

<https://eval.ucdavis.edu>

Please take 10 minutes to fill out the online course evaluation. Professor Maloof will be back shortly.

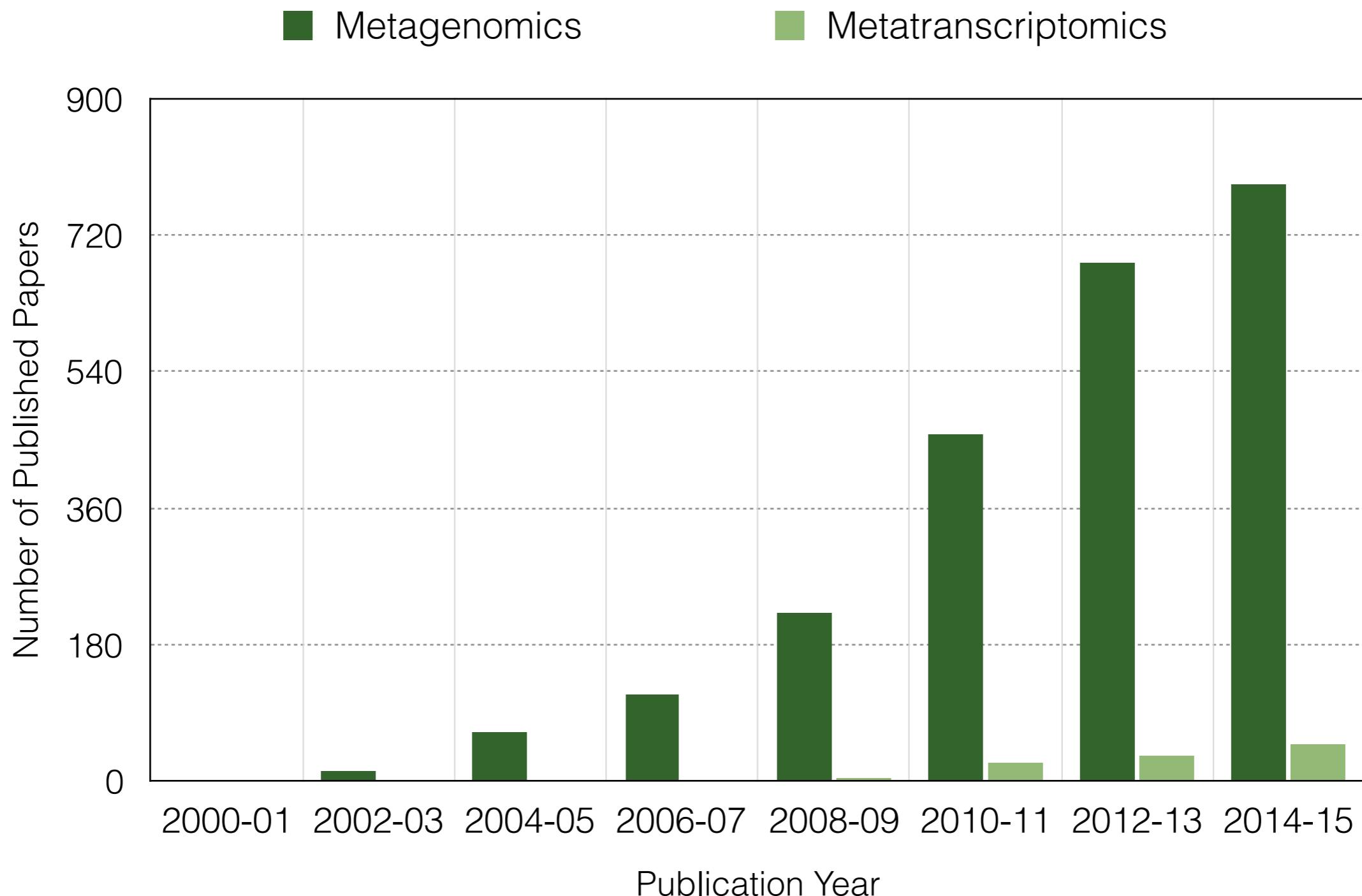
# Getting Messy with Microbes: *An Introduction to Metagenomics*

Slides from Kristen Beck, modified by Julin Maloof

# Metagenomics.

The study of a collection of genetic material from a mixed community of organisms, typically microbial.

# Metagenomics is a rapidly expanding field





Metagenomics is a lens to observe  
countless environments.





Zoological and evolutionary insights

Today's lab focuses on the rice root microbiome



Photo credit: Sundaresan Lab, UC Davis



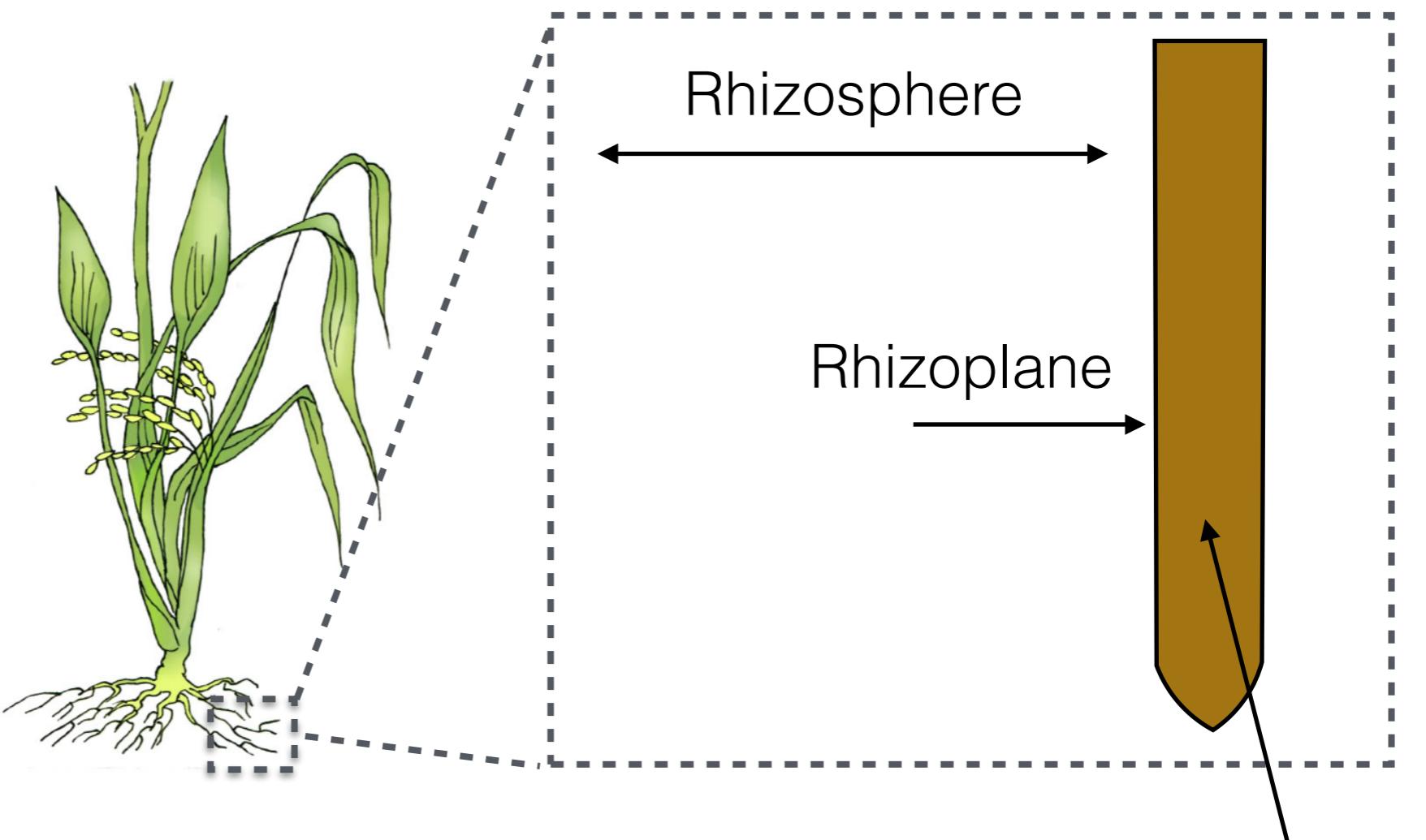
Nipponbare



M104



IR50



Endosphere

# How do we assess microbial community composition?

Sequence the 16S ribosomal RNA

# Sequencing of the 16S ribosomal RNA

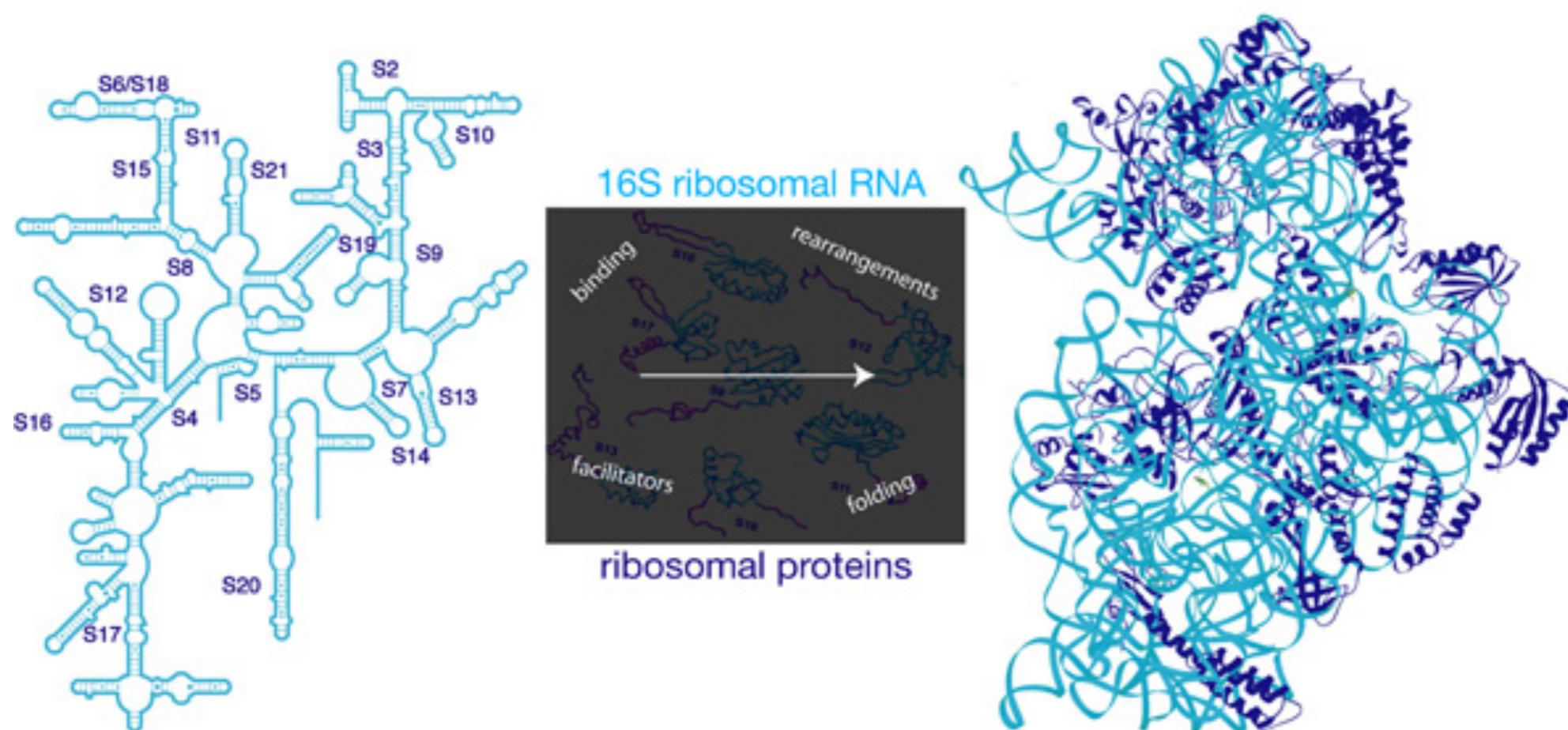


Photo credit: Culver Lab, Univ. of Rochester

What are the components of  
a metagenomics analysis?

# Operational Taxonomic Unit (OTU)



A cluster of 16S rRNA reads with >97% similarity to define “species”

# Questions that can be answered by 16S sequencing:

- Which microbes are in my sample? What might they be doing?
- Is microbe X present at different levels between samples (differential abundance)?
- Why might one phylum be more abundant in this environment?

# Alpha Diversity

- Alpha diversity: how diverse is the microbial community?
- Can compare diversity between samples
  - e.g. is the microbial community more diverse at the root surface or in bulk soil?

# Alpha Diversity Method 1

- **Count OTUs**
- Simple, but...any issues?

# Alpha Diversity Method 2

- **Shannon's Diversity**
- $s$  = number of species
- $p_i$  = proportion of counts attributable to species i
- equal representation leads to high diversity index  
(play with it in R)

$$H = - \sum_{i=1}^s (p_i \log_2 p_i)$$

# Alpha Diversity Method 3

- **Chao1**

$$chao1 = S_{obs} + \frac{F_1^2}{2F_2}$$

- $S_{obs}$  = total observed species
- $F_1$  = number of singletons (species observed only once in a sample)
- $F_2$  = number of doubletons

# Questions that can be addressed with alpha diversity:

How many taxa are in a sample? What is the richness of my sample?

Have I sequenced to a depth (coverage) that describes the diversity of my sample? (not in this lab...)

Does condition X have higher phylogenetic diversity than condition Y?

# Beta Diversity

- Beta diversity measures the diversity *between* samples.
- The distance between each pair of samples (with respect to community composition) is calculated.
- There are many ways to calculate beta diversity; we will use one: Bray-Curtis

# Bray-Curtis Dissimilarity

- For two samples,  $i$  and  $j$ :

unique species counts

total counts

- $S_i$  and  $S_j$  = sum of counts in samples  $i$  and  $j$
- $C_{ij}$  = sum of min(counts) for taxa observed in both  $i$  and  $j$

$$BC_{ij} = \frac{S_i + S_j - 2 * C_{ij}}{S_i + S_j}$$

Do samples contain significantly different microbial communities?

Which taxa have increased abundance compared to another sample?

Are there broad trends that relate many samples? Can these trends be explained by an environmental factor?