

Schedule

Day 1: (OPTIONAL) Unix Basics and Intro to Programming -- Emily Davenport

Day 2: Basics of Microbiome Analysis - From Planning to Sequencing -- Organizing Committee

Day 3: Amplicon Data Analysis in R (Sharifa Crandall and Erika Ganda)

Day 4: Amplicon Data Analysis in QIIME2 (Erika Ganda and Laura Weyrich)

QIIME2 16S rRNA Tutorials:

<https://docs.qiime2.org/2021.4/tutorials/moving-pictures/>

<https://docs.qiime2.org/2020.6/tutorials/pd-mice/>

Daily MC Video: https://psu.mediaspace.kaltura.com/media/Educational+and+Social+Opportunities+in+the+Microbiome+Center/1_5g6rxymg

Day 5: Shotgun data analysis (binning and assembly) (David Koslicki)

TA Introductions

Chad Fautt,
Erica Ryu,
Emily Van Syoc

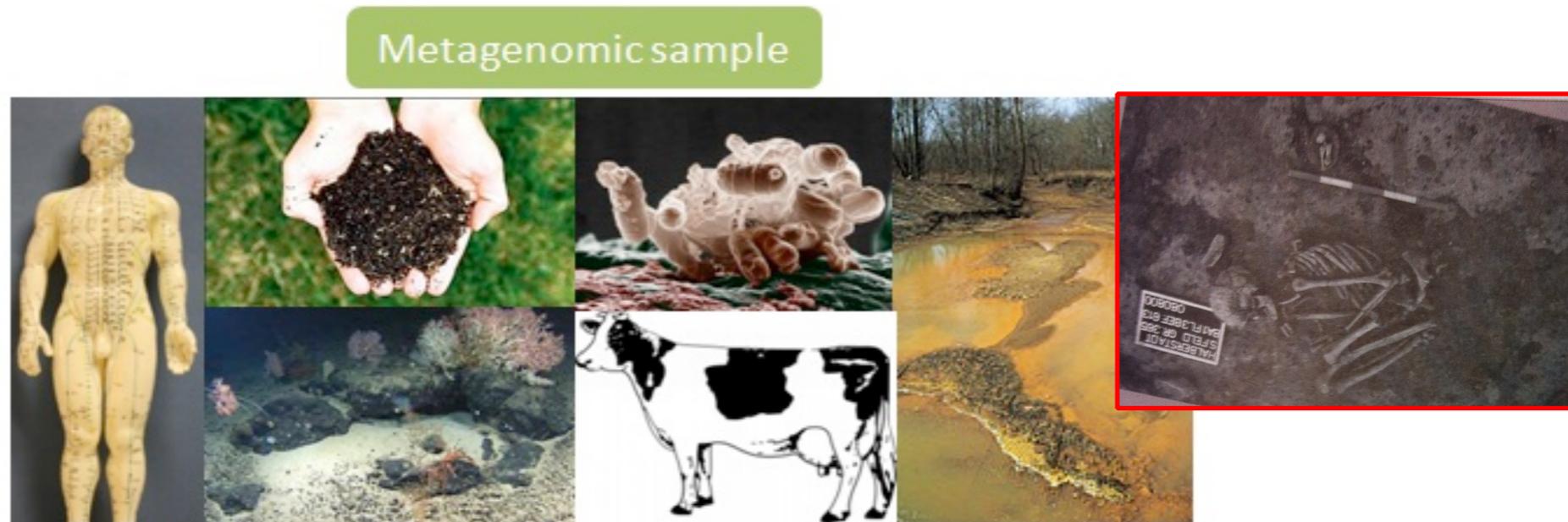
AMPLICON DATA ANALYSIS IN QIIME2



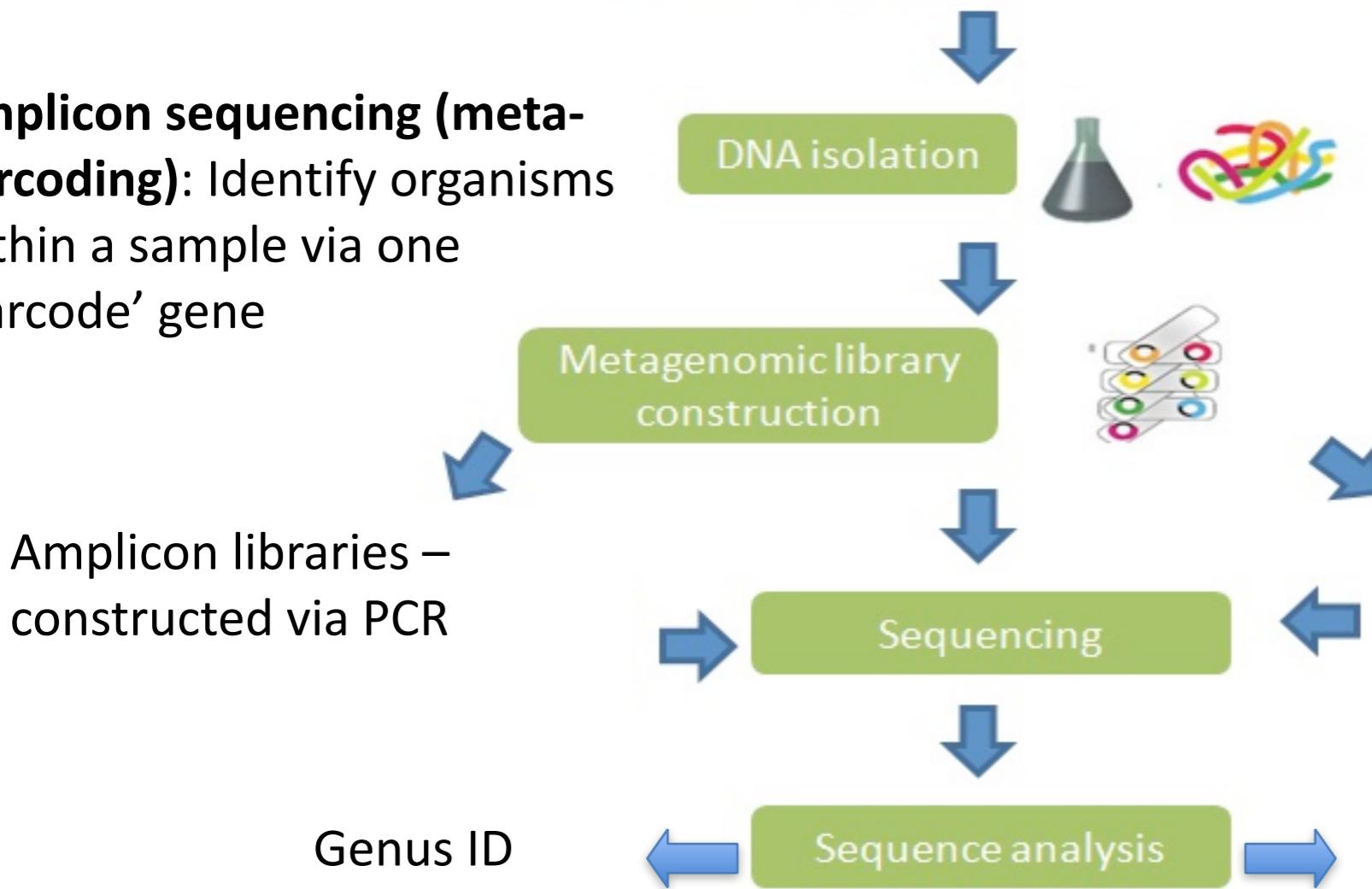
LAURA S. WEYRICH
PENNSYLVANIA STATE UNIVERSITY, USA
UNIVERSITY OF ADELAIDE, AUSTRALIA

@LSWEYRICH
LSW132@PSU.EDU

Microbiome NGS analysis

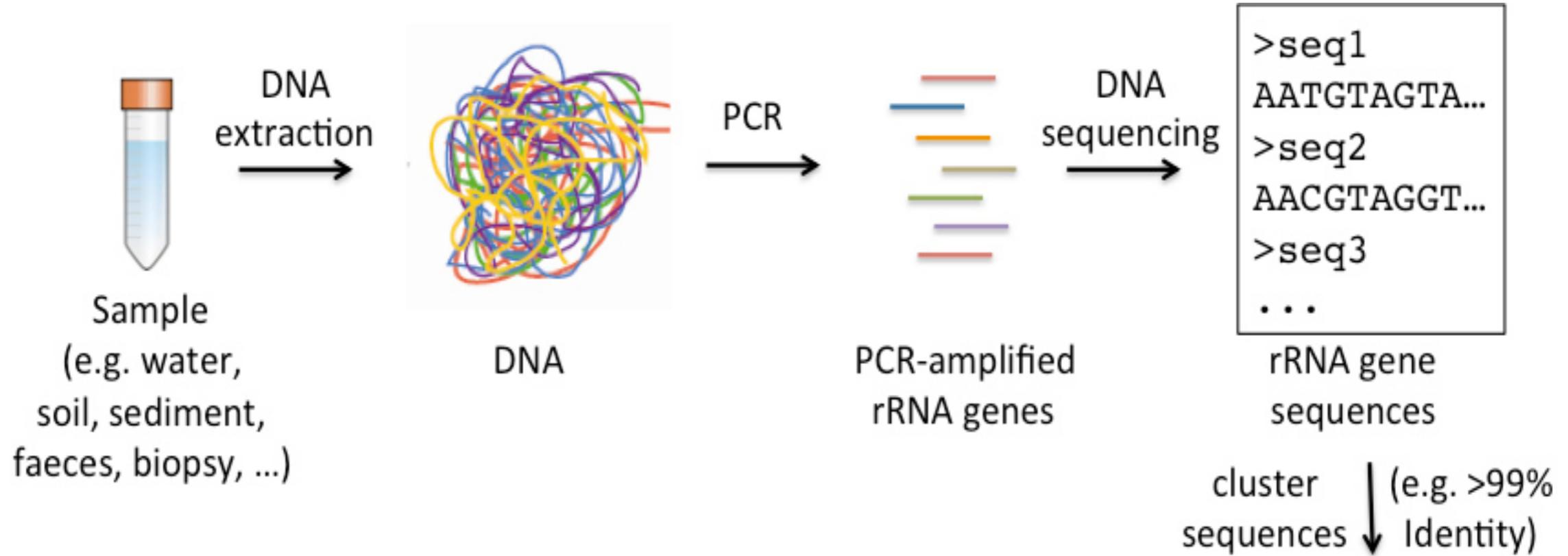


Amplicon sequencing (meta-barcoding): Identify organisms within a sample via one 'barcode' gene



Amplicon Based Sequencing

Wet Lab

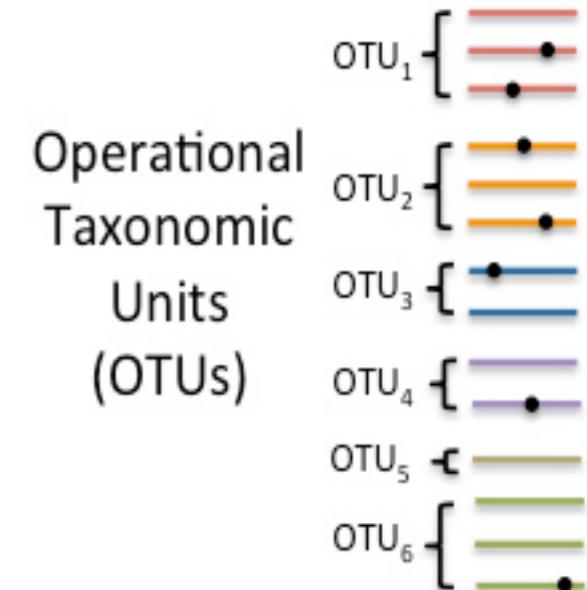


Bioinformatics

OTU	Species	Sample1	Sample2	Sample3
1	E.coli	17	0	335
2	S.aurus	231	11800	45
3	unknown	30	0	0
...

Counts of OTUs
per sample

BLAST-search
rRNA sequence
database
←
with millions
of taxonomically
classified
rRNA sequences
(e.g. RDP, Silva)

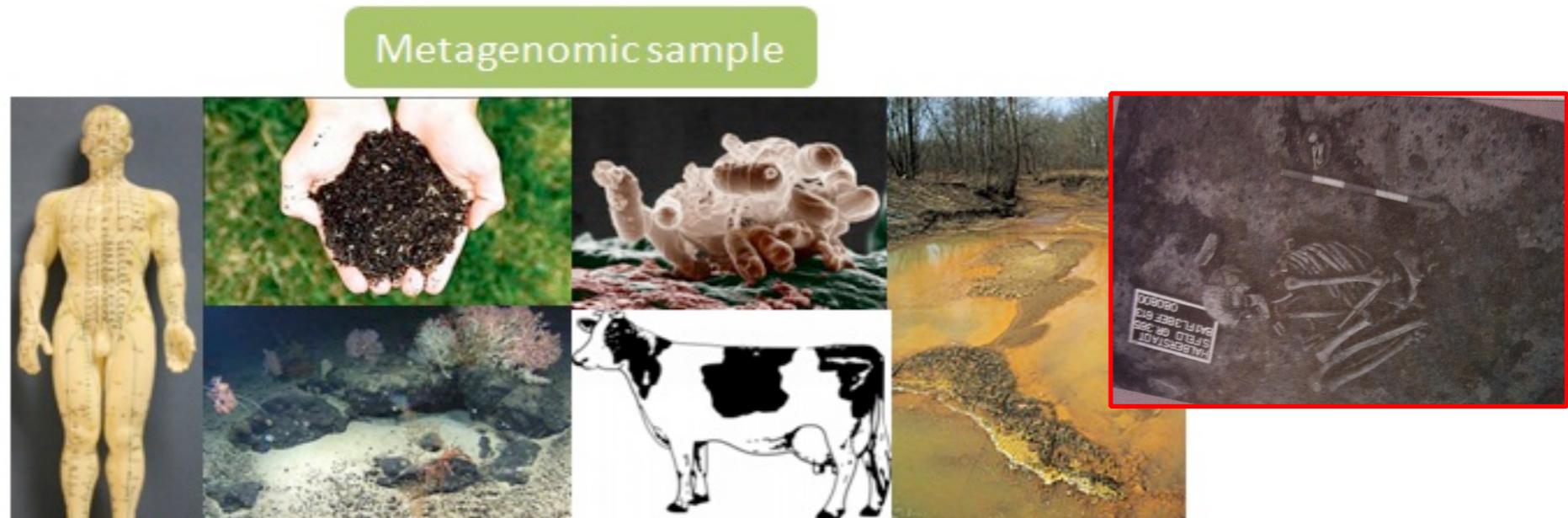


Issues with amplicon analysis

- Sample Preservation (DNA size)
- Laboratory, reagents, and technician contamination can significantly alter findings.
- Biases from amplification.
- Species IDs are difficult.
- Limited information during bioinformatic analysis.



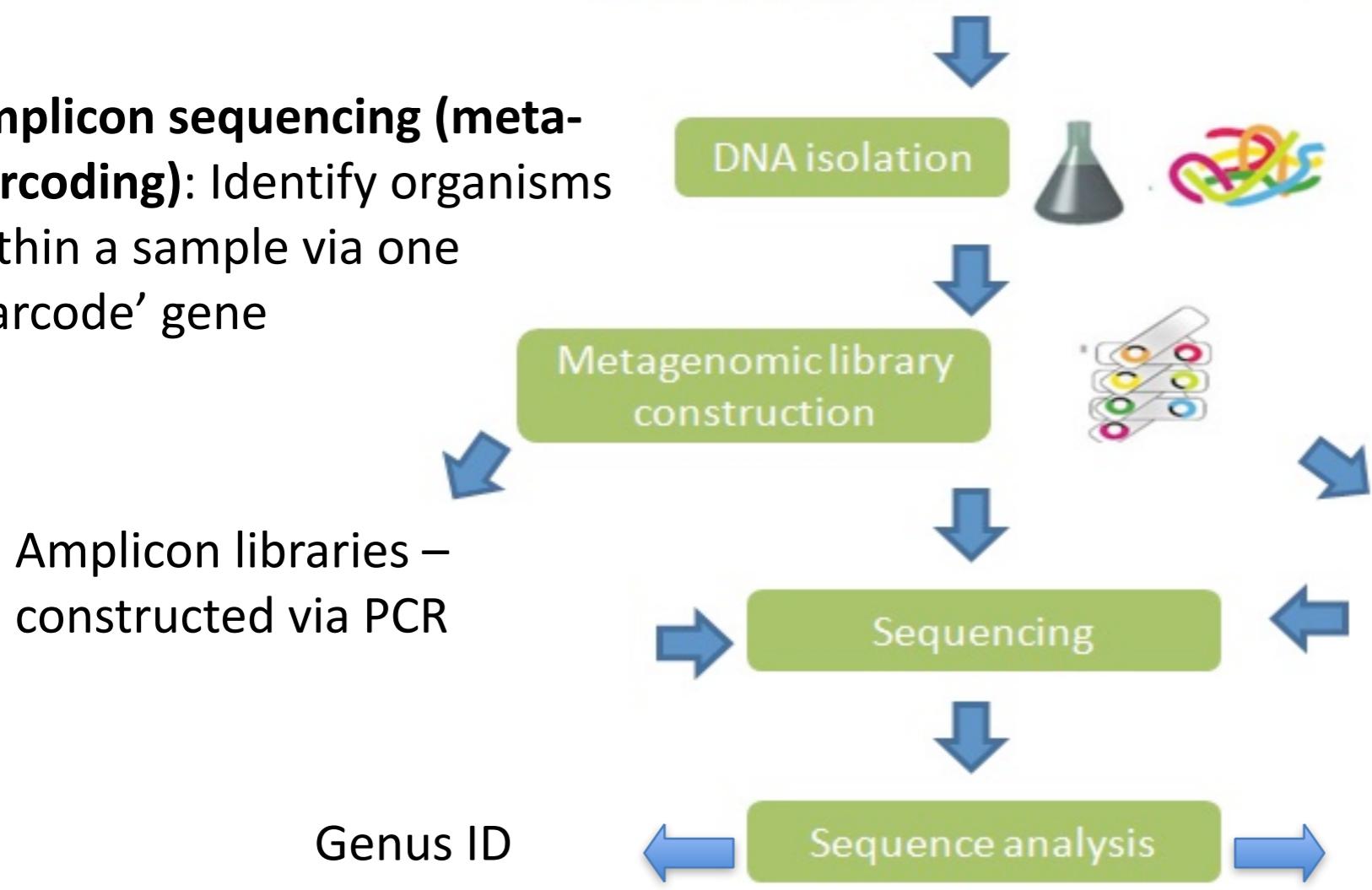
Microbiota/Microbiome Analysis: NGS



Amplicon sequencing (meta-barcoding): Identify organisms within a sample via one 'barcode' gene

Amplicon libraries – constructed via PCR

Genus ID



Shotgun sequencing: random sampling of all the DNA present

Shotgun NGS libraries – constructed via ligation of adapters

Species ID and Function

Pros/Cons of NGS Approaches

	Amplicon Sequencing	Shotgun Sequencing
Pros	Cheap Rapid Technically Easy Less Complex	Functional Information Less amplification Bias Species/Strain Level ID Genome Assembly
Cons	No Functional Data PCR Biases ID limited to genera	More expensive Labor intensive Limited Analysis Programs Need Deep Coverage?

Questions?

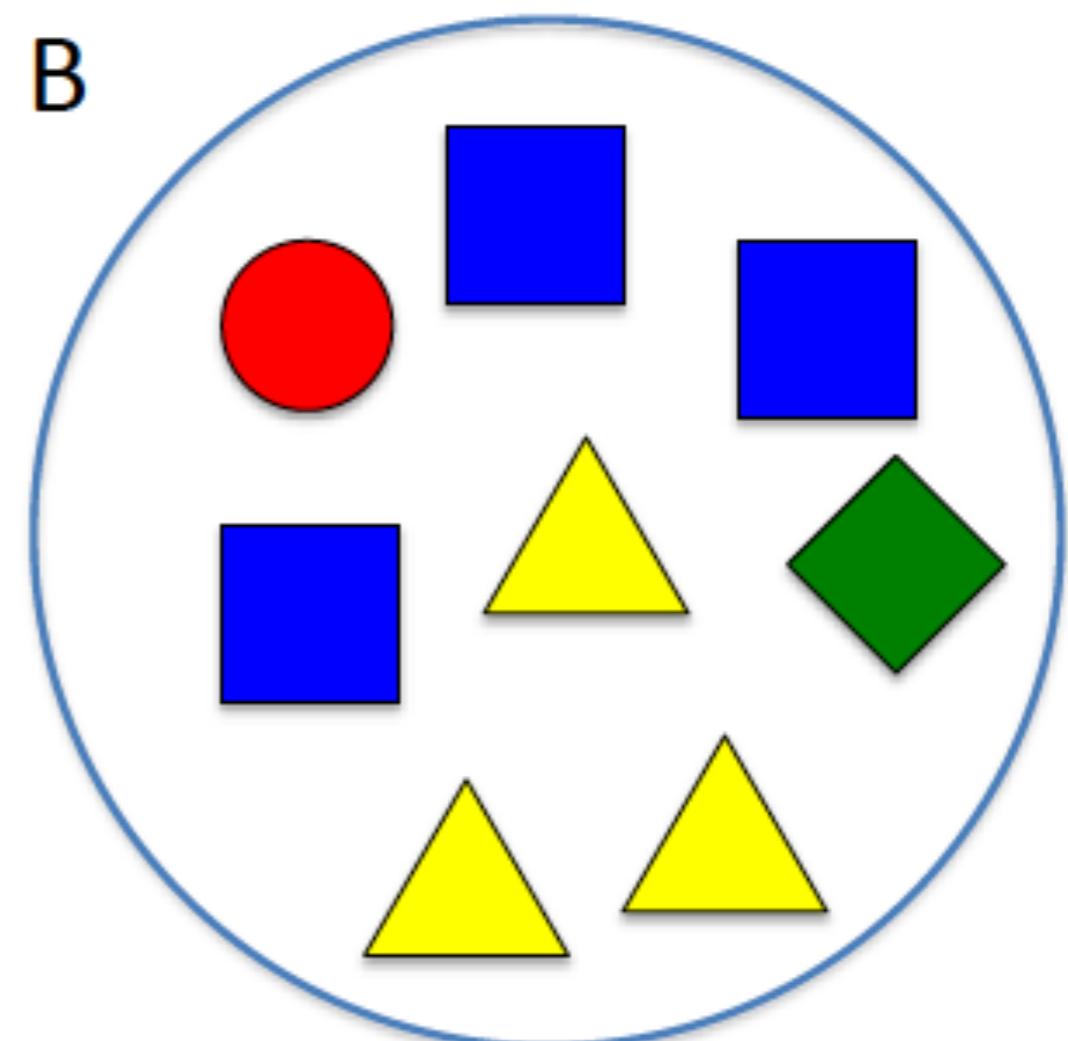
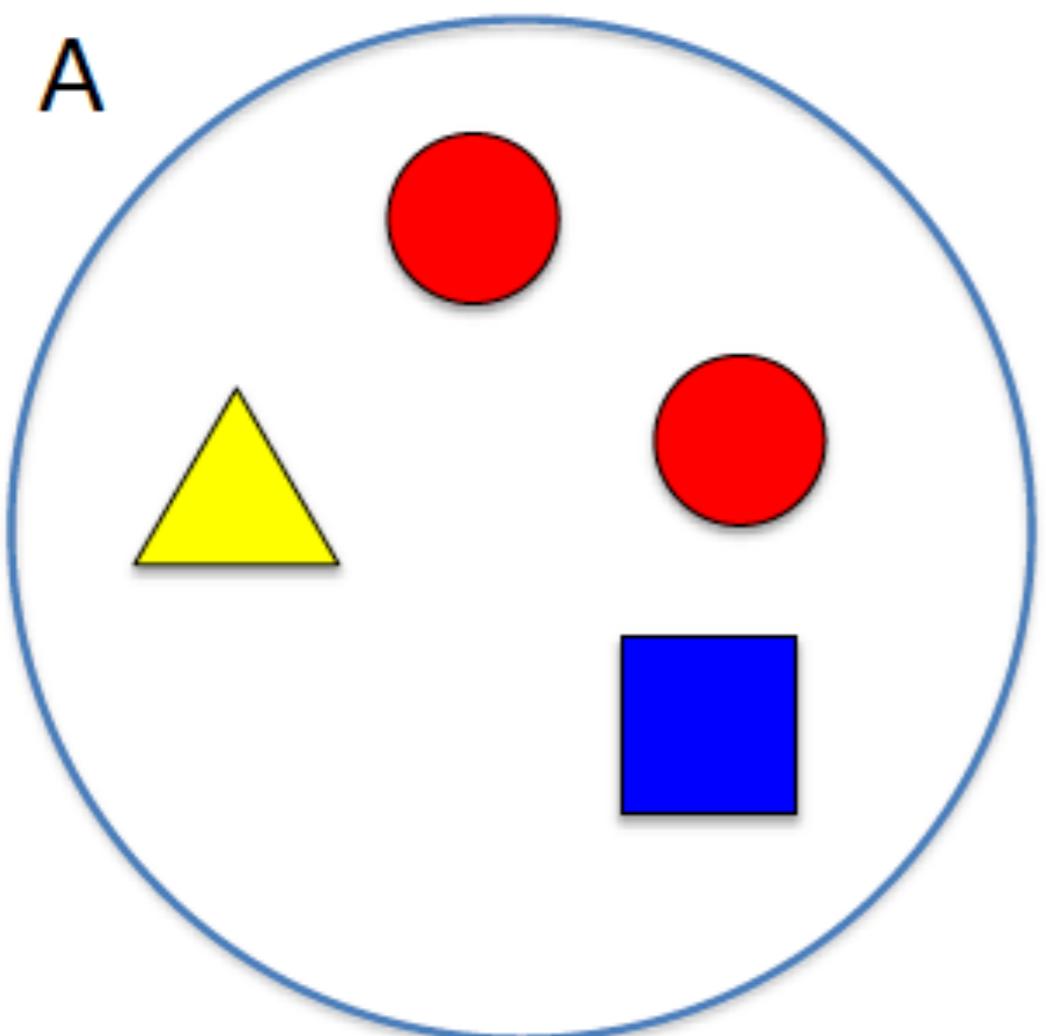
**What types of analyses
would we want to do with
amplicon data?**

Analyses

- Data cleaning
- Taxonomic Alpha diversity
- Taxonomic Beta diversity
- Taxa (ASV) Identification
 - Overview of Species Available
 - Linkages between metadata and specific ASVs

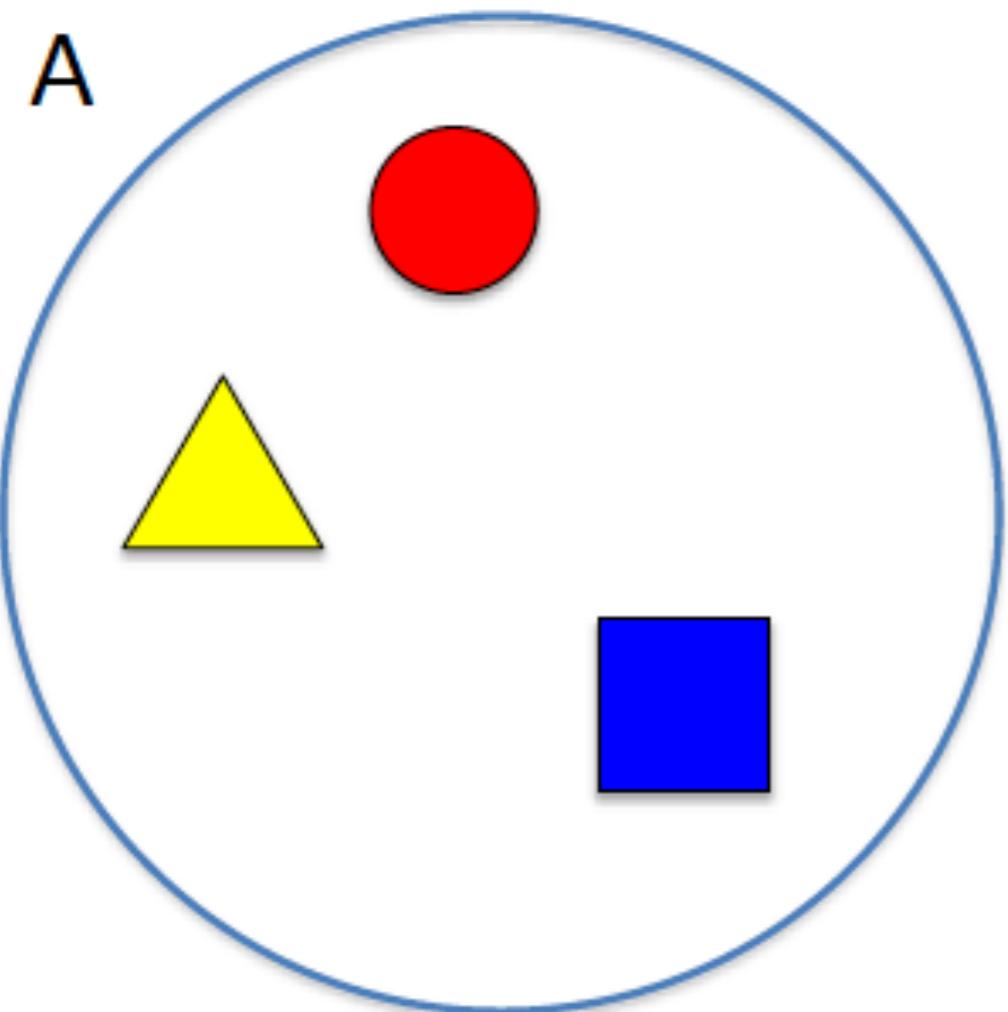
Alpha diversity

Alpha diversity

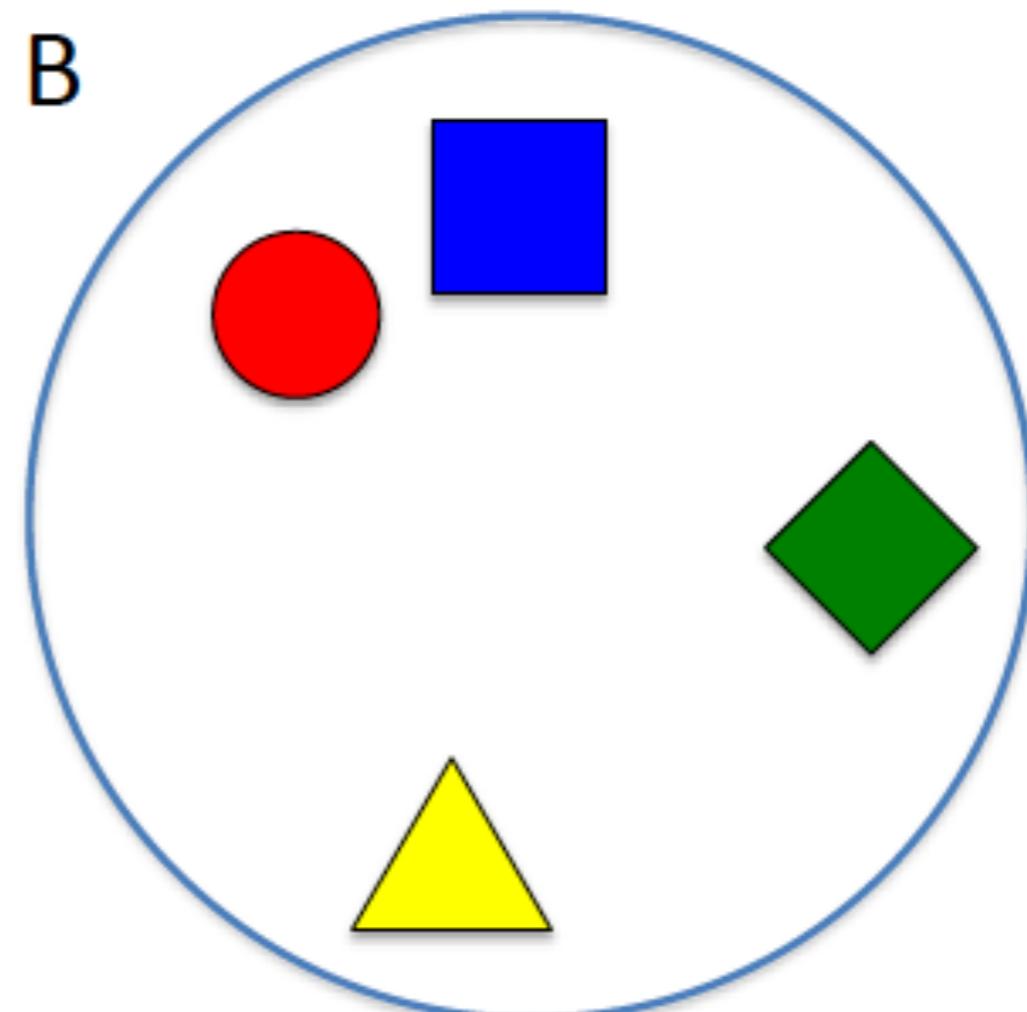


Ref: Daniel McDonald; American Gut Project

Alpha diversity



Alpha diversity = 3

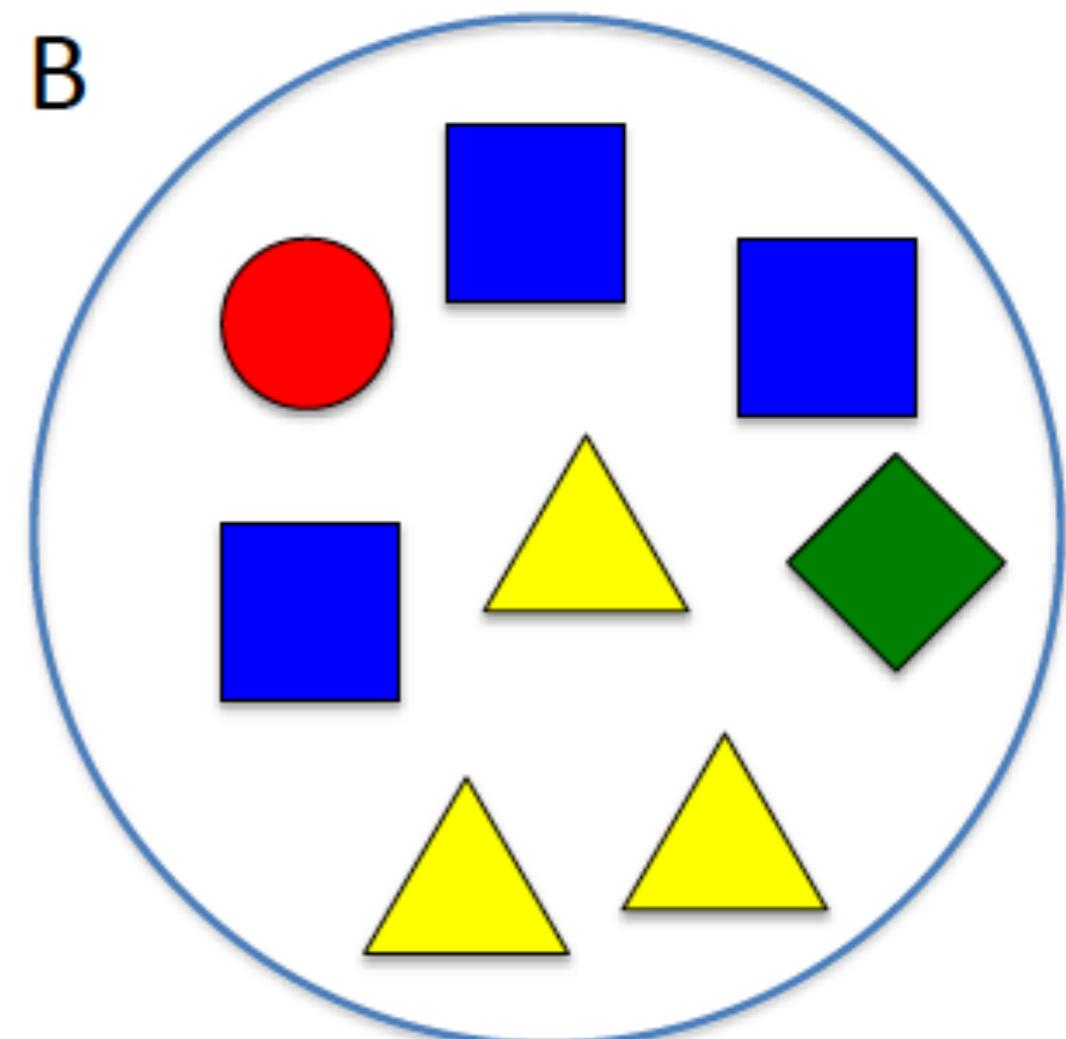
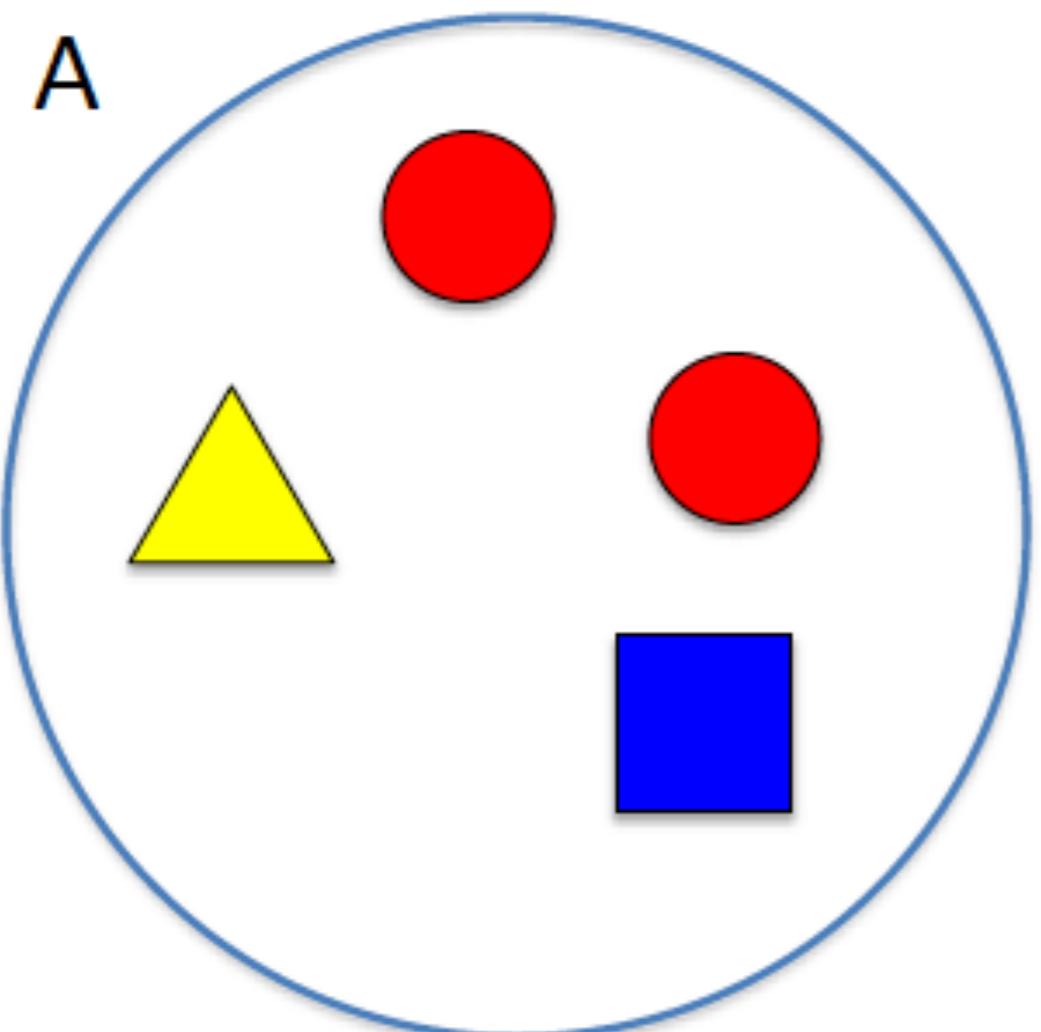


Alpha diversity = 4

Sample B is more diverse than sample A

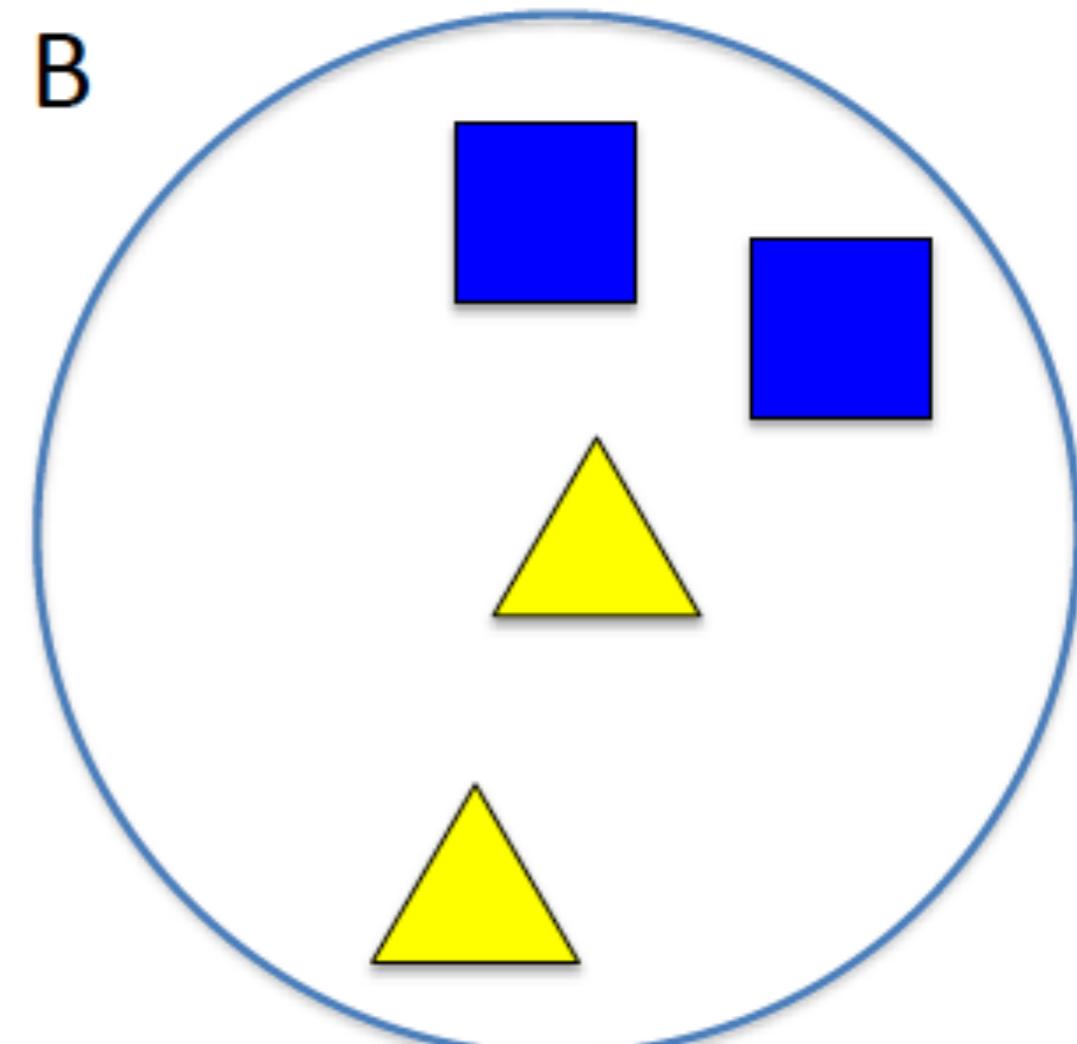
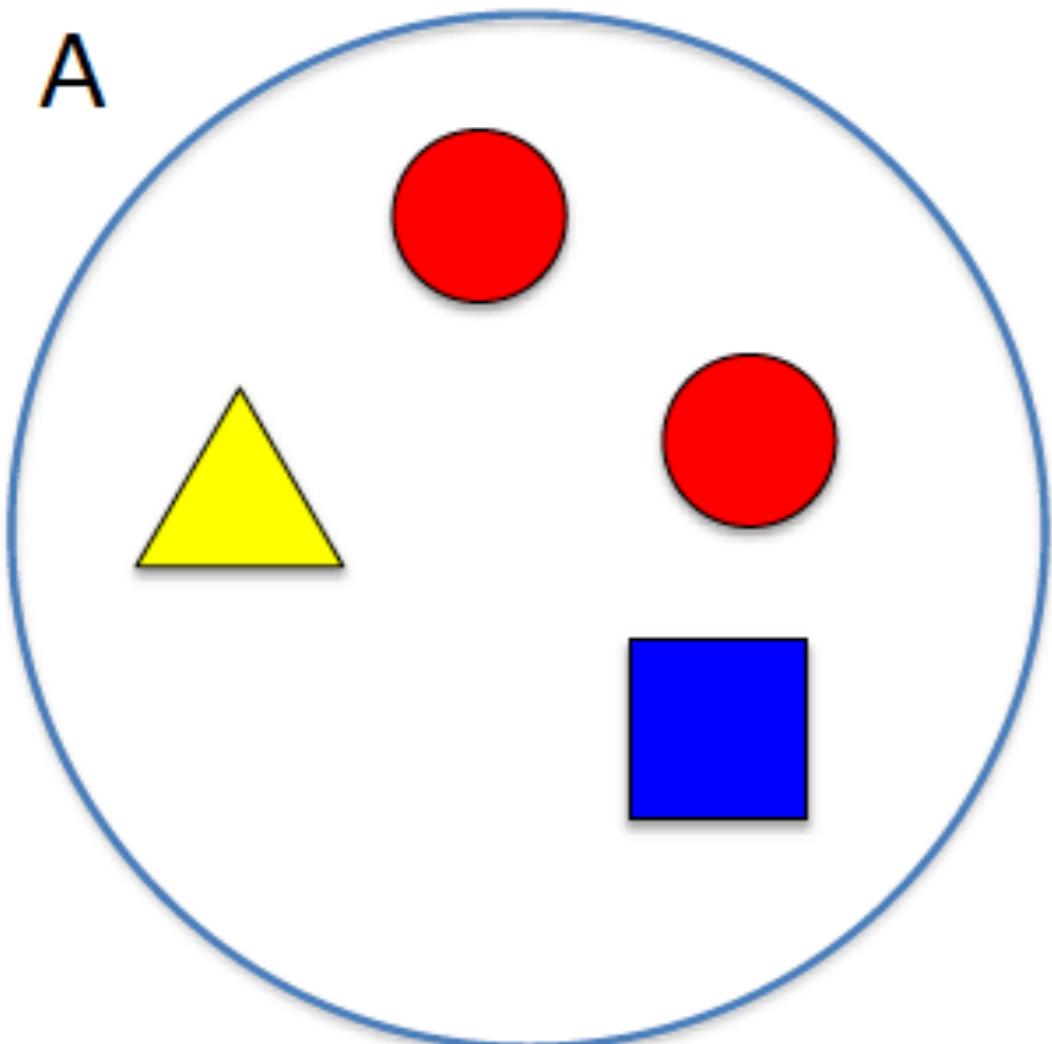
Ref: Daniel McDonald; American Gut Project

Alpha diversity



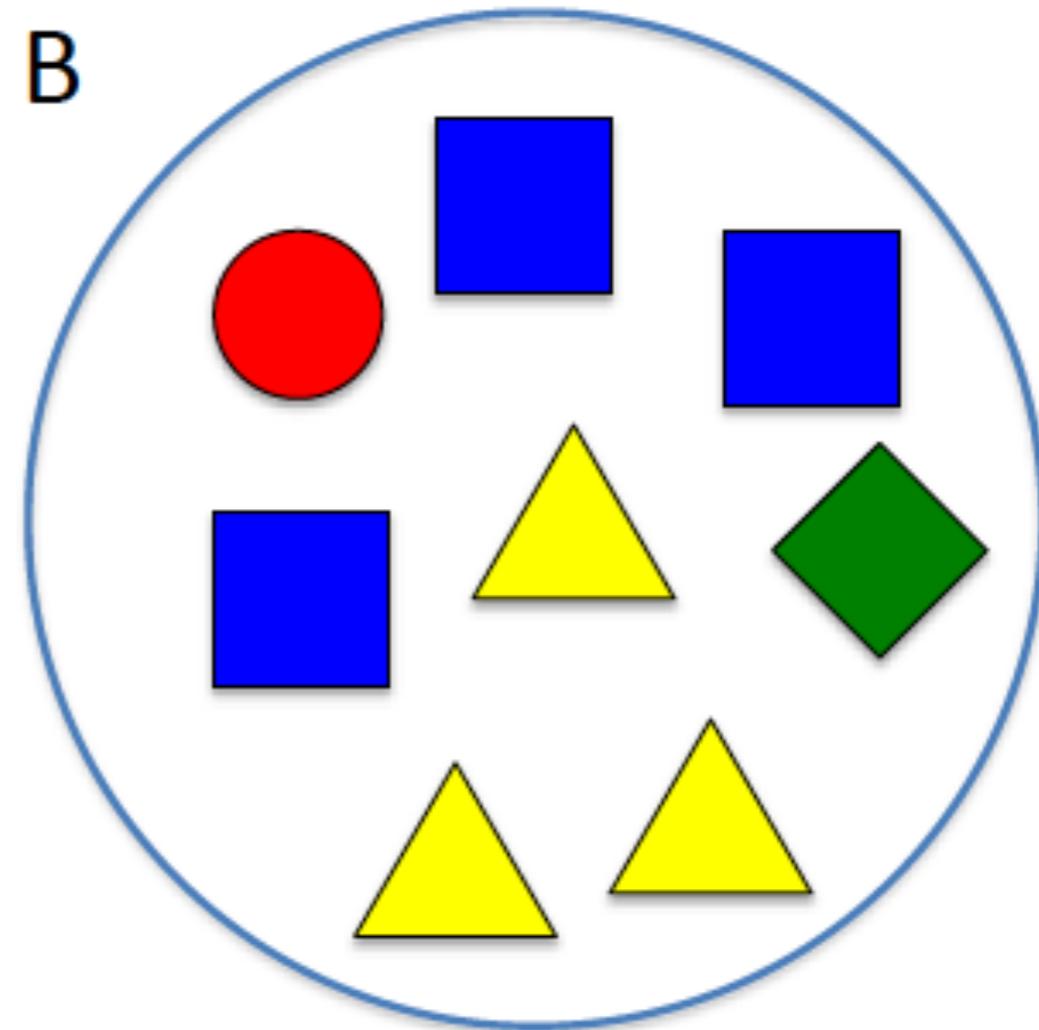
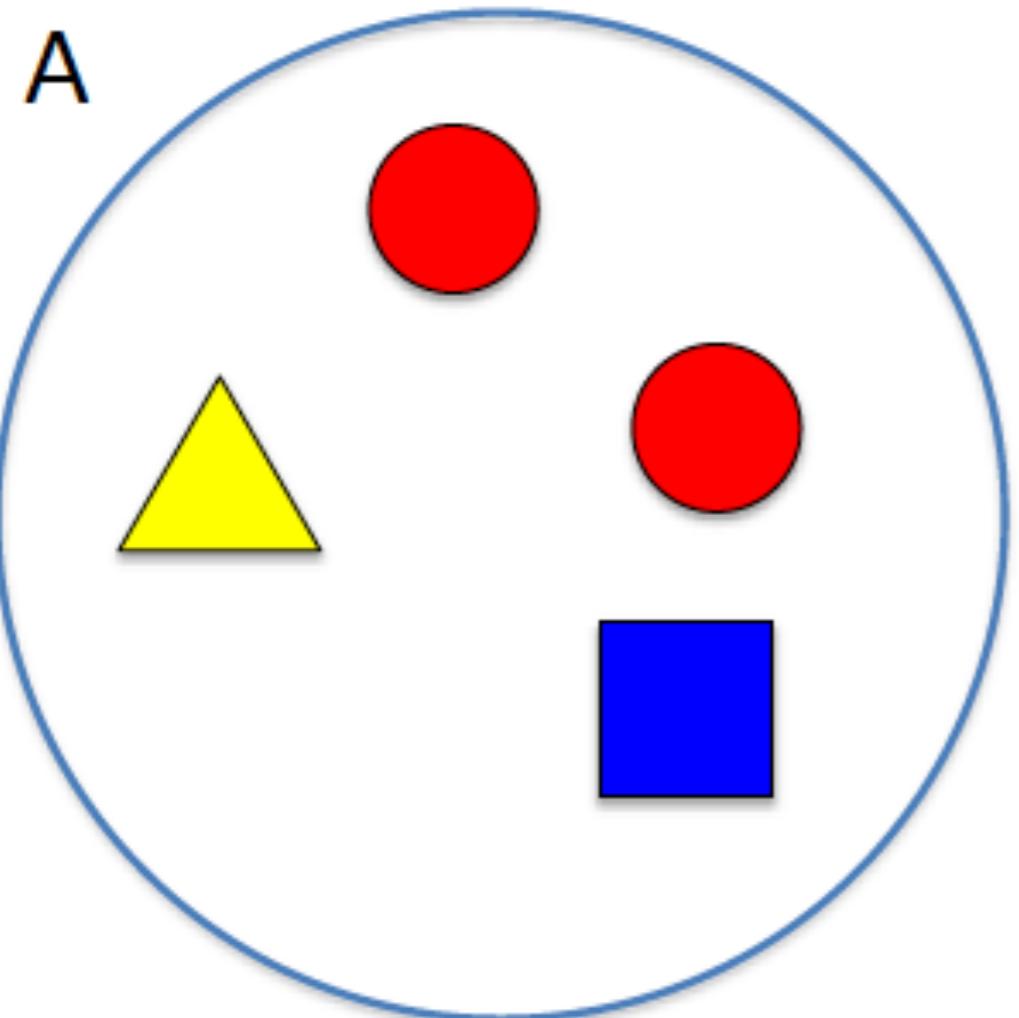
Ref: Daniel McDonald; American Gut Project

Rarefaction



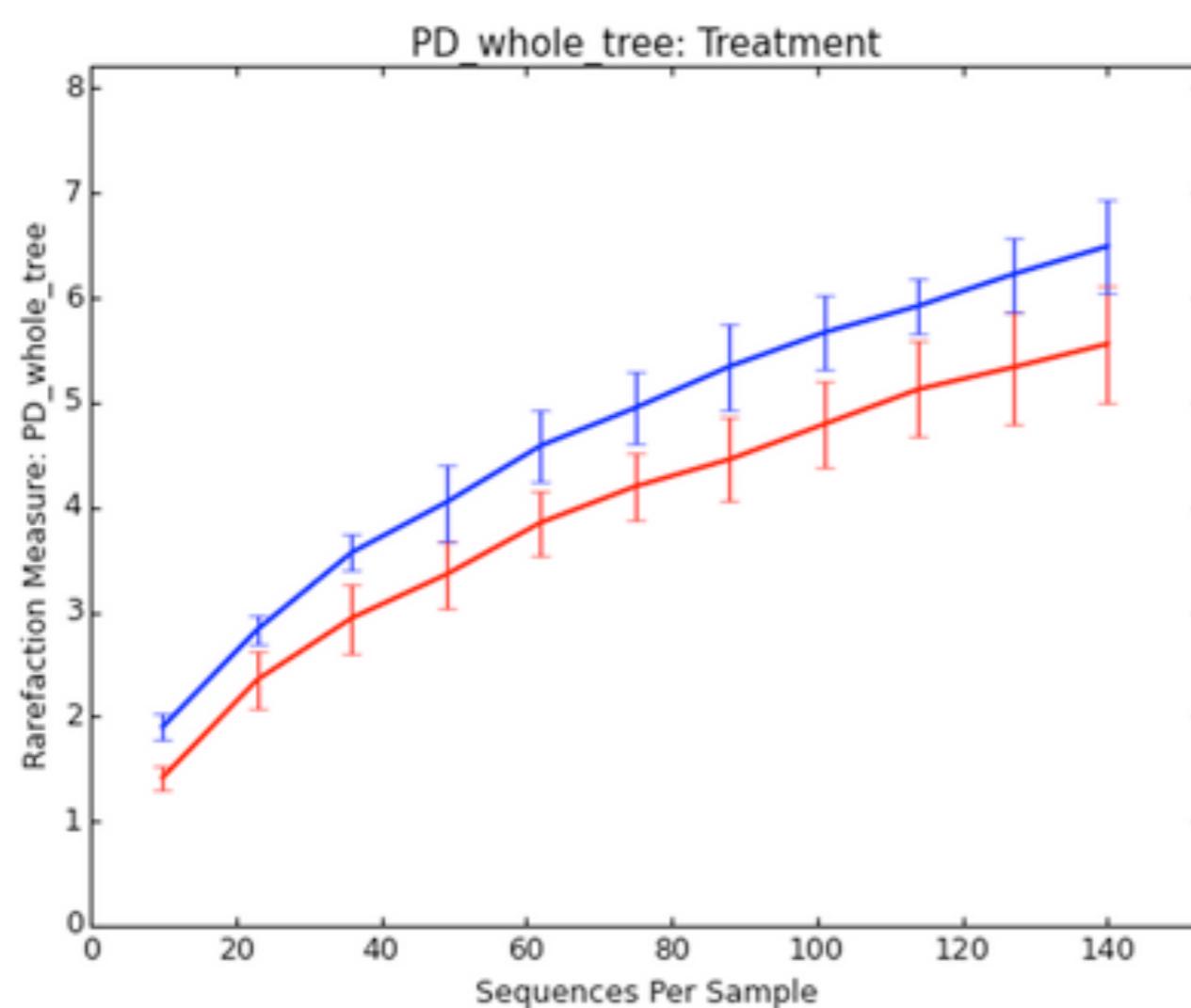
Ref: Daniel McDonald; American Gut Project

Rarefaction



Rarefy to 4 sequences

Select a Metric: PD_whole_tree ▾ Select a Category: Treatment ▾



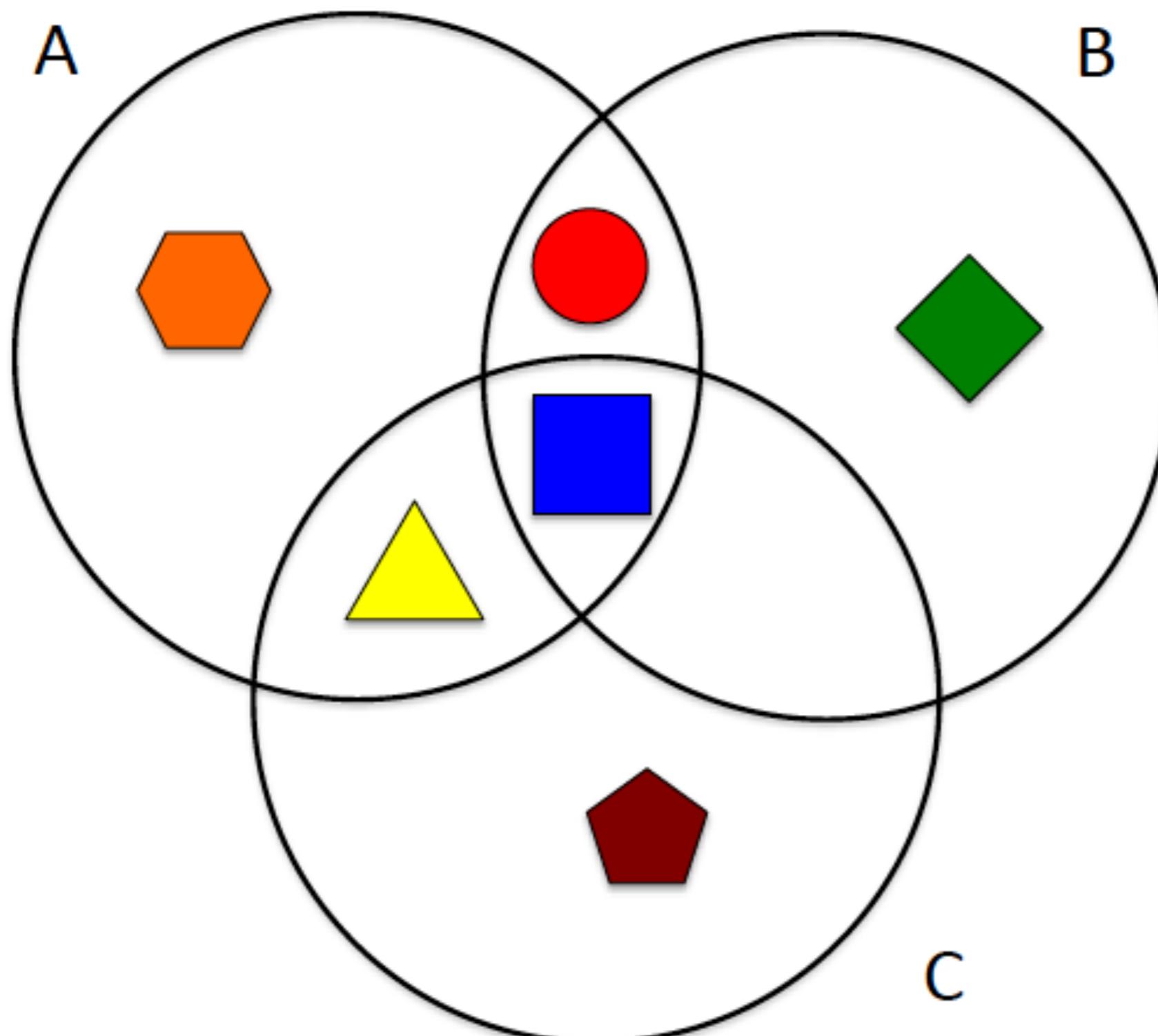
Show Categories: ▾

Legend

- Control
- Fast

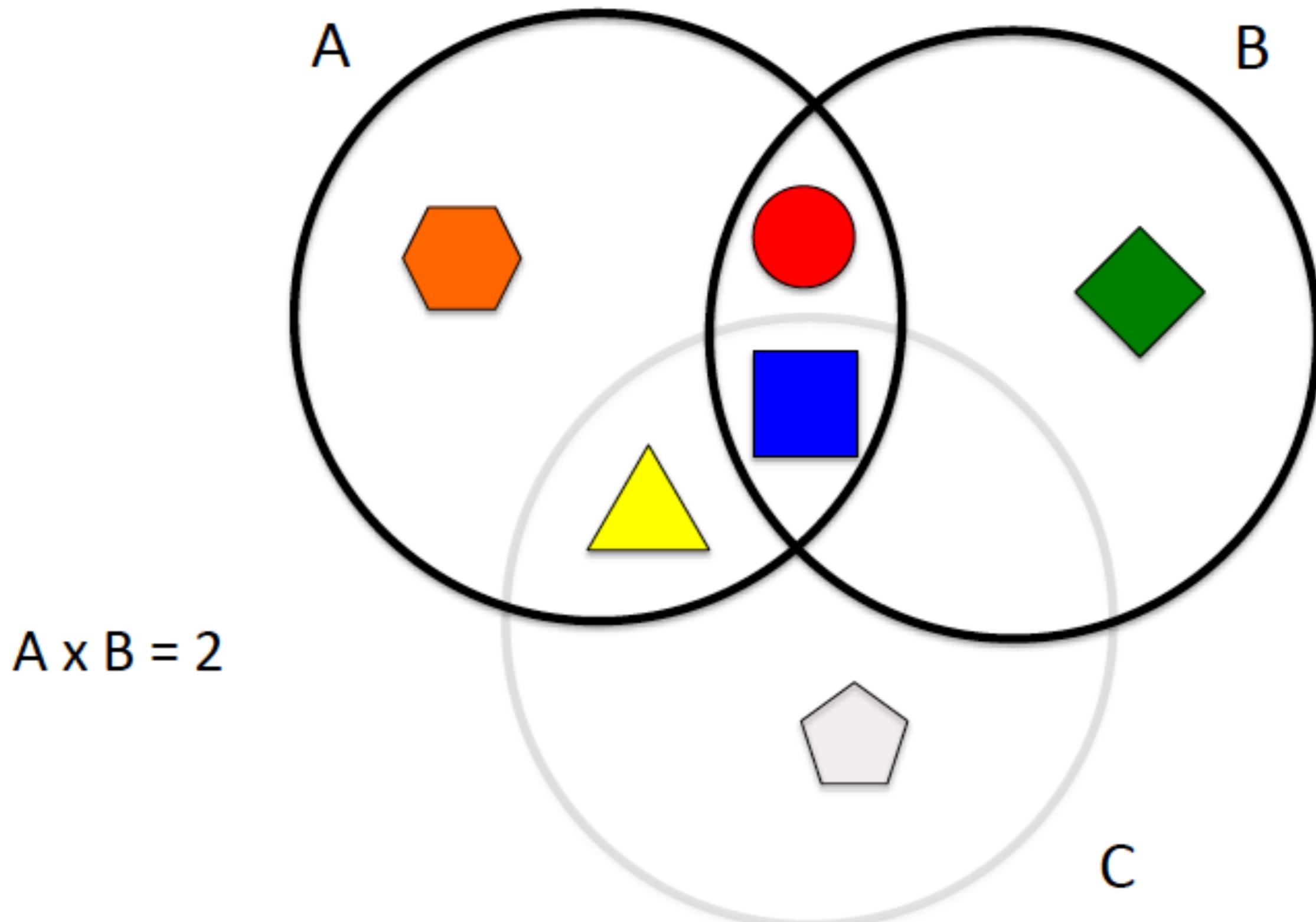
Beta diversity

Beta diversity



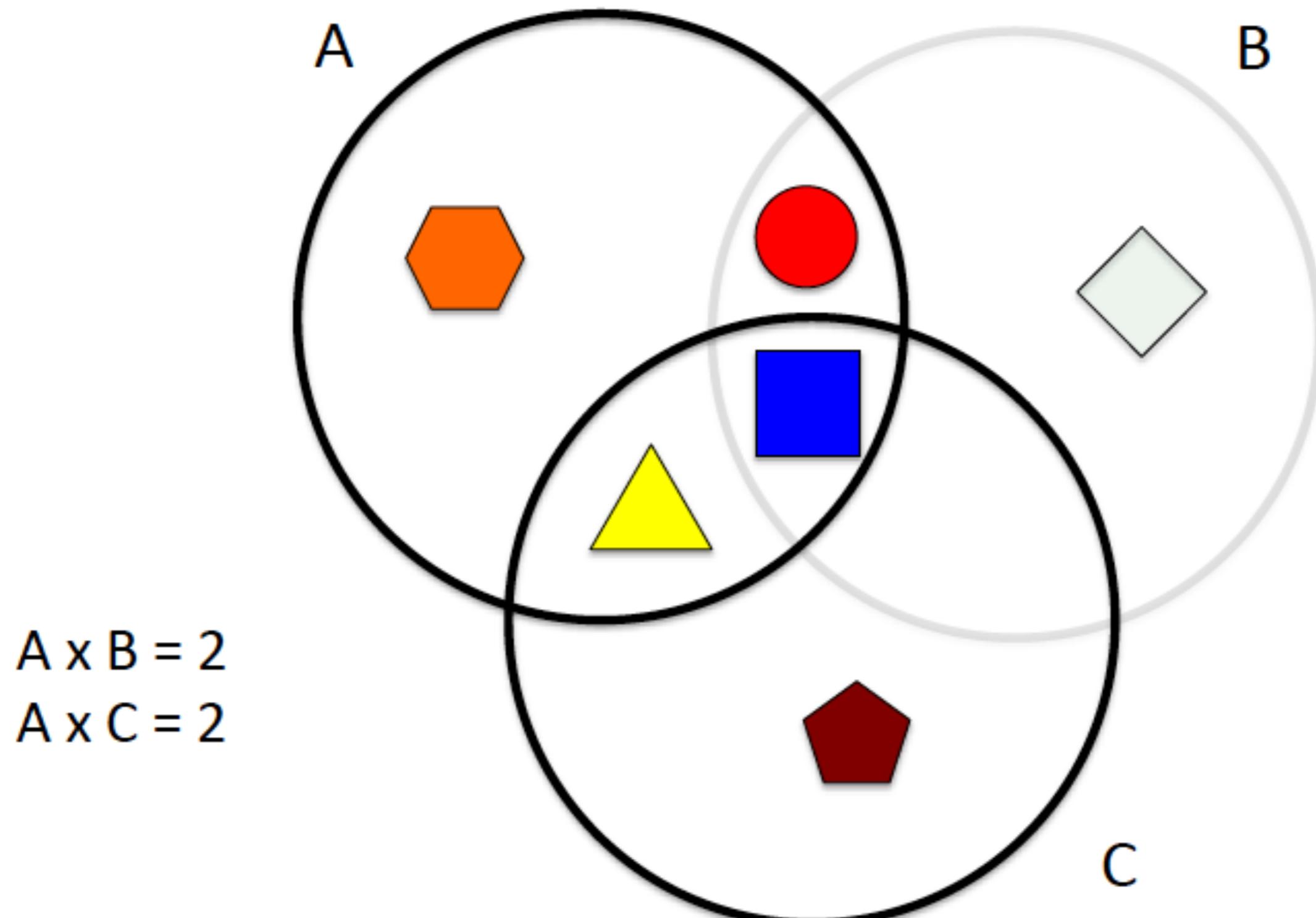
Ref: Daniel McDonald; American Gut Project

Beta diversity



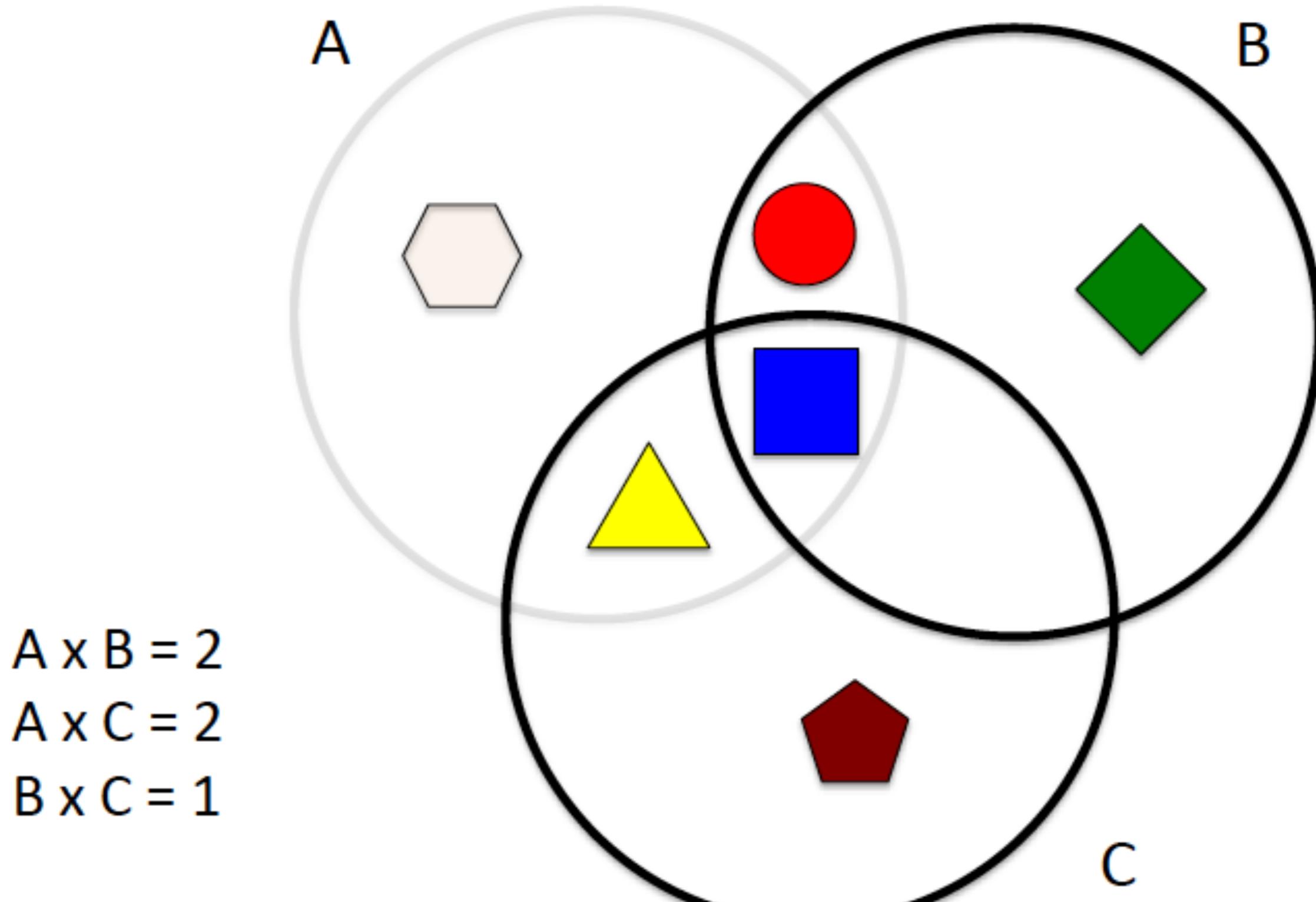
Ref: Daniel McDonald; American Gut Project

Beta diversity

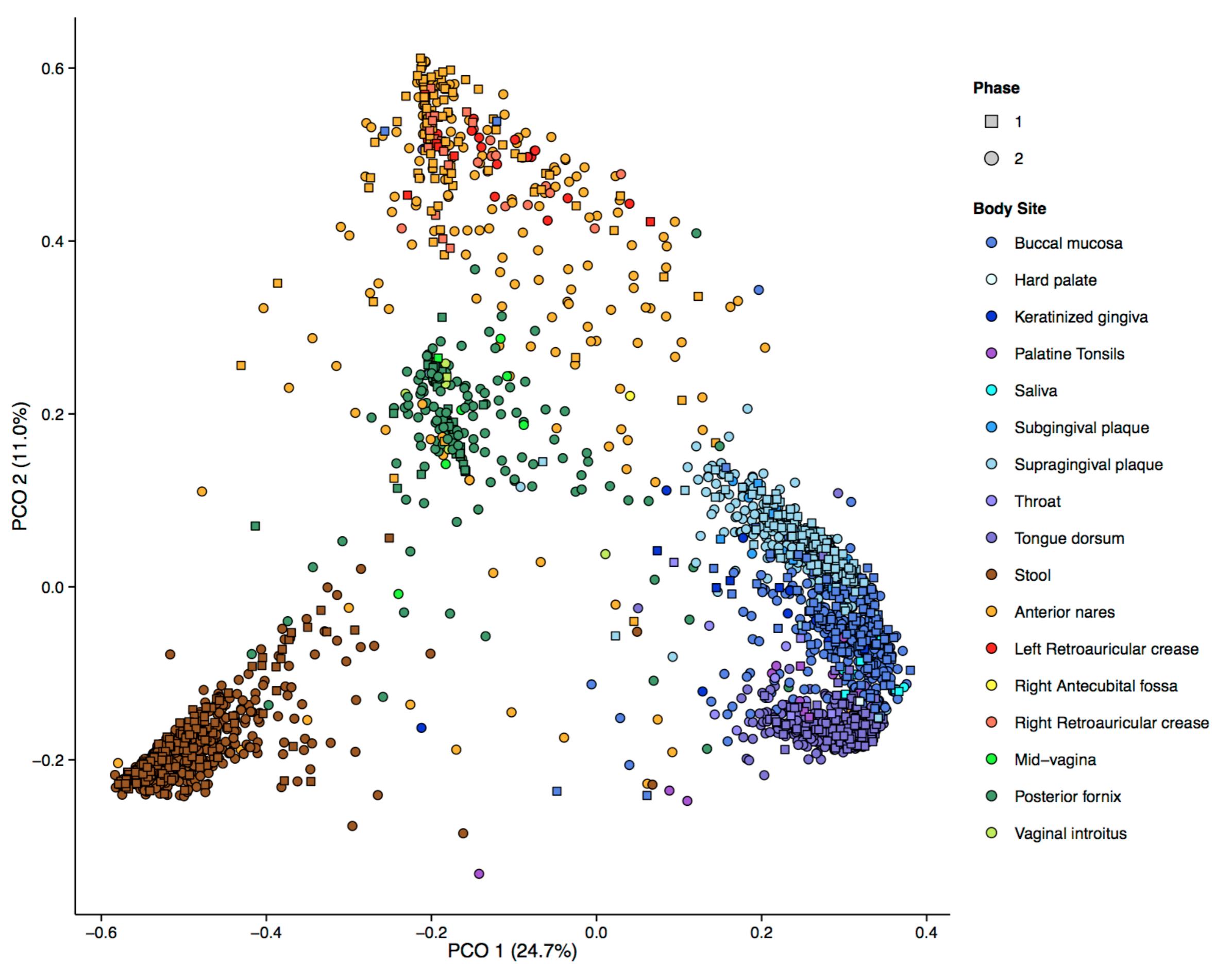


Ref: Daniel McDonald; American Gut Project

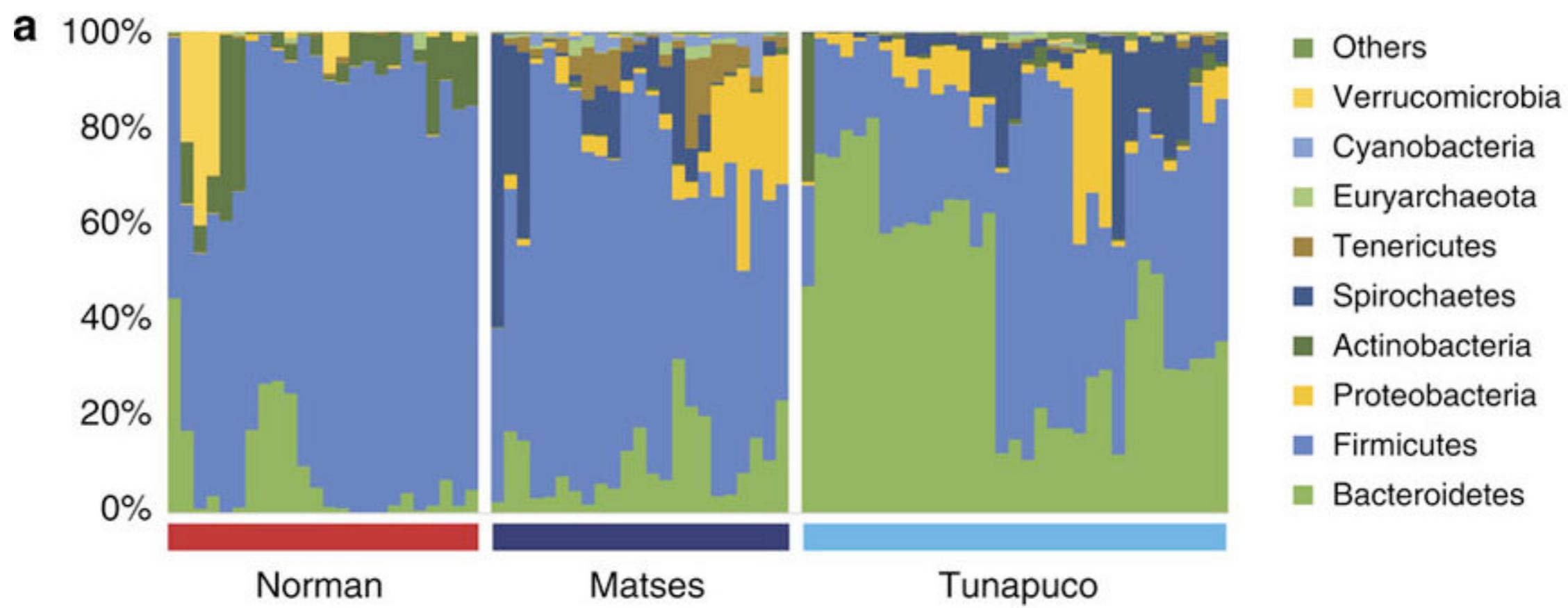
Beta diversity



Ref: Daniel McDonald; American Gut Project



Species Analyses



A. Obregon-Tito, et al. Nat Comms, 2015

Questions?