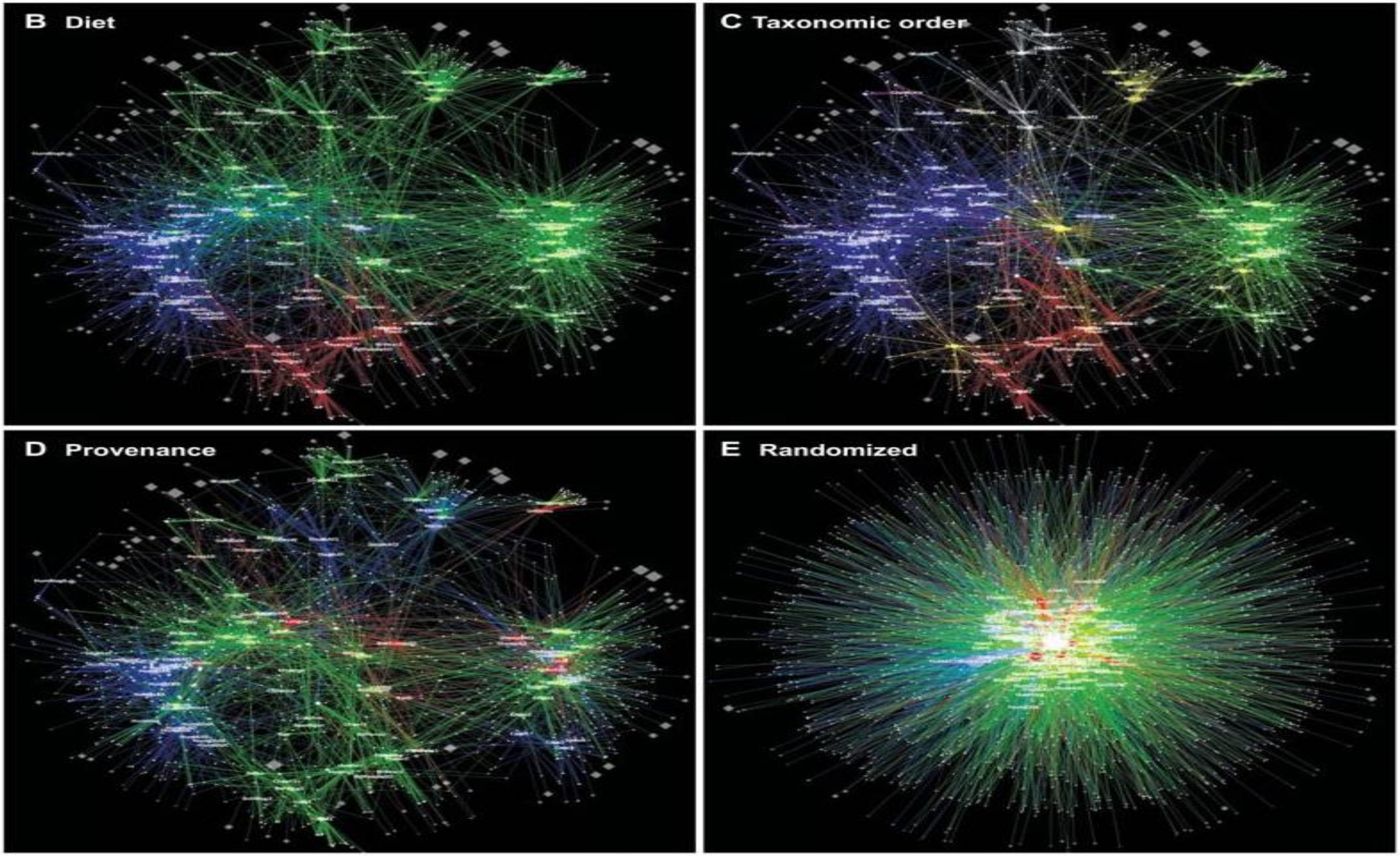
Thicker edge = more
sequences (shared
OTUs only)

Diamonds contain OTUs
unique to individual
samples (bigger =
more OTUs)



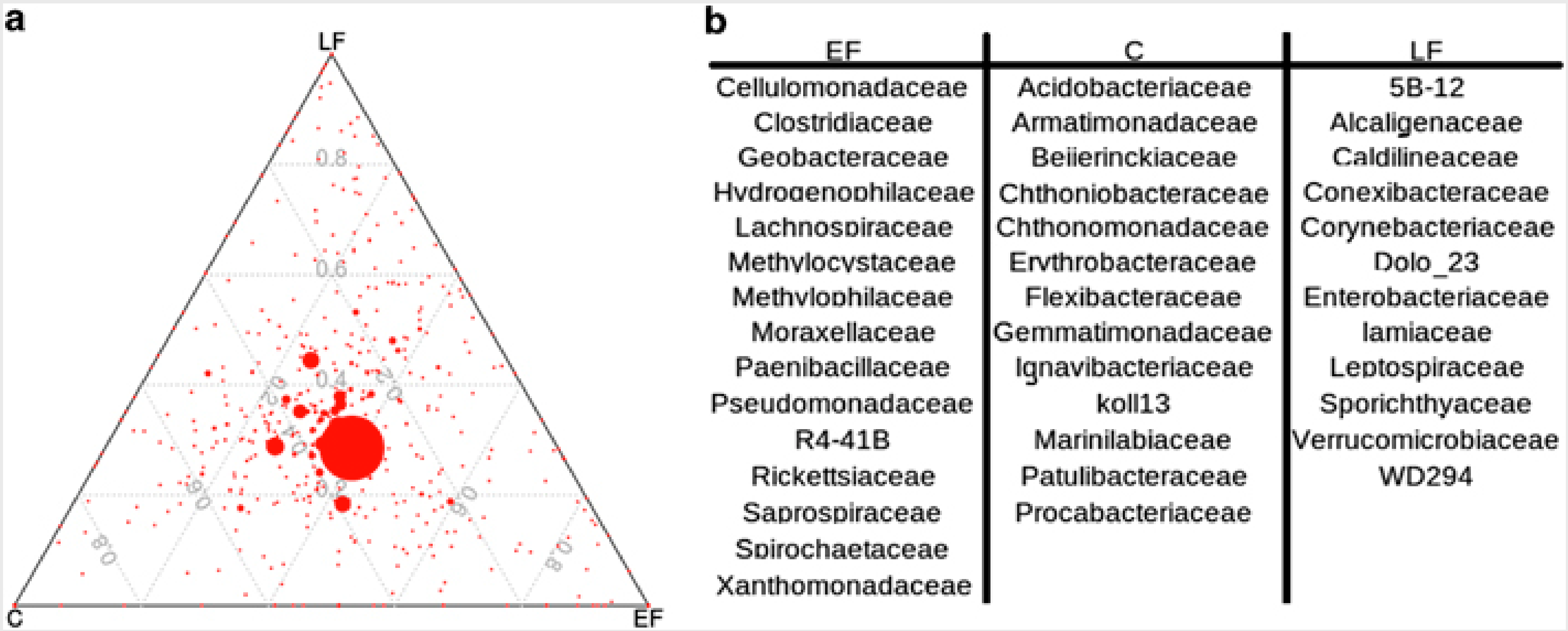
Cytoscape Network Maps

Mammal Gut Microbiomes[3]



Ternary Plots

R: Package ‘ggtern’

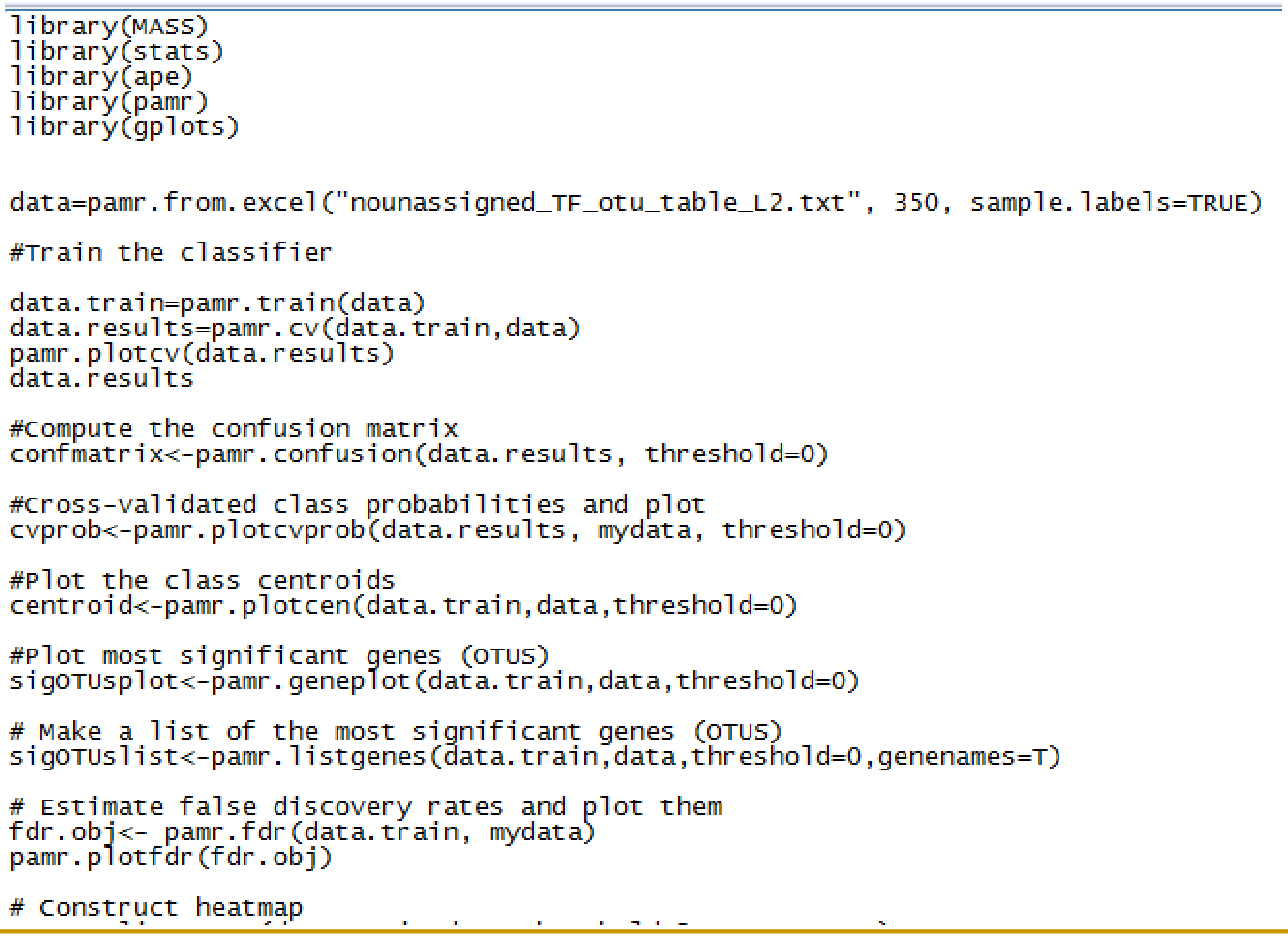


EF	C	LF
Cellulomonadaceae	Acidobacteriaceae	5B-12
Clostridiaceae	Armatimonadaceae	Alcaligenaceae
Geobacteraceae	Beijerinckiaceae	Caldilineaceae
Hydrogenophilaceae	Chthoniobacteraceae	Conexibacteraceae
Lachnospiraceae	Chthonomonadaceae	Corynebacteriaceae
Methylocystaceae	Erythrobacteraceae	Dolo_23
Methylophilaceae	Flexibacteraceae	Enterobacteriaceae
Moraxellaceae	Gemmatimonadaceae	Iamiaceae
Paenibacillaceae	Ignavibacteriaceae	Leptospiraceae
Pseudomonadaceae	koll13	Sporichthyaceae
R4-41B	Marinilabiaceae	Verrucomicrobiaceae
Rickettsiaceae	Patulibacteraceae	WD294
Saprospiraceae	Procabacteriaceae	
Spirochaetaceae		
Xanthomonadaceae		

Prediction Analysis of Microarrays for R (PAMR)

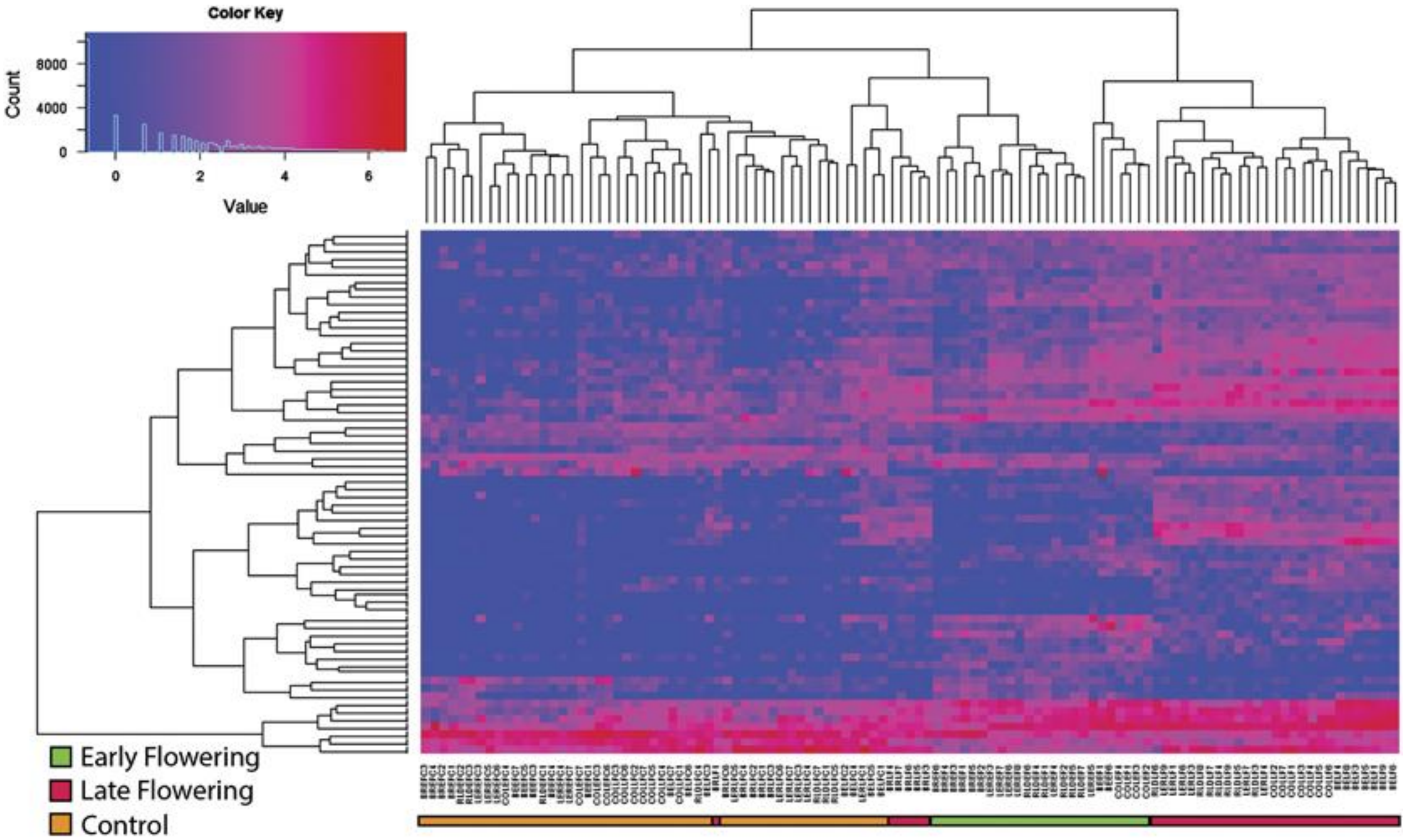
- Nearest Shrunken Centroid-based classification[7]
- Can be repurposed for use in community analysis
- Gene = OTU, Expression = RA, Class = Treatment
- Standardized centroid calculated for each class
- Class centroids “shrunken” toward centroid of overall data by threshold value
  - Reduces noise

R Code

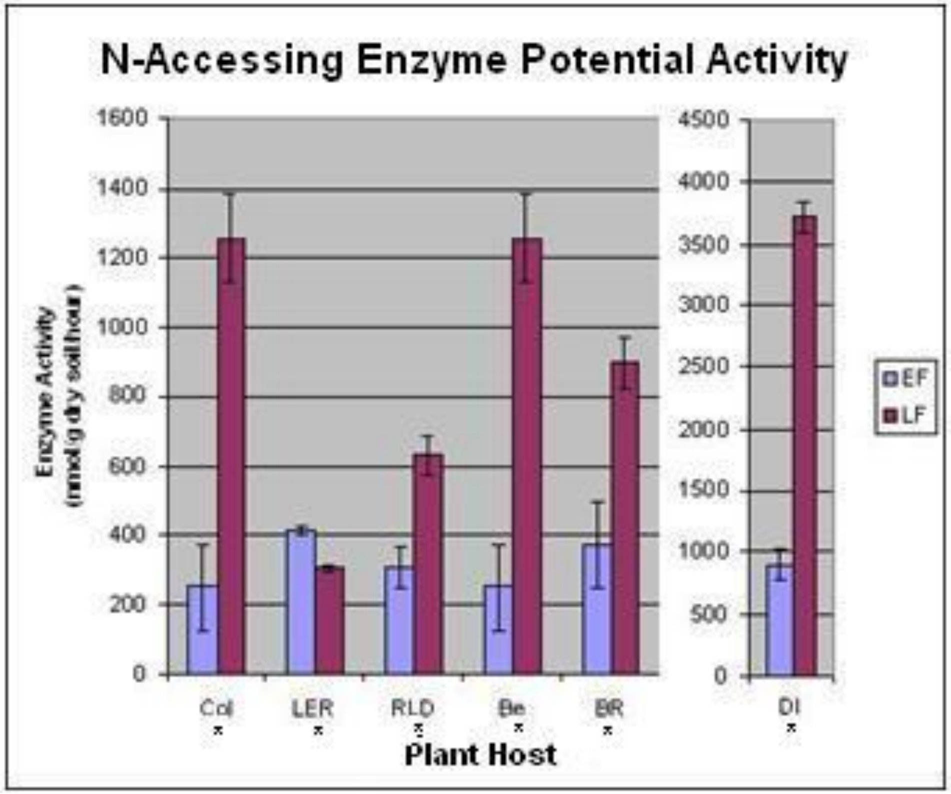


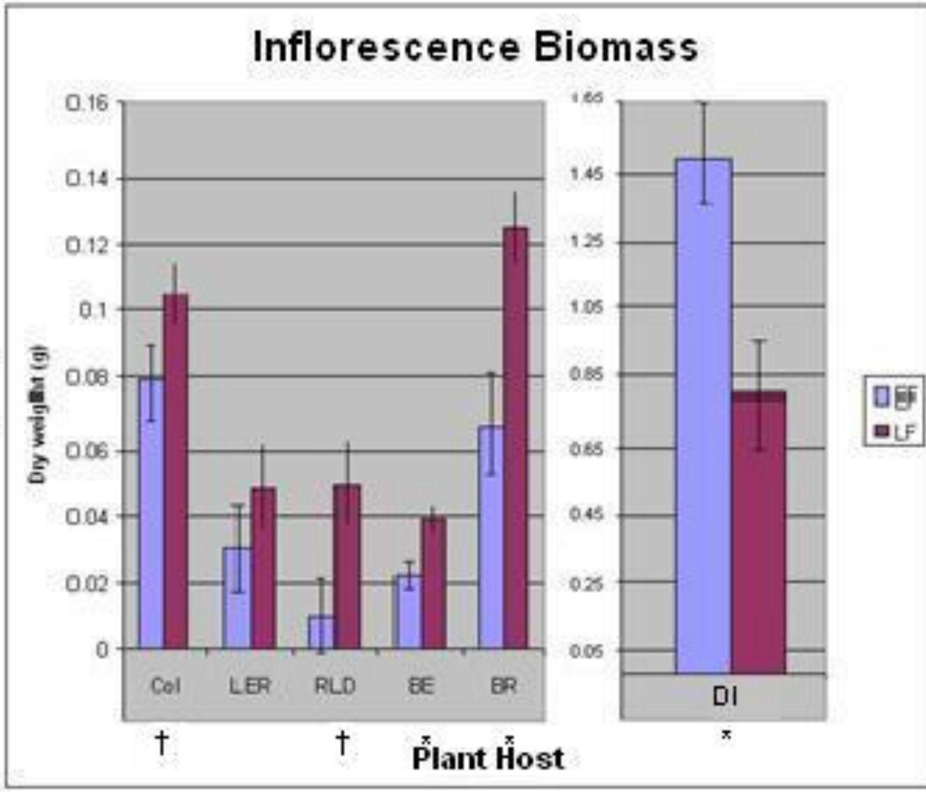
```
library(MASS)
library(stats)
library(ape)
library(pamr)
library(gplots)


data=pamr.from.excel("nounassigned_TF_otu_table_L2.txt", 350, sample.labels=TRUE)

#Train the classifier

data.train=pamr.train(data)
data.results=pamr.cv(data.train,data)
pamr.plotcv(data.results)
data.results

#Compute the confusion matrix
confmatrix<-pamr.confusion(data.results, threshold=0)

#Cross-validated class probabilities and plot
cvprob<-pamr.plotcvprob(data.results, mydata, threshold=0)

#Plot the class centroids
centroid<-pamr.plotcen(data.train,data,threshold=0)

#Plot most significant genes (OTUS)
sigOTUsplot<-pamr.geneplot(data.train,data,threshold=0)

# Make a list of the most significant genes (OTUS)
sigOTUslist<-pamr.listgenes(data.train,data,threshold=0,genenames=T)

# Estimate false discovery rates and plot them
fdr.obj<- pamr.fdr(data.train, mydata)
pamr.plotfdr(fdr.obj)

# Construct heatmap
```

PAMR Heatmap



LER

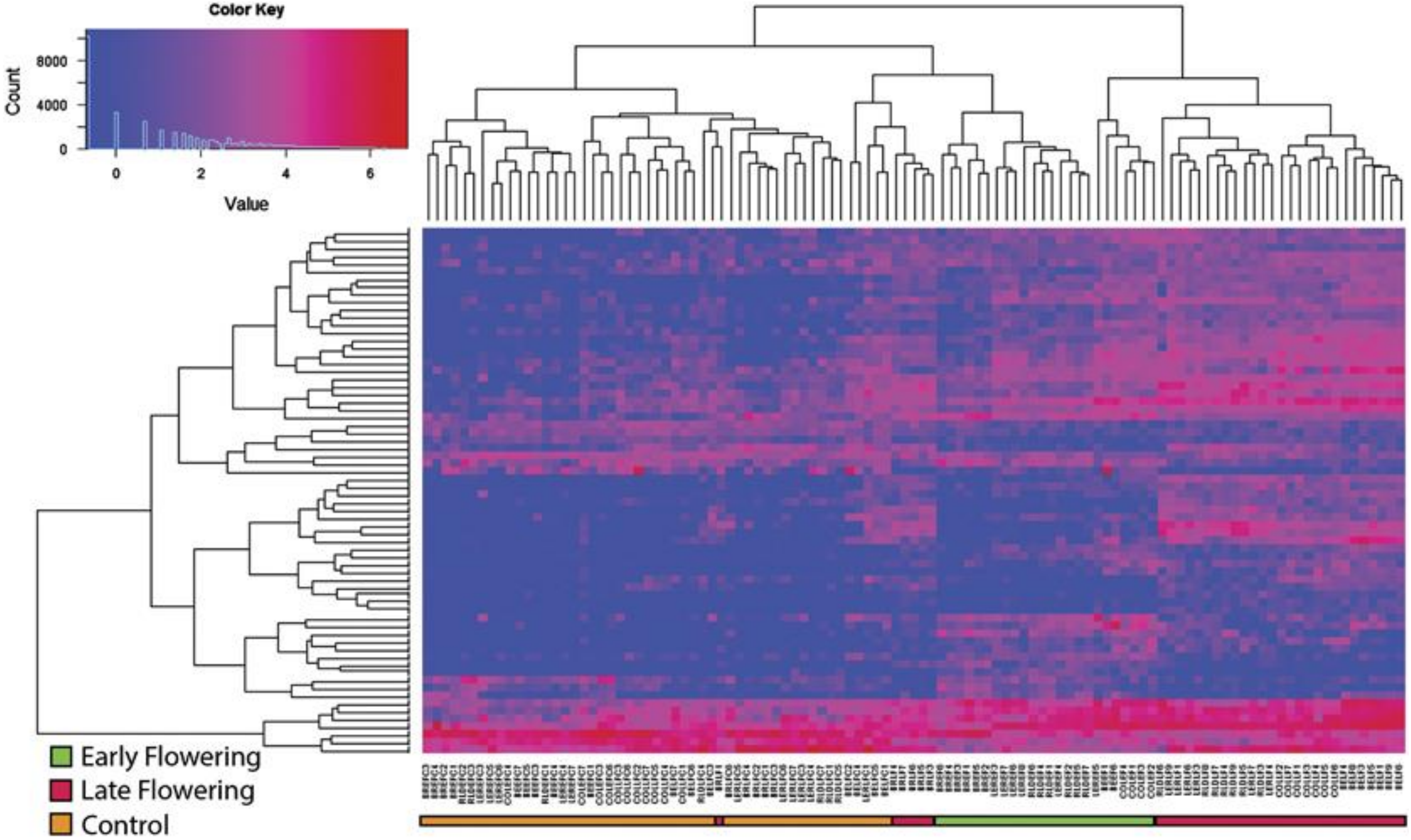

N-Accessing Enzyme Potential Activity


Inflorescence Biomass

Sequential Filtering

- Application of successive filters
- Identify what differs between LER and all other hosts
- Taxa switching

PAMR Heatmap



Thank you for your time.

References

1) Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. (2010). QIIME allows analysis of high-throughput community sequencing data. Nat. Methods 7:335–336.Cline MS, Smoot M, Cerami E, Kuchinsky A, Landys N, Workman C, et al. (2007). Integration of biological networks and gene expression data using Cytoscape. Nat. Protoc. 2:2366–2382.
2) Guerrero R, Margulis L, Berlanga M. (2013). Symbiogenesis: The holobiont as a unit of evolution. Int. Microbiol. 16:133–143.
3) Ley RE, Hamady M, Lozupone C, Turnbaugh PJ, Ramey RR, Bircher JS, et al. (2008). Evolution of mammals and their gut microbes. Science 320:1647–1651.
4) Panke-Buisse K, Poole AC, Goodrich JK, Ley RE, Kao-Kniffin J. (2015). Selection on Soil Microbiomes Reveals Reproducible Impacts on Plant Function. Isme J.
5) Saito R, Smoot ME, Ono K, Ruscheinski J, Wang P, Lotia S, et al. (2012). A travel guide to Cytoscape plugins. Nat. Methods 9:1069–76. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3649846&tool=pmcentrez&rendertype=abstract.
6) Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T. (2011). Cytoscape 2.8: New features for data integration and network visualization. Bioinformatics 27:431–432.
7) Tibshirani R, Hastie T, Narasimhan B, Chu G. (2002). Diagnosis of multiple cancer types by shrunken centroids of gene expression. Proc. Natl. Acad. Sci. U. S. A. 99:6567–6572.
8) Yeung N, Cline MS, Kuchinsky A, Smoot ME, Bader GD. (2008). Exploring biological networks with cytoscape software. Curr. Protoc. Bioinforma.
9) Swenson W, Wilson DS, Elias R. (2000). Artificial ecosystem selection. Proc. Natl. Acad. Sci. U. S. A. 97:9110–9114. <Go to ISI>://WOS:000088608000064.