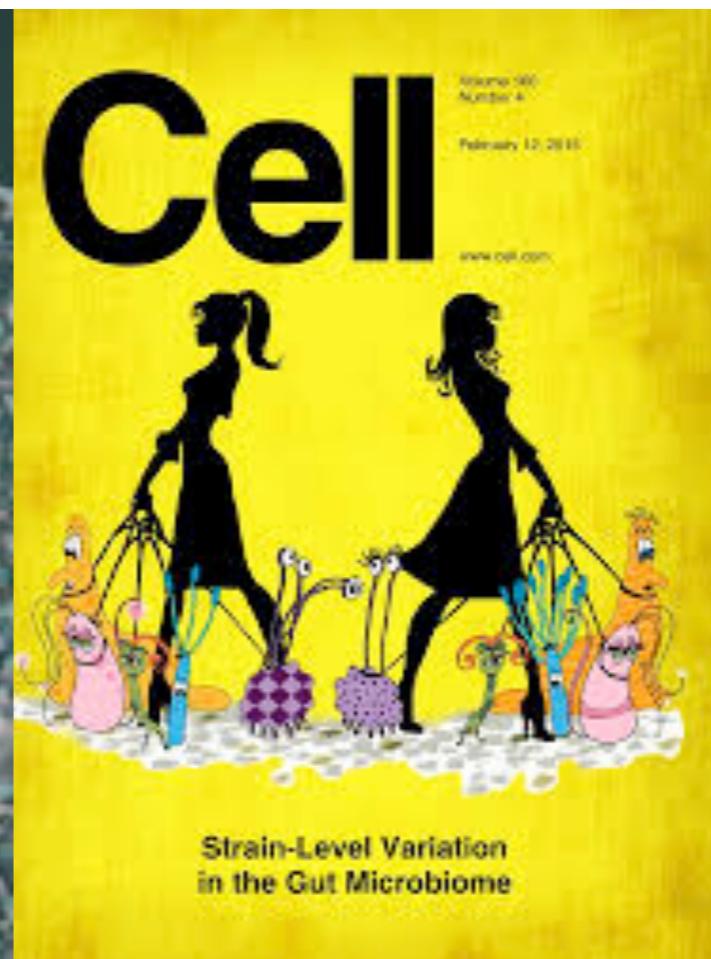
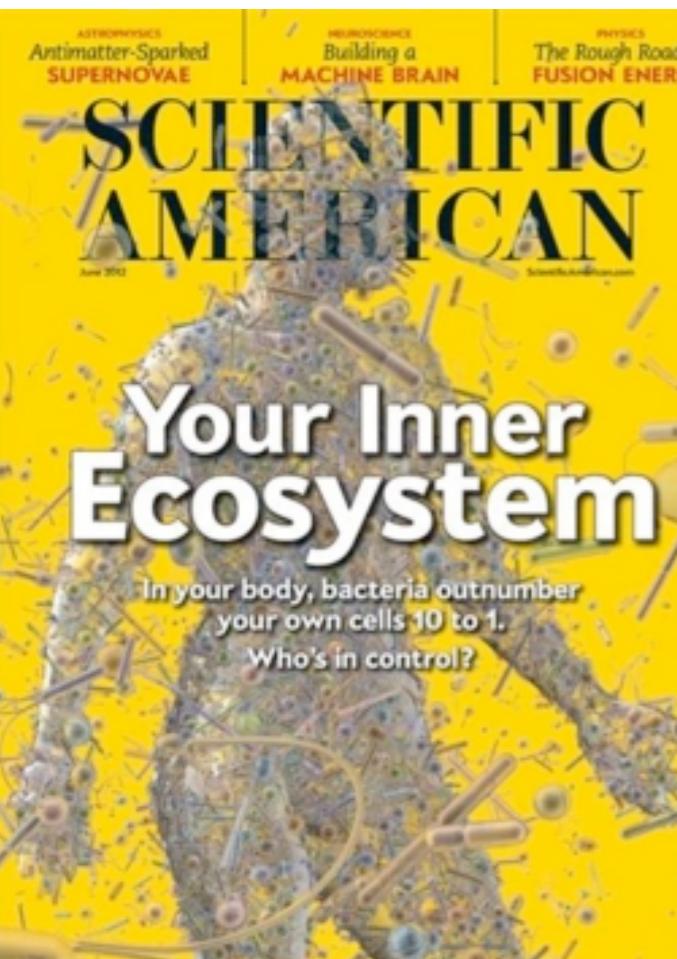


# Amplicon Analysis via QIIME2

Laura Weyrich and Erika Ganda

lsw132@psu.edu and ganda@psu.edu

@lsweyrich and @erika\_ganda



# **Introductions:**

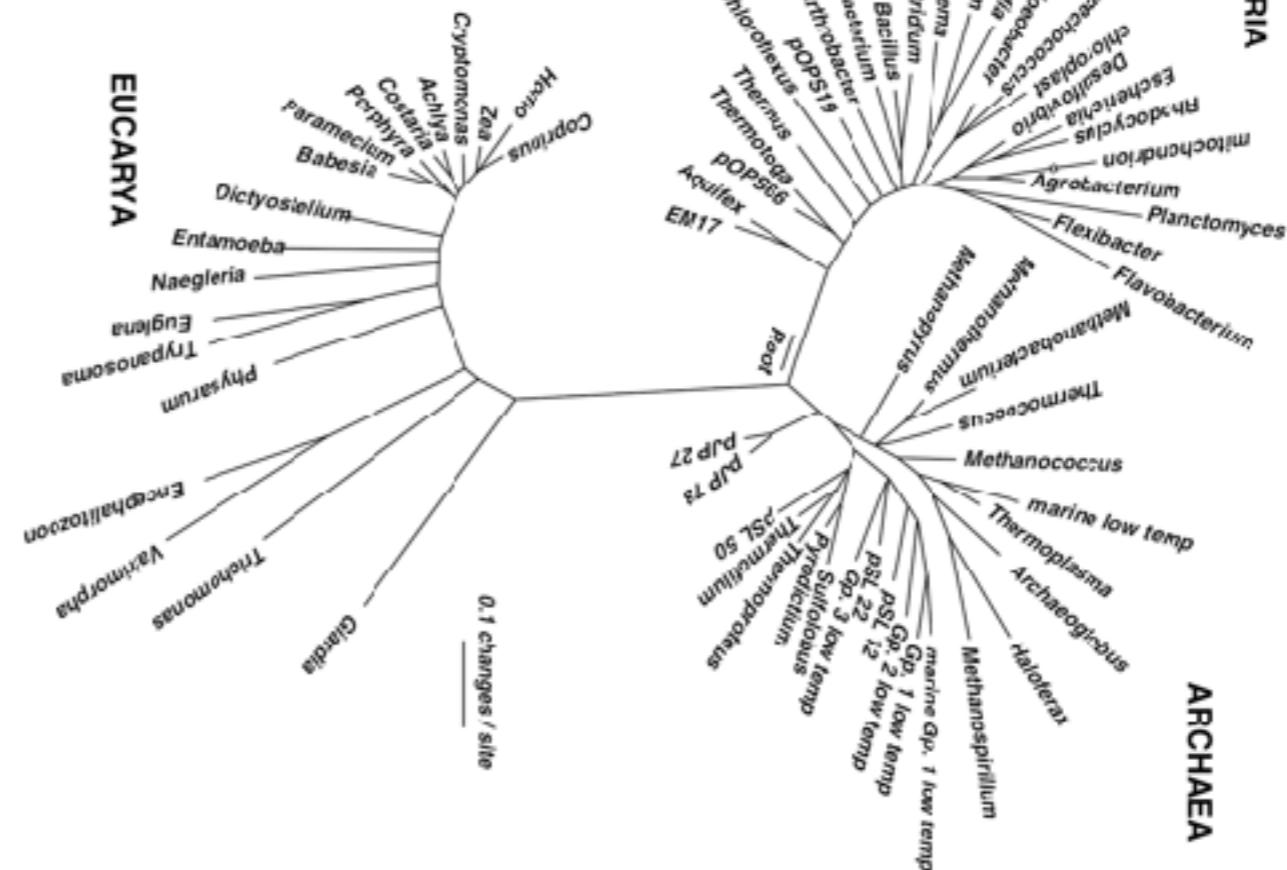
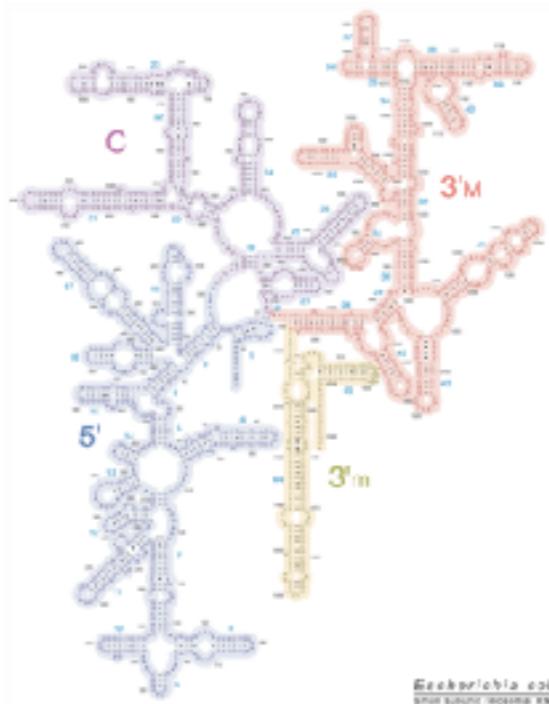
**Laura Weyrich**  
**Erika Ganda**

**Abby Gancz**  
**Christine Ta**  
**Susan Tian**  
**Marysabel Mendez**

**\*Slides are on GitHub:**

# The experimental history

- Culturing microbes = 2% of total bacterial diversity
- Ribosomal RNA (rRNA) sequences are ‘barcodes’ for bacteria species
  - 1986 by Norman Pace
- 

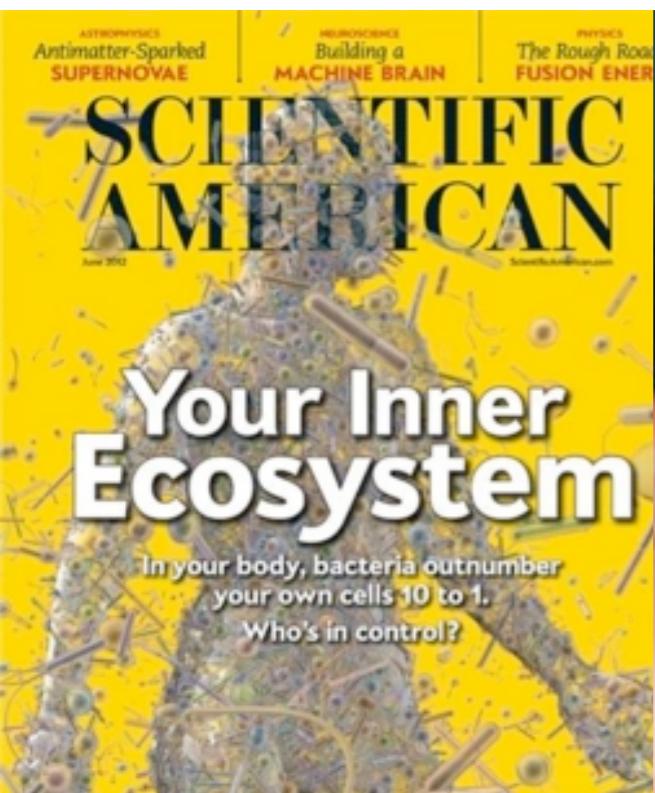


Pace, Norman R., David A. Stahl, David J. Lane, and Gary J. Olsen. 1986. “The Analysis of Natural Microbial Populations by Ribosomal RNA Sequences.” In *Advances in Microbial Ecology*, edited by K. C. Marshall, 1–55. Advances in Microbial Ecology 9. Springer US. [http://link.springer.com.proxy.library.adelaide.edu.au/chapter/10.1007/978-1-4757-0611-6\\_1](http://link.springer.com.proxy.library.adelaide.edu.au/chapter/10.1007/978-1-4757-0611-6_1).

# HUMAN EVOLUTION MAY BE TIGHTLY LINKED TO OUR MICROBIAL EVOLUTION

**Microbiota:**  
the microorganisms (bacteria, fungi, viruses) that live in your body

**100 trillion** bacterial cells  
**>50%** of total cells  
**>1,000** species  
**1.4 kg** of body weight



# MICROBIOTA FUNCTIONS ARE MORE CRITICAL THAN THE SPECIES THEMSELVES

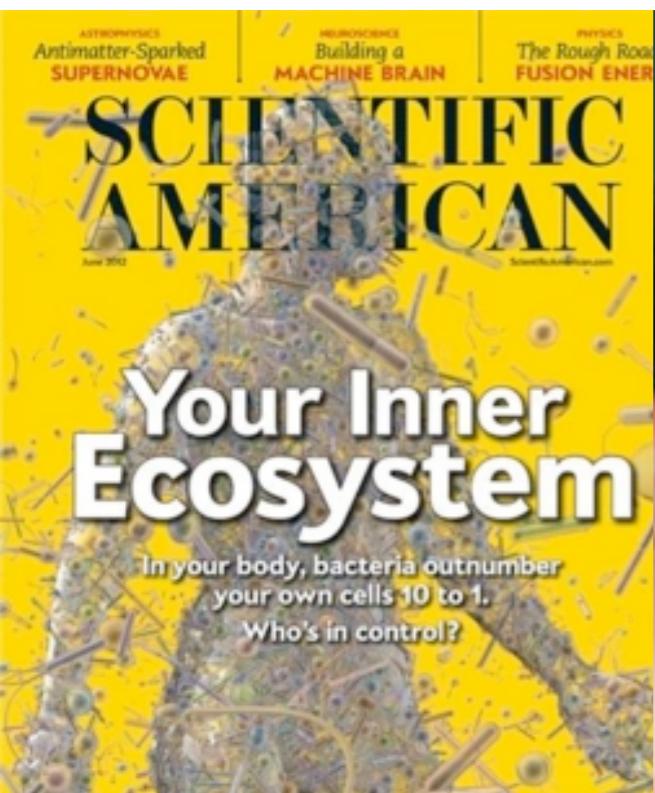
## Microbiome:

the genetic and environmental content of the microbiota present in the body

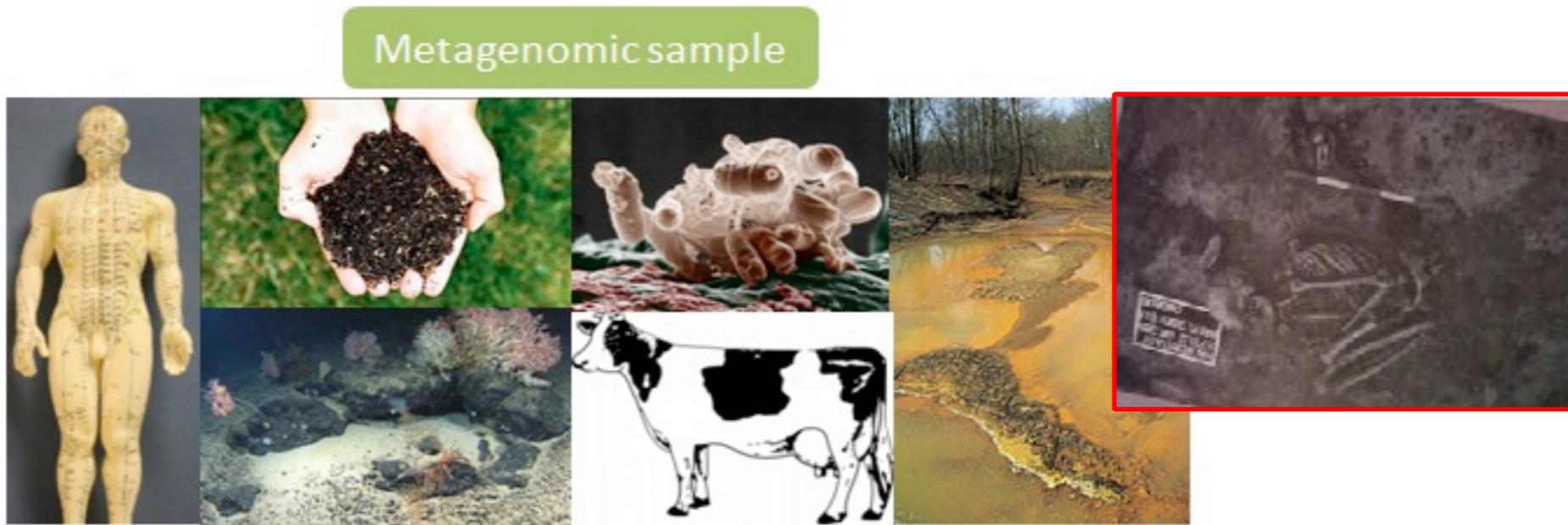
Outnumbers human cells **100 to 1!**

**2-5 million** genes per individual

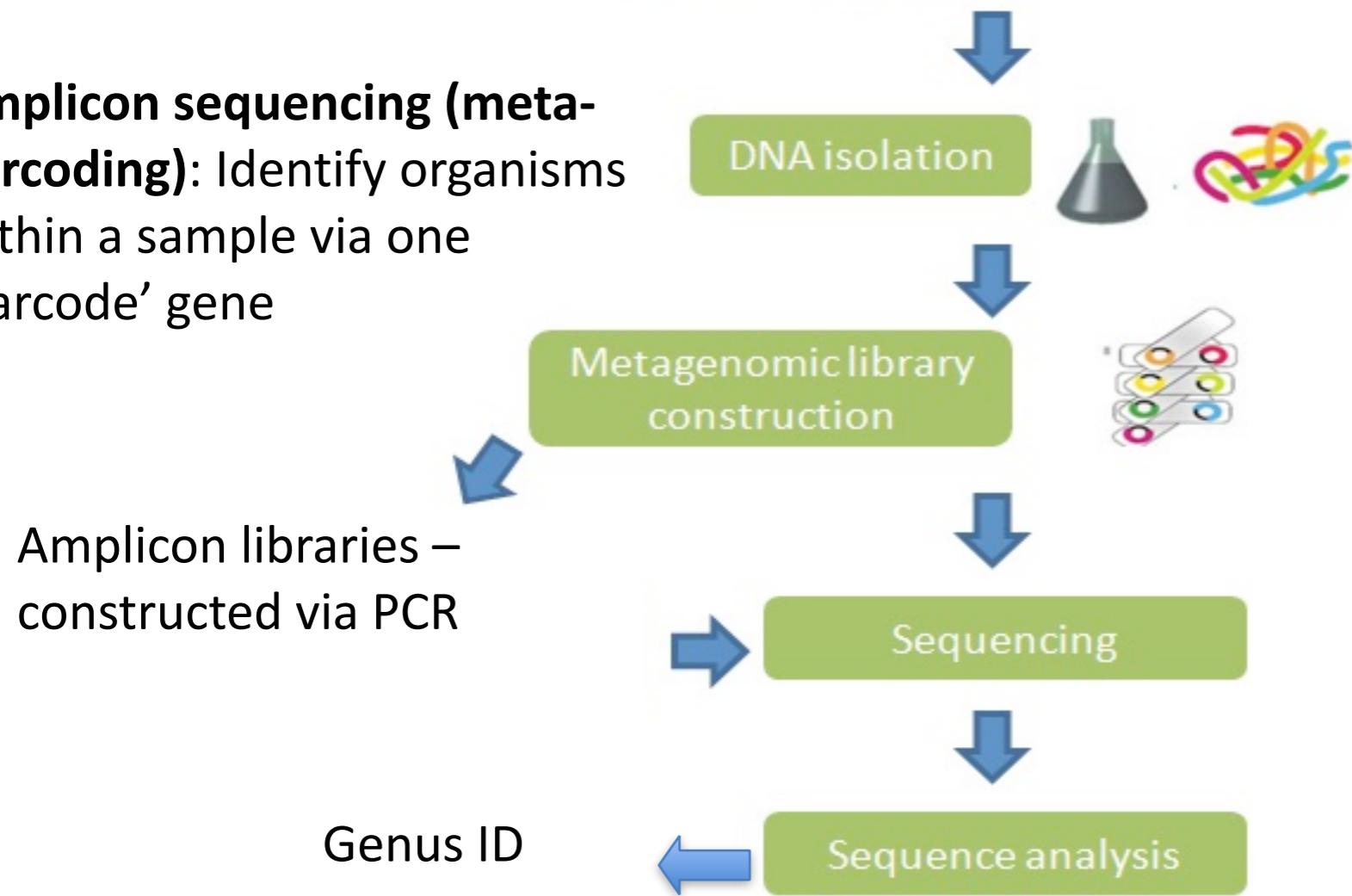
**99%** of your genetic makeup!



# Microbiome NGS analysis

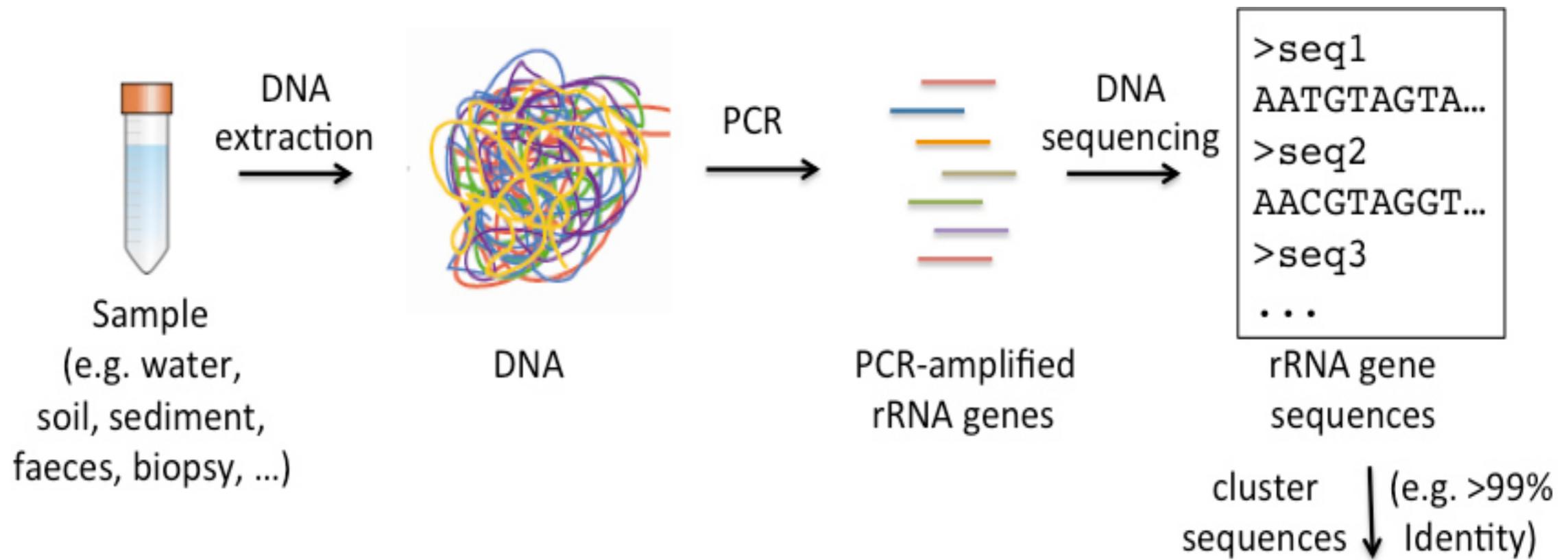


**Amplicon sequencing (meta-barcoding):** Identify organisms within a sample via one 'barcode' gene



# Amplicon Based NGS 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

Compare to Known  
References

Sequences

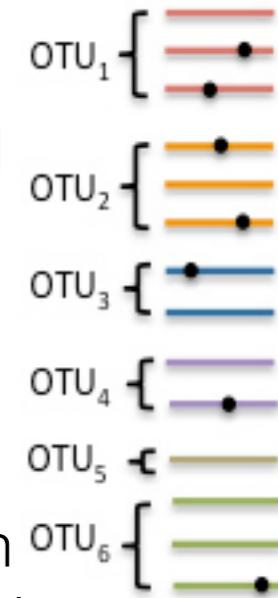


with millions  
of taxonomically  
classified  
rRNA sequences  
(e.g. RDP, Silva)

Operational  
Taxonomic  
Units  
(OTUs)

Or

Amplicon  
Sequence Variants  
(ASVs)

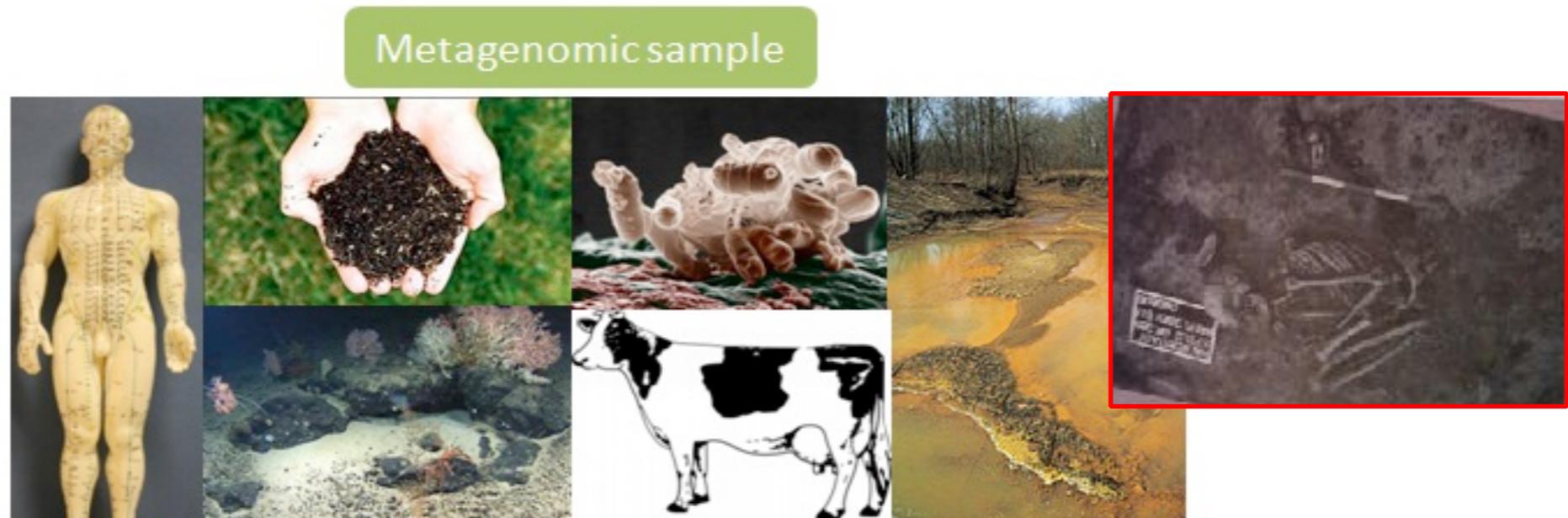


# Issues with 16S analysis

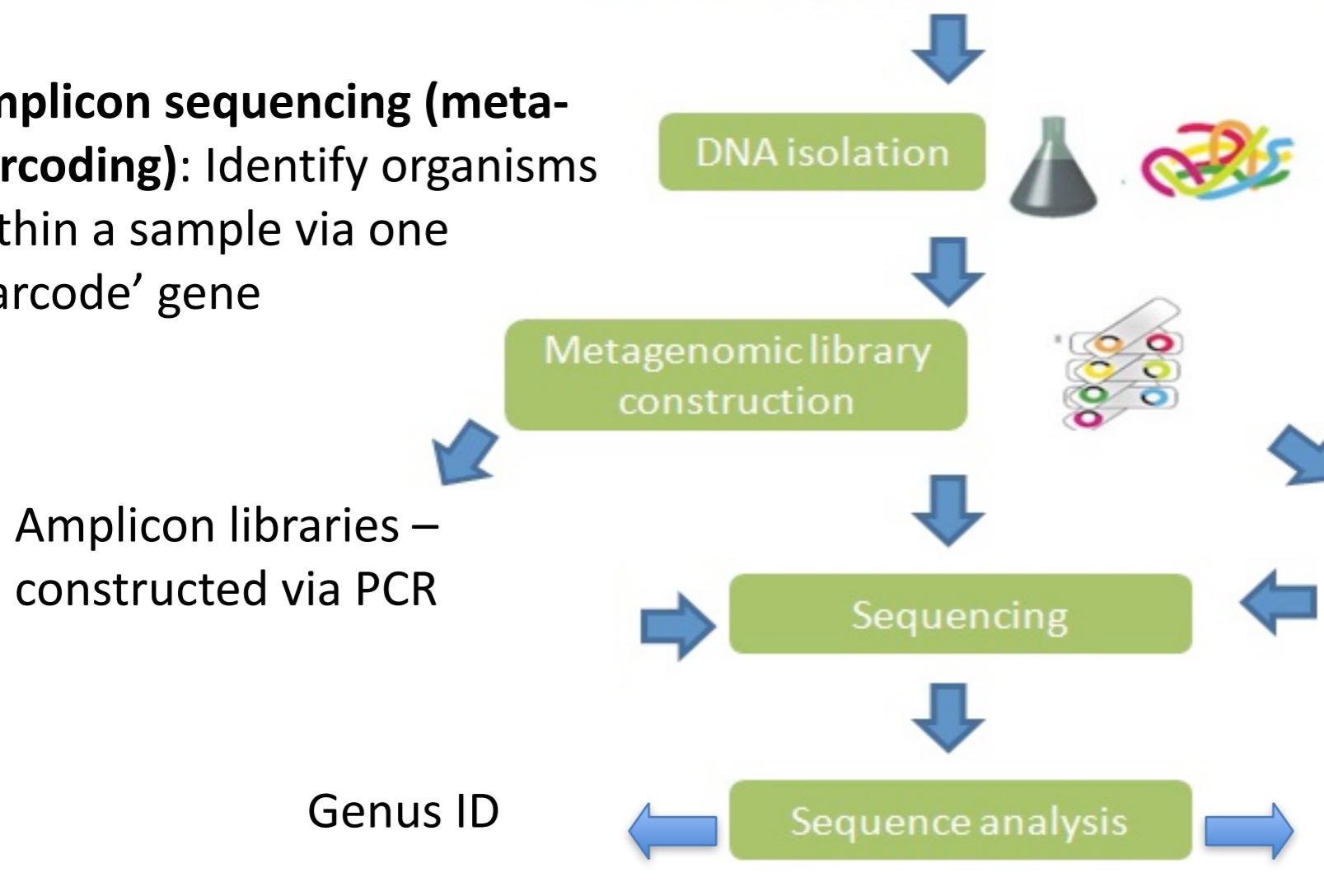
- Sample Preservation
  - Frozen? Ancient?
- Contamination
  - Laboratory, reagents, and technicians
- Over amplification
- Inappropriate bioinformatic analysis



# Microbiota/Microbiome Analysis: NGS



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



**Shotgun sequencing:** random sampling of all the DNA present

Amplicon libraries – constructed via PCR

Shotgun NGS libraries – constructed via ligation of adapters

Genus ID

Species ID and Function

# Pros/Cons of NGS Approaches

|      | Amplicon Sequencing   | Shotgun Sequencing  |
|------|---|---|
| Pros | Inexpensive<br>Fast<br>Technically Easy   | Genomes<br>Functional Information<br>Little Technical Bias<br>Species Level ID    |
| Cons | No Functional Data<br>Technical Biases<br>ID limited to genera<br>No genomic info | Very expensive<br>Labor intensive<br>Poor Analysis Programs<br>Need Deep Coverage |

# **What types of analyses can you do with amplicon data?**

- Data assessment and cleaning
- Diversity (alpha-diversity)
- Composition (beta-diversity)
- Taxa identification (ASVs)
- Predict functions based on known information
- Assess what factors drives changes in microbiota

# **Questions?**

To Github!

# Qiime2

## What is QIIME 2?

---

QIIME 2 is a powerful, extensible, and decentralized microbiome analysis package with a focus on data and analysis transparency. QIIME 2 enables researchers to start an analysis with raw DNA sequence data and finish with publication-quality figures and statistical results.

Key features:

- Integrated and automatic tracking of data provenance
- Semantic type system
- Plugin system for extending microbiome analysis functionality
- Support for multiple types of user interfaces (e.g. API, command line, graphical)

QIIME 2 is a complete redesign and rewrite of the [QIIME 1](#) microbiome analysis pipeline. QIIME 2 will address many of the limitations of QIIME 1, while retaining the features that makes QIIME 1 a powerful and widely-used analysis pipeline.

QIIME 2 currently supports an initial end-to-end microbiome analysis pipeline. New functionality will regularly become available through QIIME 2 plugins. You can view a list of plugins that are currently available on the [QIIME 2 plugin availability](#) page. The [future plugins](#) page lists plugins that are being developed.



# Glossary

## **action**

A general term for a [method](#), a [visualizer](#), or a [pipeline](#). Actions are always defined by QIIME 2 [plugins](#).

## **artifact**

Artifacts are QIIME 2 [results](#) that are generally considered to represent intermediate data in an analysis, meaning that an artifact is generated by QIIME 2 and intended to be consumed by QIIME 2 (rather than by a human). Artifacts can be generated either by importing data into QIIME 2 or as output from a QIIME 2 [action](#). When written to file, artifacts typically have the extension `.qza`, which stands for *QIIME Zipped Artifact*. Artifacts can be provided as input to QIIME 2 [actions](#), loaded with tools such as the QIIME 2 Artifact API for use with Python 3 or [qiime2R](#) for use with R, or exported from QIIME 2 for use with other software.

## **data provenance**

See [decentralized data provenance](#).

## **data format**

A view of an [artifact](#) as a file or multiple files stored on disk. QIIME 2 supports many data (or file) formats, and multiple data formats are sometimes available for importing or exporting of QIIME 2 [artifacts](#) of a given [semantic type](#).

## **data type**

A view of an [artifact](#) as an in-memory data representation. Data types are generally only encountered by Artifact API users or plugin developers. QIIME 2 supports many data types, and multiple data types are sometimes available for viewing QIIME 2 [artifacts](#) of a given [semantic type](#).

## **decentralized data provenance**

Information describing how a QIIME 2 [result](#) was generated. This will include details on all of the QIIME 2 [actions](#) that led to the creation of an [artifact](#), including the values of all [parameters](#), and references to all [inputs](#) and [results](#) as [UUIDs](#). Data provenance additionally contains the literature citations that are relevant to the generation of a QIIME 2 [result](#). Those citations should be included in all published work that derives from a given QIIME 2 [result](#).

All QIIME 2 [results](#) contain embedded data provenance which can be visualized with [QIIME 2 View](#). Because the data provenance is embedded in the [results](#) themselves, as opposed to being stored in a centralized database that maintains records on all [results](#) (for example), QIIME 2's data provenance is described as being decentralized.



## Glossary

### feature

A unit of observation, such as an operational taxonomic unit, a sequence variant, a gene, a metabolite, etc. This generic term is used because QIIME 2 can support many different types of features.

### input

An [artifact](#) (i.e., non-primitive) provided to an [action](#). For example, `table` is an [input](#) to the [filter-features action](#) in the [q2-feature-table plugin](#).

### method

A type of QIIME 2 [action](#) that takes one or more [artifacts](#) or [parameters](#) as input, and produces one or more [artifacts](#) as output. For example, the [filter-features action](#) in the [q2-feature-table plugin](#) is a [method](#).

### output

A [result](#) generated by running an [action](#). For example, `filtered-table` is an [output](#) from the [filter-features action](#) in the [q2-feature-table plugin](#).

### parameter

A primitive (i.e., non-artifact) provided to an [action](#). For example, `min-frequency` is a [parameter](#) to the [filter-features action](#) in the [q2-feature-table plugin](#). See [primitive type](#).

### pipeline

A type of QIIME 2 [action](#) that typically combines two or more other [actions](#). A pipeline takes one or more [artifacts](#) or [parameters](#) as input, and produces one or more [results](#) ([artifacts](#) and/or [visualizations](#)) as output. For example, the [core-metrics action](#) in the [q2-diversity plugin](#) is a [pipeline](#).

### plugin

A plugin provides analysis functionality in the form of [actions](#). All plugins can be accessed through all interfaces. Plugins can be developed and distributed by anyone. As of this writing, a collection of plugins referred to as the “core distribution” is provided on installation of QIIME 2. Additional plugins can be installed, and the primary resource enabling discovery of additional plugins is the [QIIME 2 Library](#). Anyone with a QIIME 2 Forum account can share their plugins on the QIIME 2 Library. We plan to phase out the core distribution as we move toward distributing all QIIME 2 plugins through the QIIME 2 Library.

### provenance

See [decentralized data provenance](#).

### primitive type

A type used to define a [parameter](#) to a QIIME 2 [action](#). For example, strings (i.e., text), integers, and booleans (i.e., true or false values) are primitives. Primitives are only used as input to [actions](#), and never generated as output by QIIME 2.



## Glossary

### **qza**

See [artifact](#).

### **qzv**

See [visualization](#).

### **result**

A general term for an [artifact](#) or a [visualization](#).

### **sample**

An individual unit of study in an analysis.

### **semantic type**

A semantic type describes the meaning of data in QIIME 2. All [results](#) in QIIME 2 have a single semantic type associated with them, and when importing data into QIIME 2, the user must provide the semantic type of that data.

The use of semantic types by QIIME 2 provides an unambiguous way to communicate with others about data, and allows QIIME 2 to reason about data and help users prevent error. An example is helpful for illustrating what semantic types are and how they're used by QIIME 2. QIIME 2 contains two related semantic types, `Phylogeny[Rooted]` and `Phylogeny[Unrooted]`, which represent rooted and unrooted phylogenetic trees, respectively. Both rooted and unrooted phylogenetic trees can be stored in newick files, and it isn't possible to easily tell if a phylogenetic tree is rooted or not without parsing the file. Some [actions](#), such as the [beta-phylogenetic method](#) in the [q2-diversity plugin](#), should be applied only to a rooted phylogenetic tree. By associating a semantic type with a phylogenetic tree artifact, QIIME 2 can determine if the correct type of data is being provided to an [action](#), without having to first parse the file (which might be slow, and therefore delay the amount of time before an error can be presented to a user), and then possibly make assumptions based on what is observed. If a user accidentally provides data of a semantic type that is not acceptable for a QIIME 2 [action](#), QIIME 2 can quickly detect this mismatch and provide the user with detailed information on the error and how to correct it.

Semantic types shouldn't be confused with [data formats](#) which define how data is represented on disk. For example, another QIIME 2 semantic type, the `FeatureTable[Frequency]`, can be written to a BIOM-formatted file or to a tab-separated text file. By differentiating [data formats](#) from semantic types, QIIME 2 can support import and export of different file formats based on a user's needs. Semantic types should also not be confused with [data types](#). For example, the `FeatureTable[Frequency]` semantic type could be represented in memory as a `biom.Table` object or a `pandas.DataFrame` object, and for different applications one of these representations might be more useful than the other. Regardless of which in-memory representation is used, the meaning of the data is the same. By differentiating data types and semantic types, QIIME 2 allows developers and users to choose the data structure that is most convenient for them for a given task.



# What is metadata and why is it important?

Metadata is information about your samples (e.g. date collected, patient age, sex, pH, etc). This information should be contained within a single spreadsheet that has samples as rows and variables as columns.

In order to use QIIME, you **must** have your mapping spreadsheet correctly formatted. Set this file up as a google sheet, using [this example](#). In order to check whether your file is correctly formatted, use the [Keemei plugin](#) for google sheets. Once Keemei says your spreadsheet is correctly formatted for QIIME2, you're ready to proceed!

Create a .txt version of your metadata spreadsheet and transfer it to the server so that it resides in your working directory. You can then inspect further if you want using:



# What is a FastQ file?

## Format [\[edit\]](#)

---

A FASTQ file normally uses four lines per sequence.

- Line 1 begins with a '@' character and is followed by a sequence identifier and an *optional* description (like a [FASTA](#) title line).
- Line 2 is the raw sequence letters.
- Line 3 begins with a '+' character and is *optionally* followed by the same sequence identifier (and any description) again.
- Line 4 encodes the quality values for the sequence in Line 2, and must contain the same number of symbols as letters in the sequence.

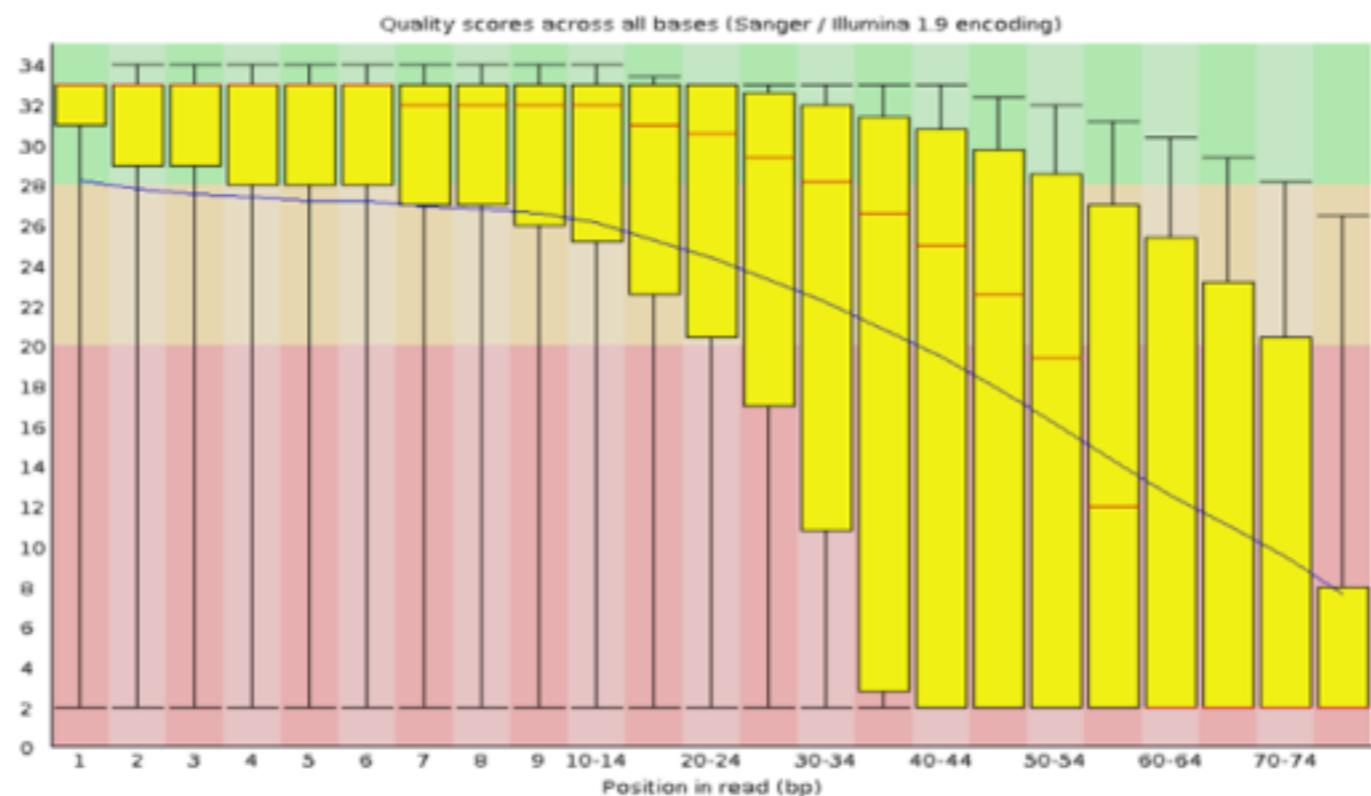
A FASTQ file containing a single sequence might look like this:

```
@SEQ_ID
GATTTGGGGTTCAAAGCAGTATCGATCAAATAGTAAATCCATTGTTCAACTCACAGTTT
+
!''*((((***+))%%%++)(%%%).1***-+*'')**55CCF>>>>CCCCCCCC65
```



# What is a Phred Score?

Next Generation Sequencing techniques have brought new insights into -omics data analysis, mostly thanks to their reliability in detecting biological variants. This reliability is usually measured using a value called Phred quality score (or Q score).



<https://medium.com/@robertopreste/phred-quality-score-2837415f0af>



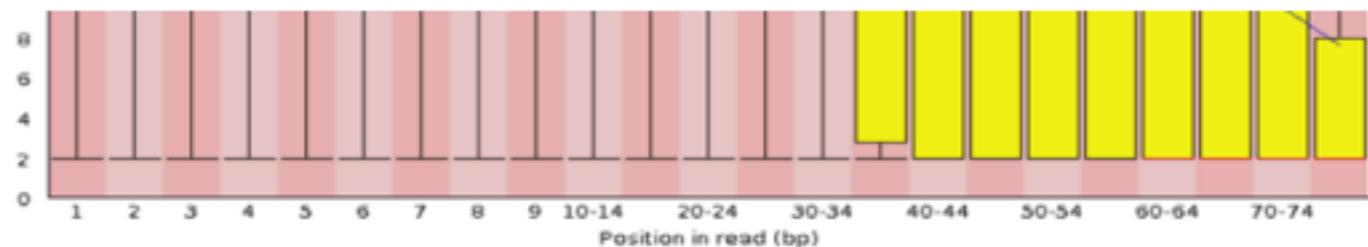
# What is a Phred Score?

Next Generation Sequencing techniques have brought new insights into -omics data analysis, mostly thanks to their reliability in detecting biological variants. This reliability is usually measured using a value called Phred quality score (or Q score).

Search this file...

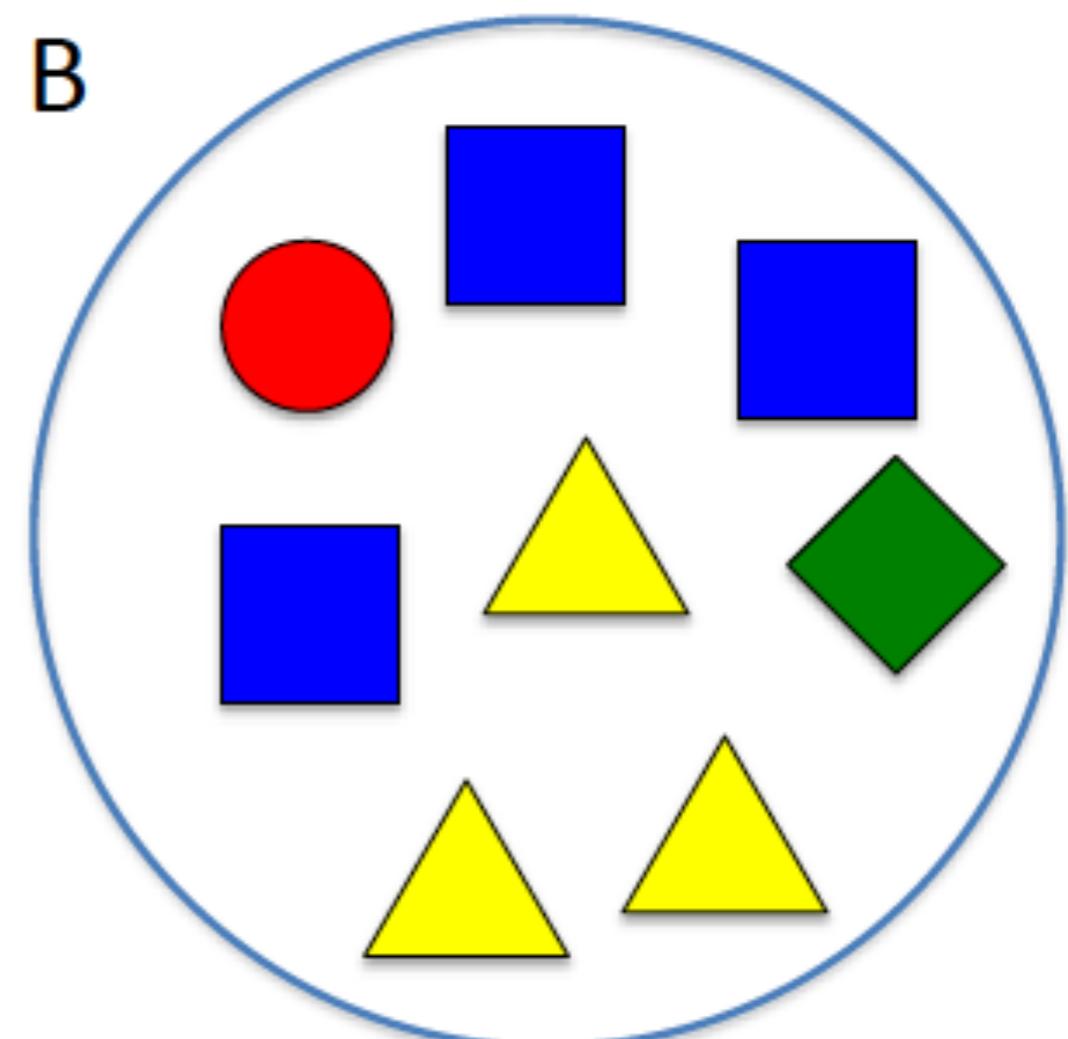
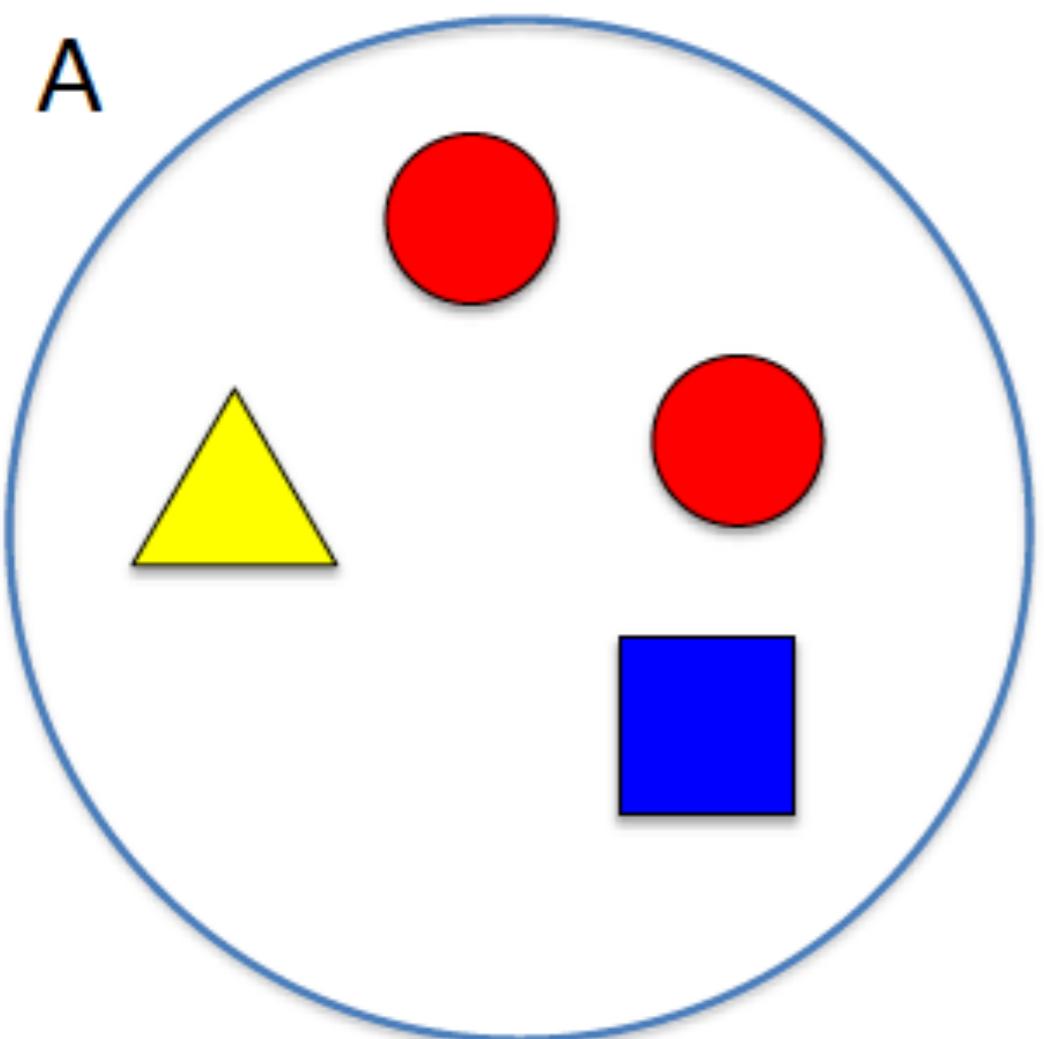
| 1 | Phred Quality Score | Incorrect base call prob | Base call accuracy |
|---|---------------------|--------------------------|--------------------|
| 2 | 10                  | 1 in 10                  | 90%                |
| 3 | 20                  | 1 in 100                 | 99%                |
| 4 | 30                  | 1 in 1000                | 99.9%              |
| 5 | 40                  | 1 in 10000               | 99.99%             |

phred\_quality\_score\_1.csv hosted with ❤ by GitHub [view raw](#)



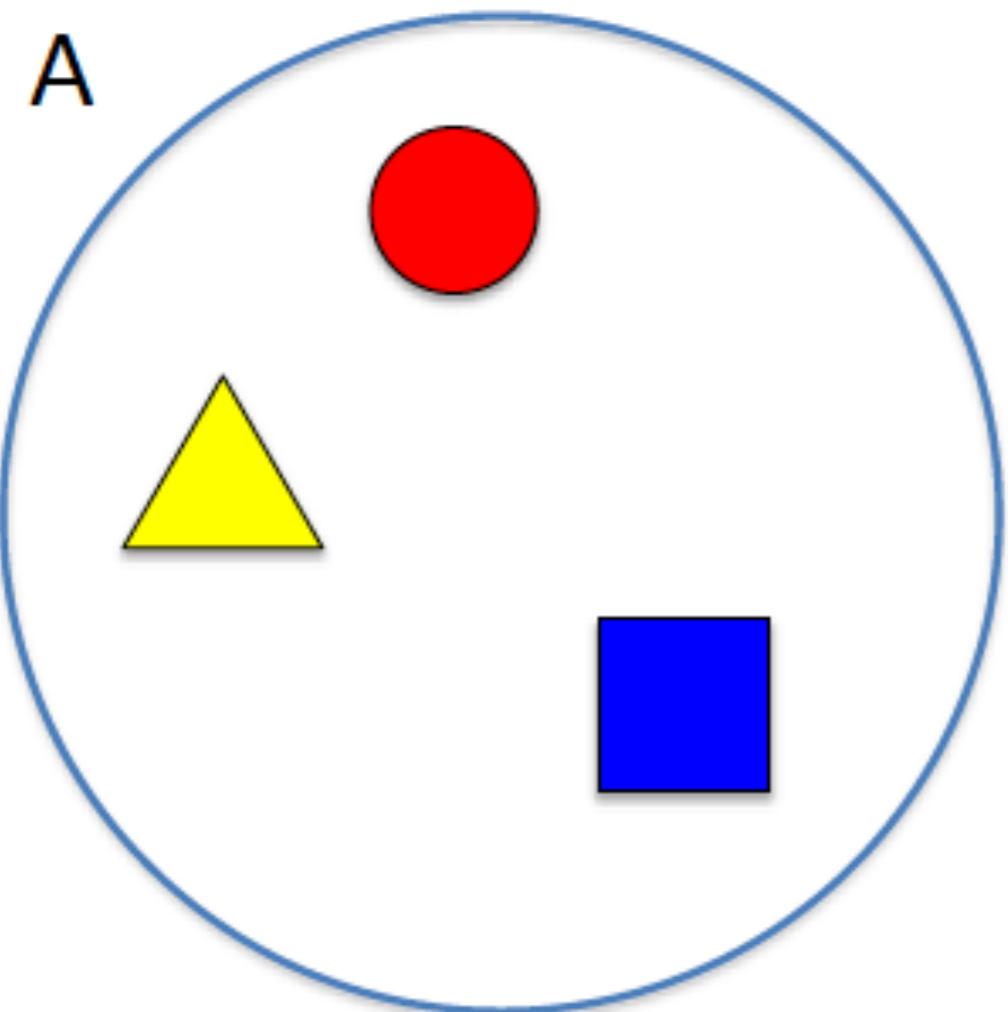
<https://medium.com/@robertopreste/phred-quality-score-2837415f0af>

# Alpha diversity

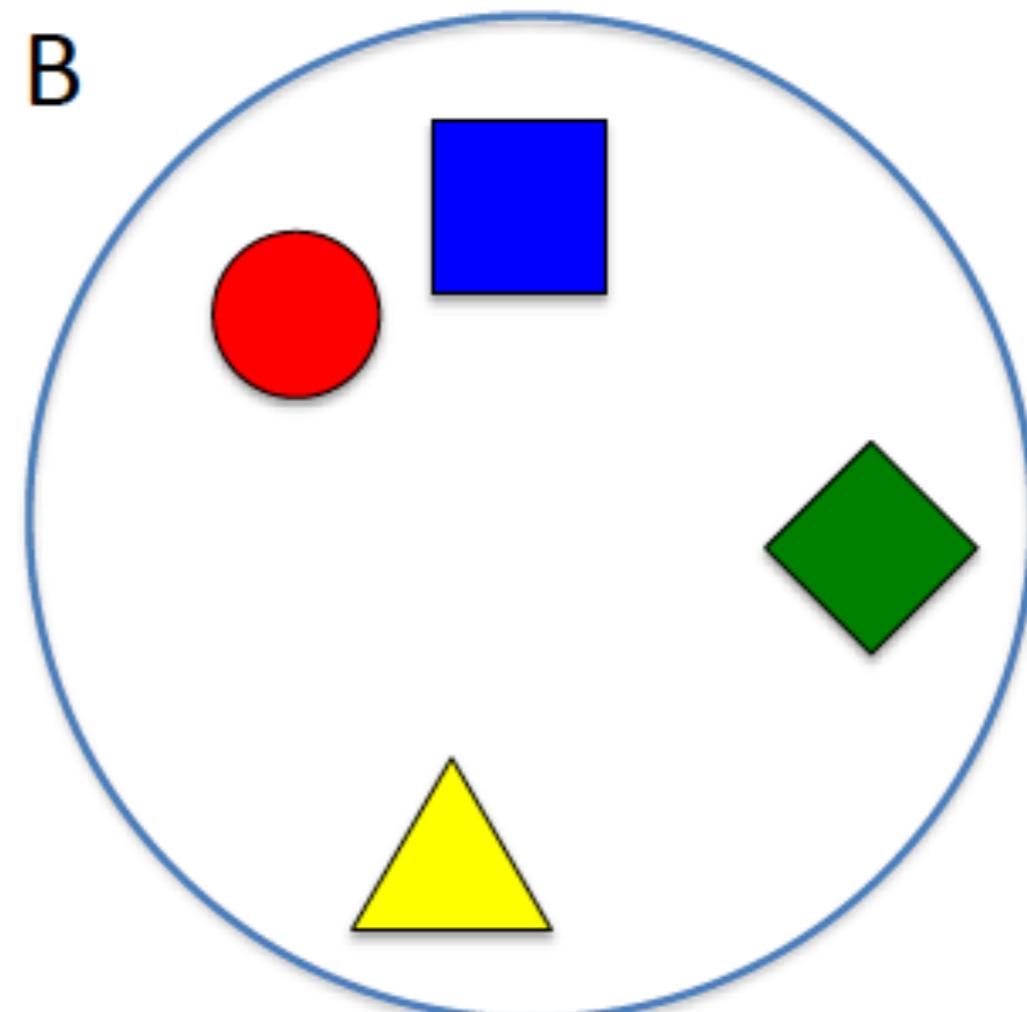


Ref: Daniel McDonald; American Gut Project

# Alpha diversity



Alpha diversity = 3

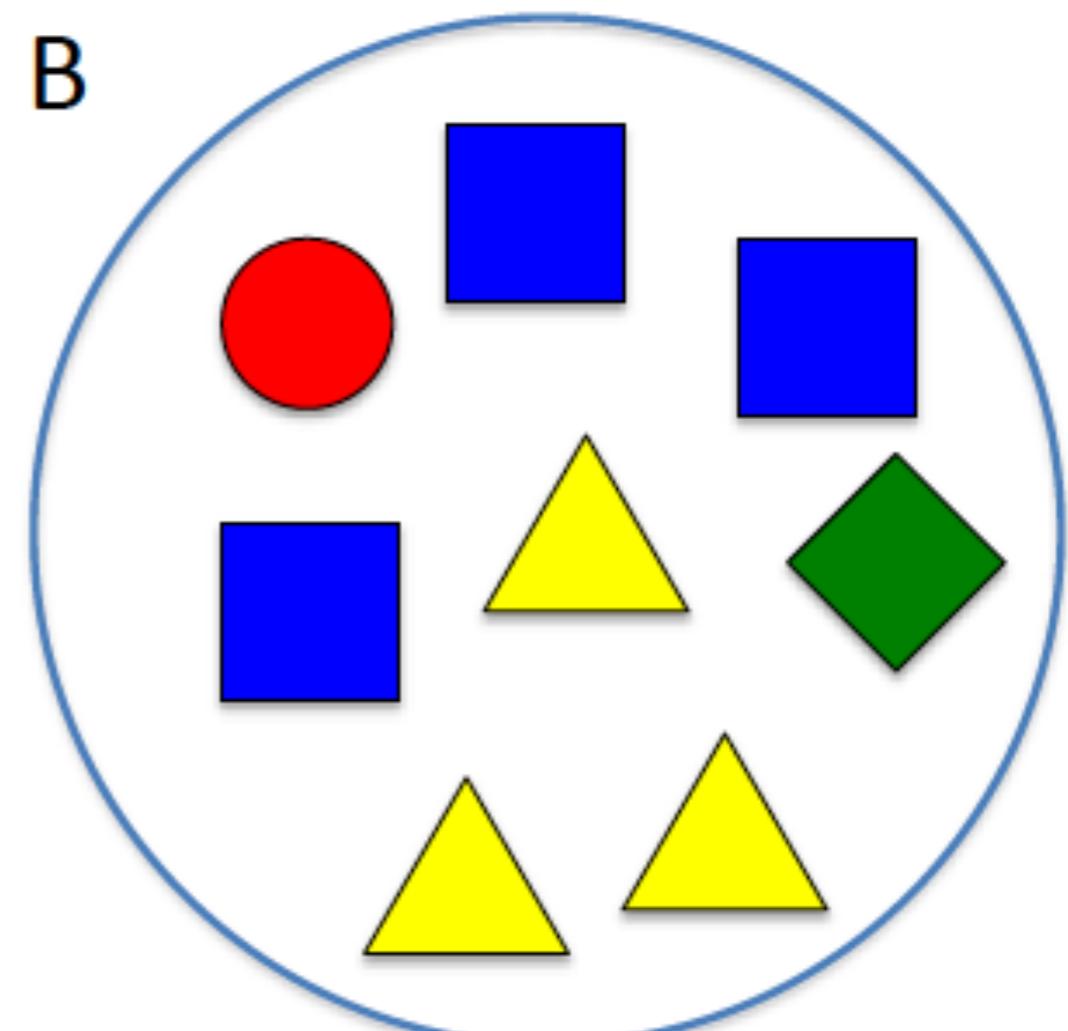
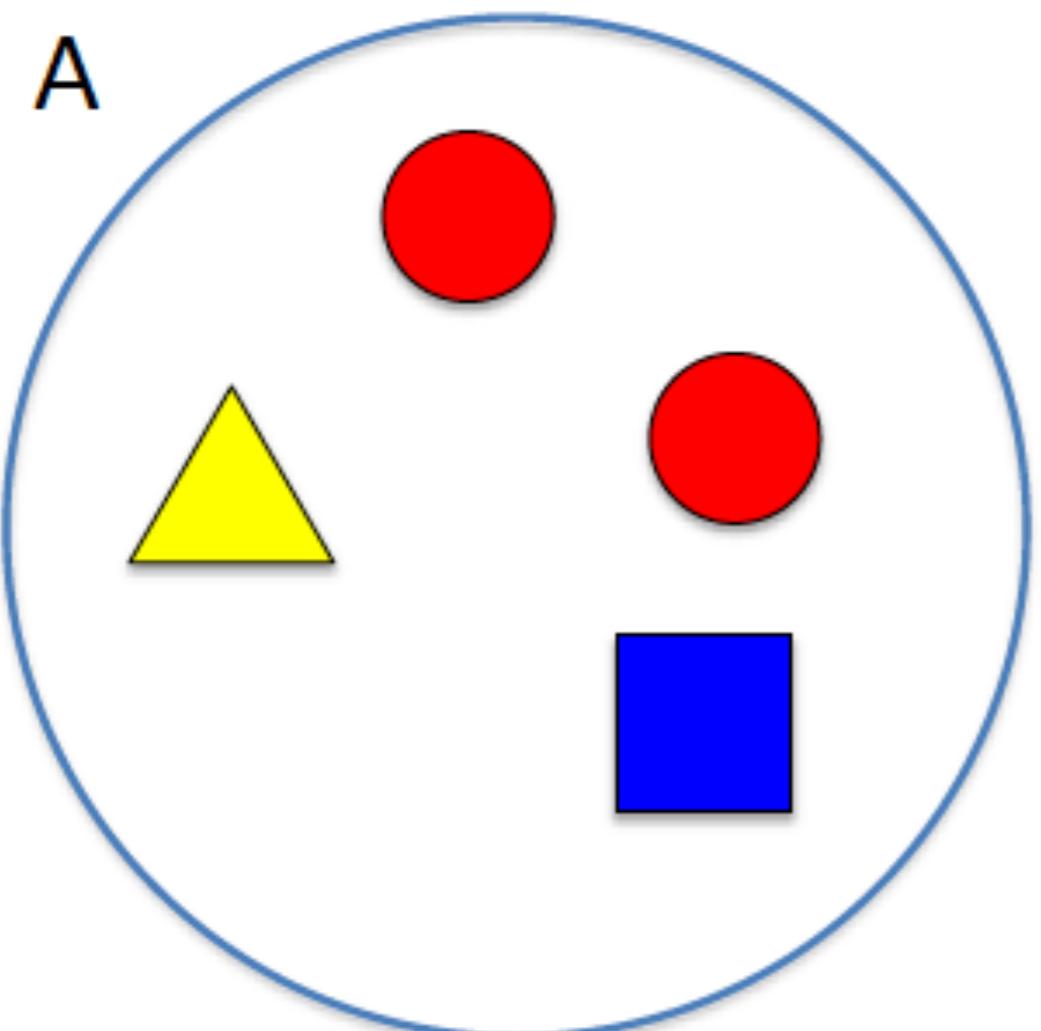


Alpha diversity = 4

Sample B is more diverse than sample A

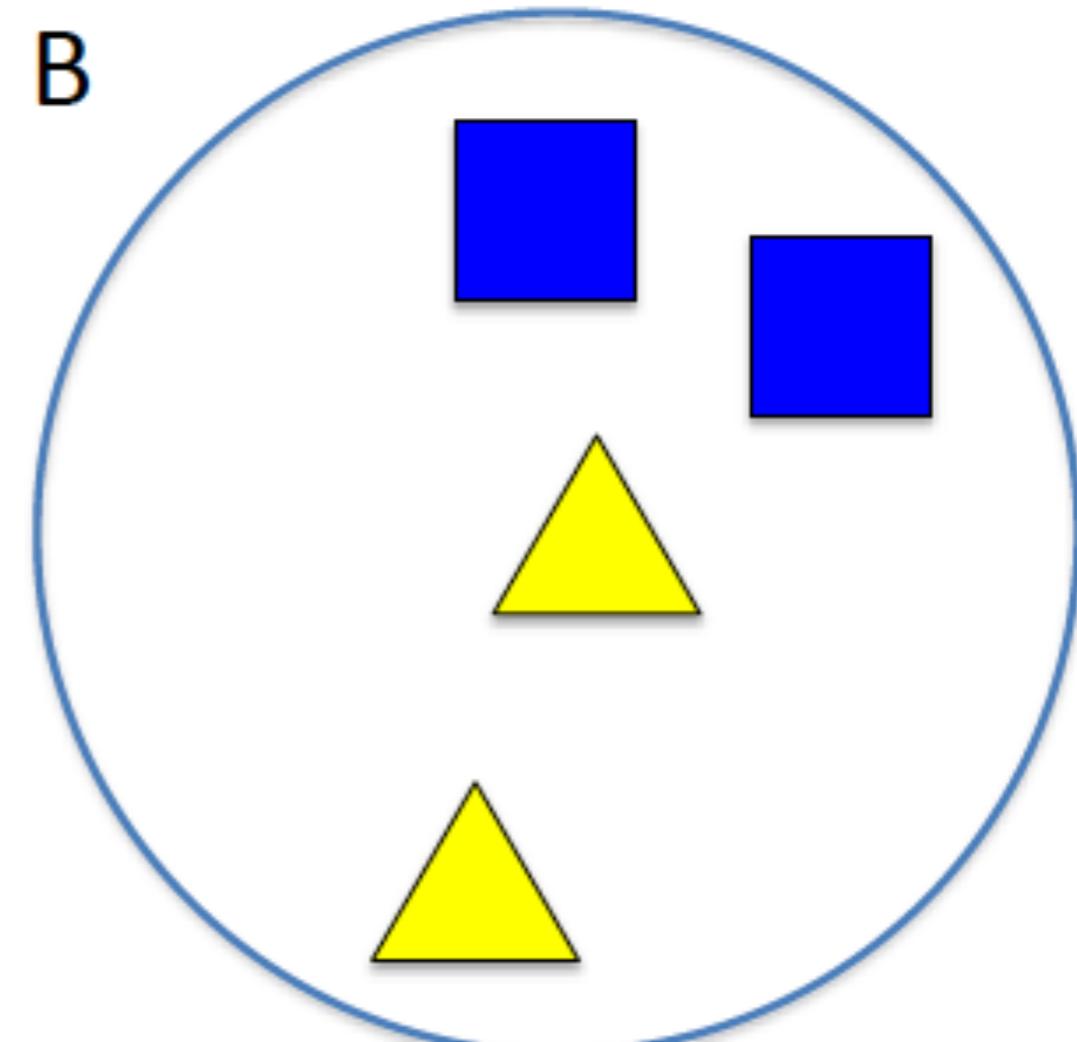
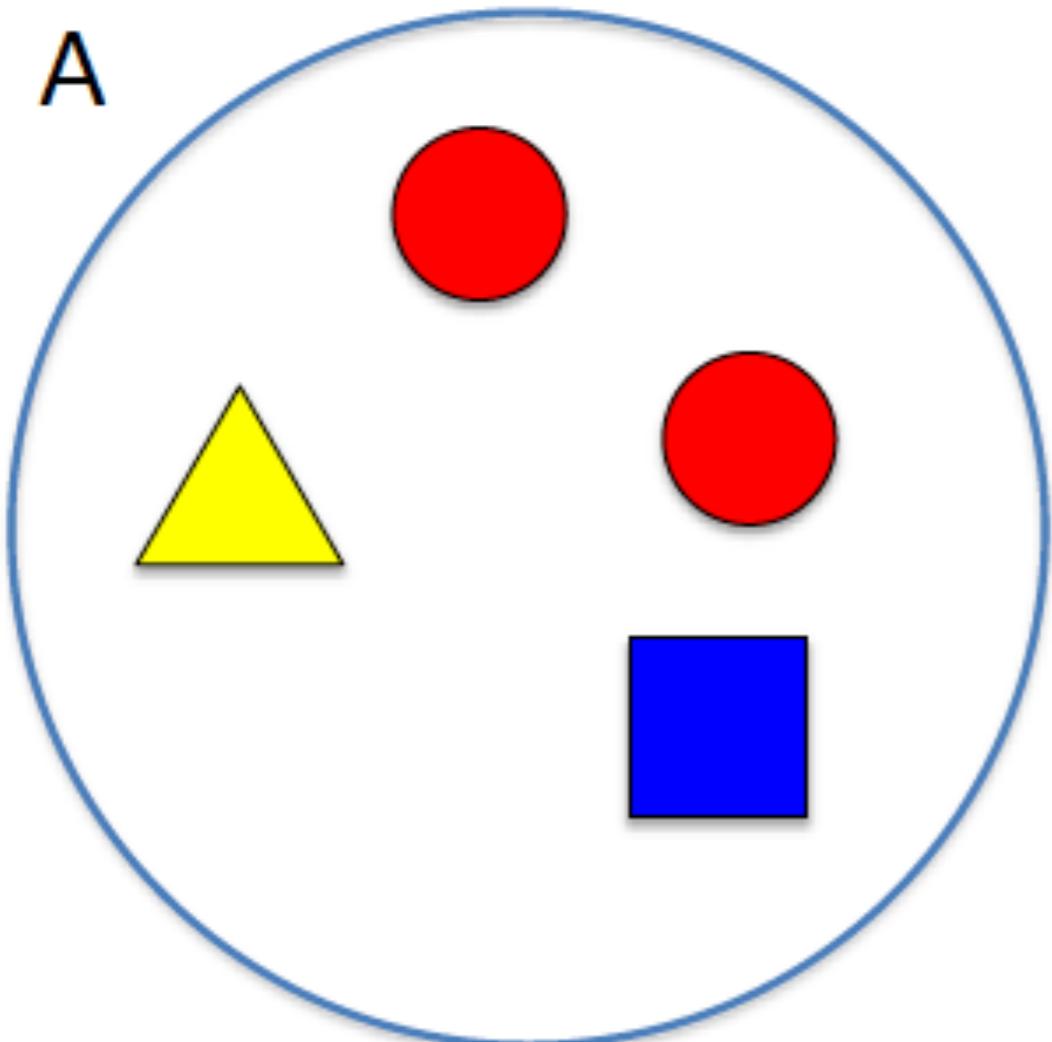
Ref: Daniel McDonald; American Gut Project

# Alpha diversity



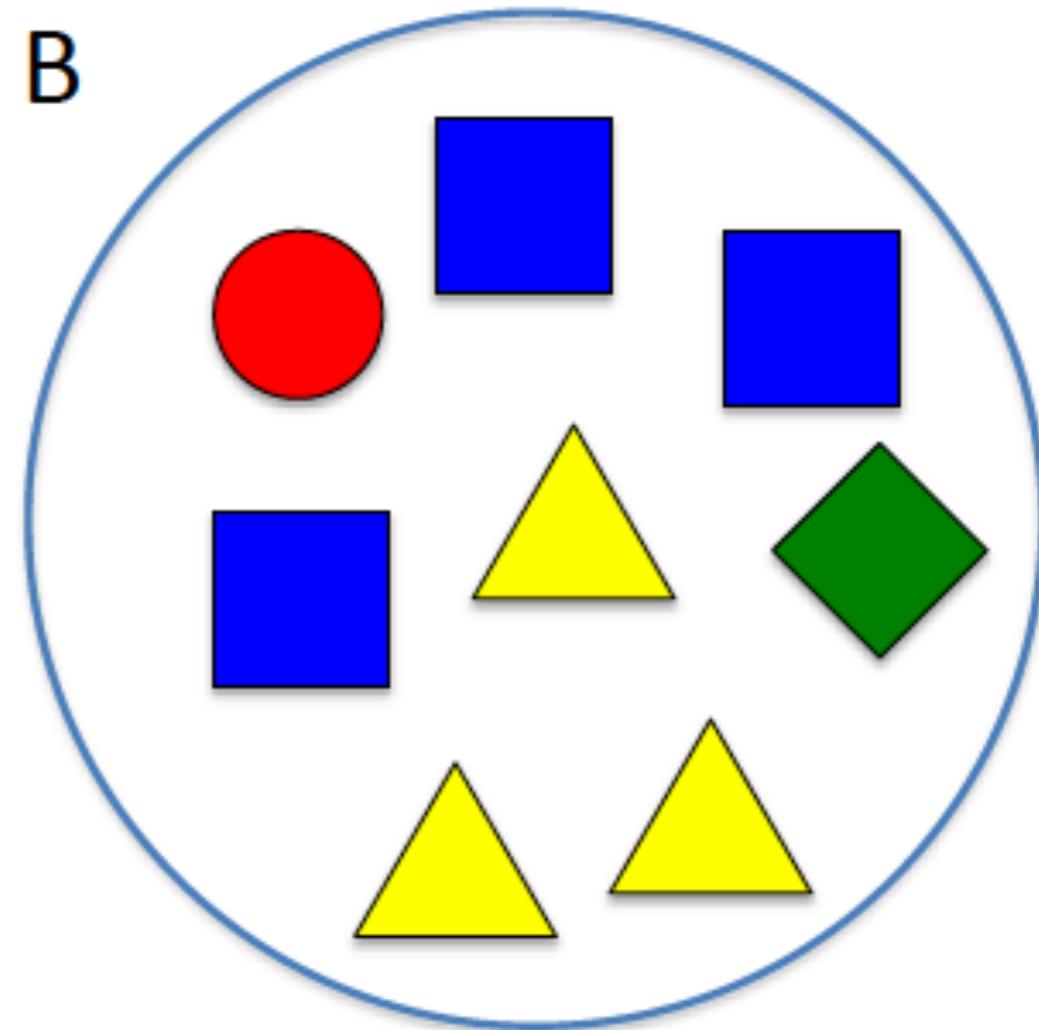
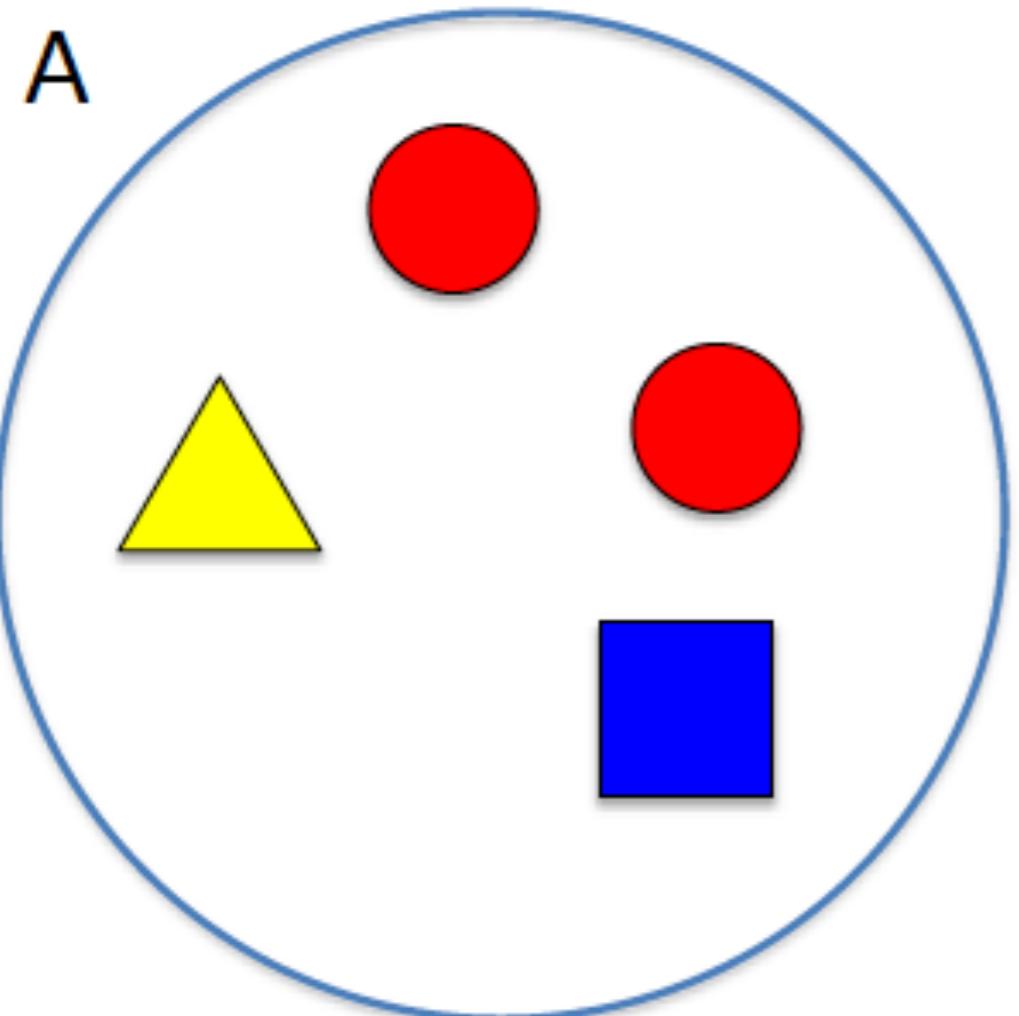
Ref: Daniel McDonald; American Gut Project

# Rarefaction



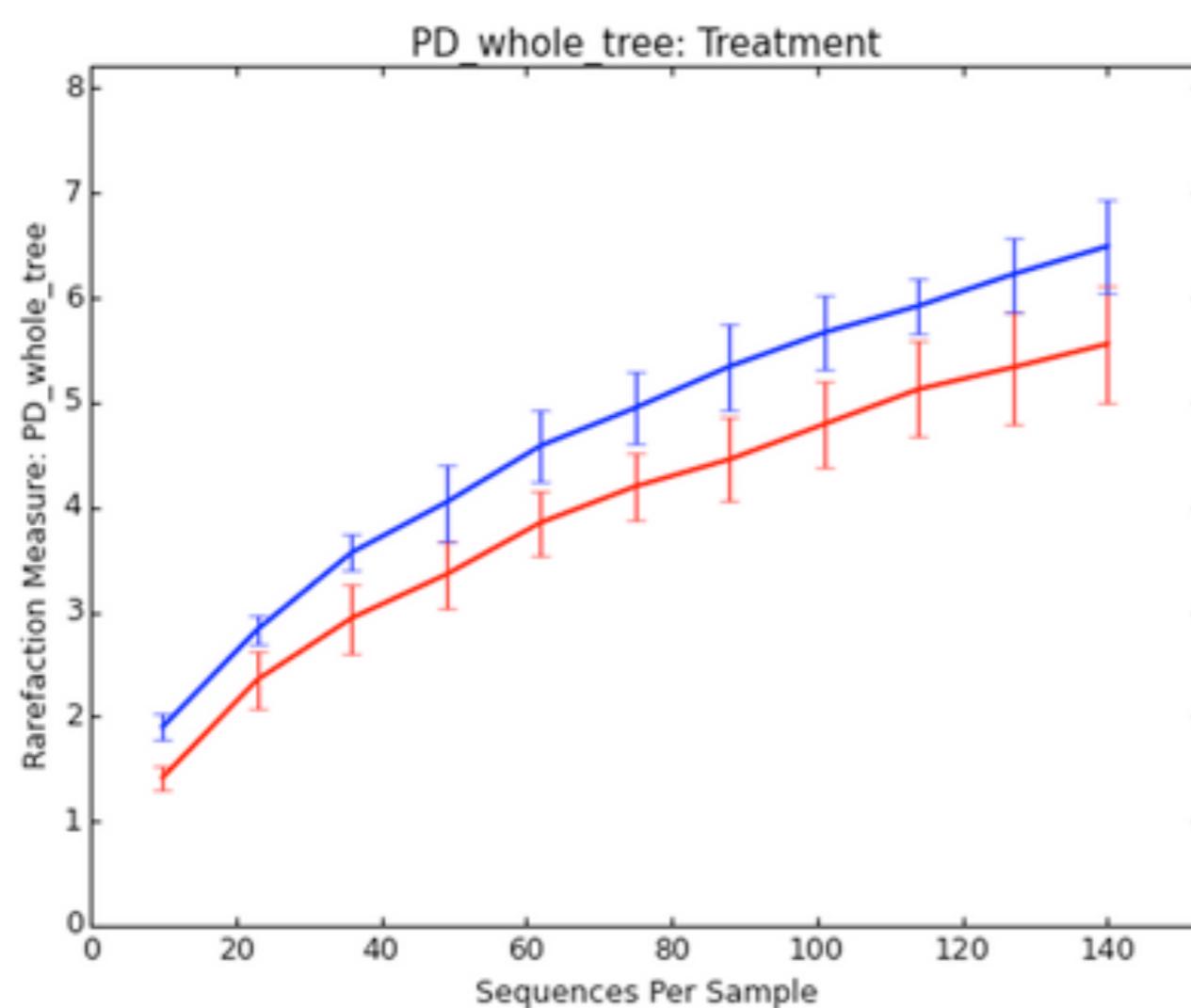
Ref: Daniel McDonald; American Gut Project

# Rarefaction



**Rarefy to 4 sequences**

Select a Metric: PD\_whole\_tree ▾ Select a Category: Treatment ▾



Show Categories: ▾

Legend

- Control
- Fast

# **Alpha Diversity Metrics**

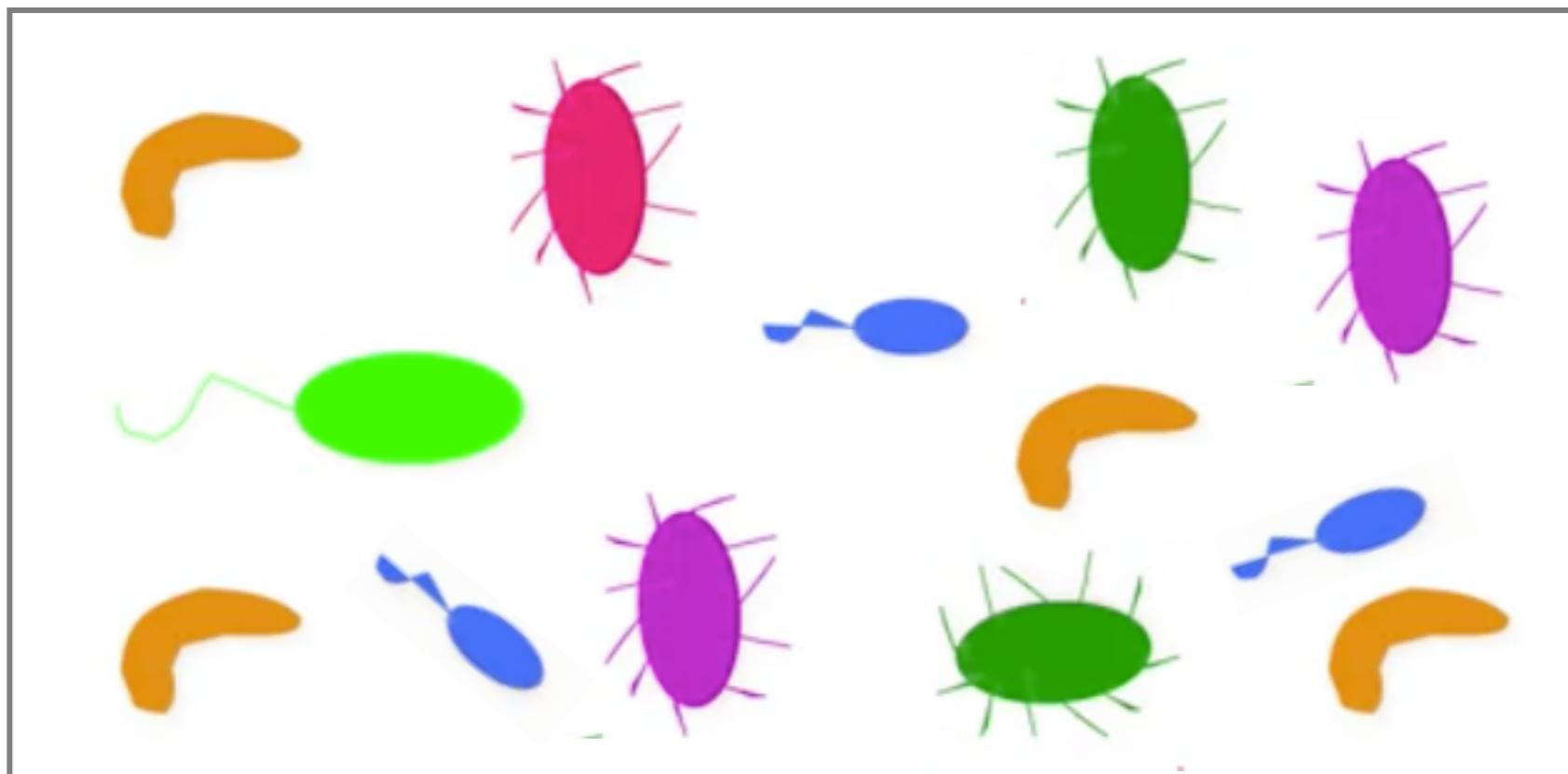
- Richness (e.g., Observed Species or PD Whole Tree)
- Evenness (e.g., Shannon's or Simpson's)



© E. K. Ganda

## $\alpha$ Diversity

# Within Sample Diversity



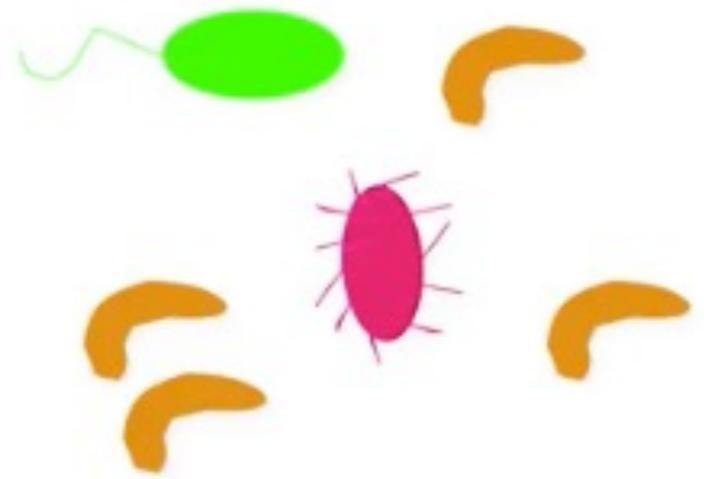
How diverse is this sample?



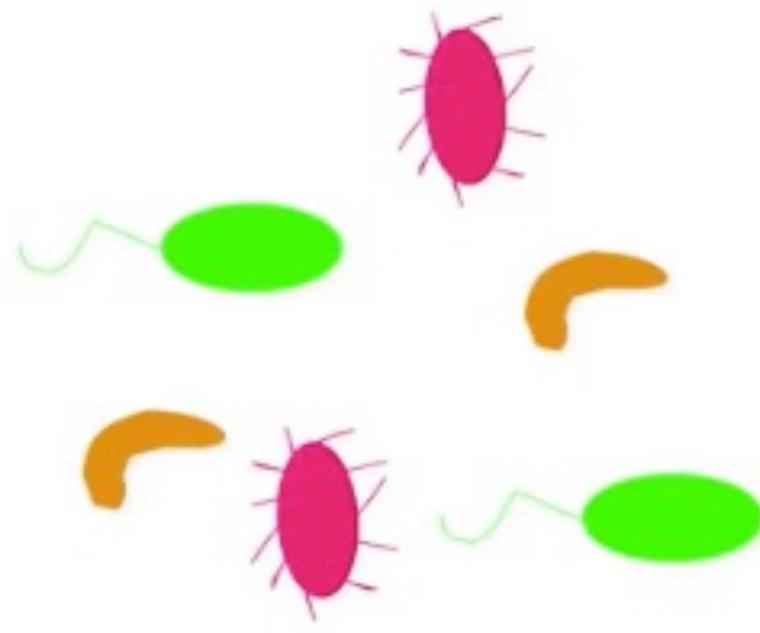
© E. K. Ganda

## $\alpha$ Diversity

# Richness



Community 1



Community 2

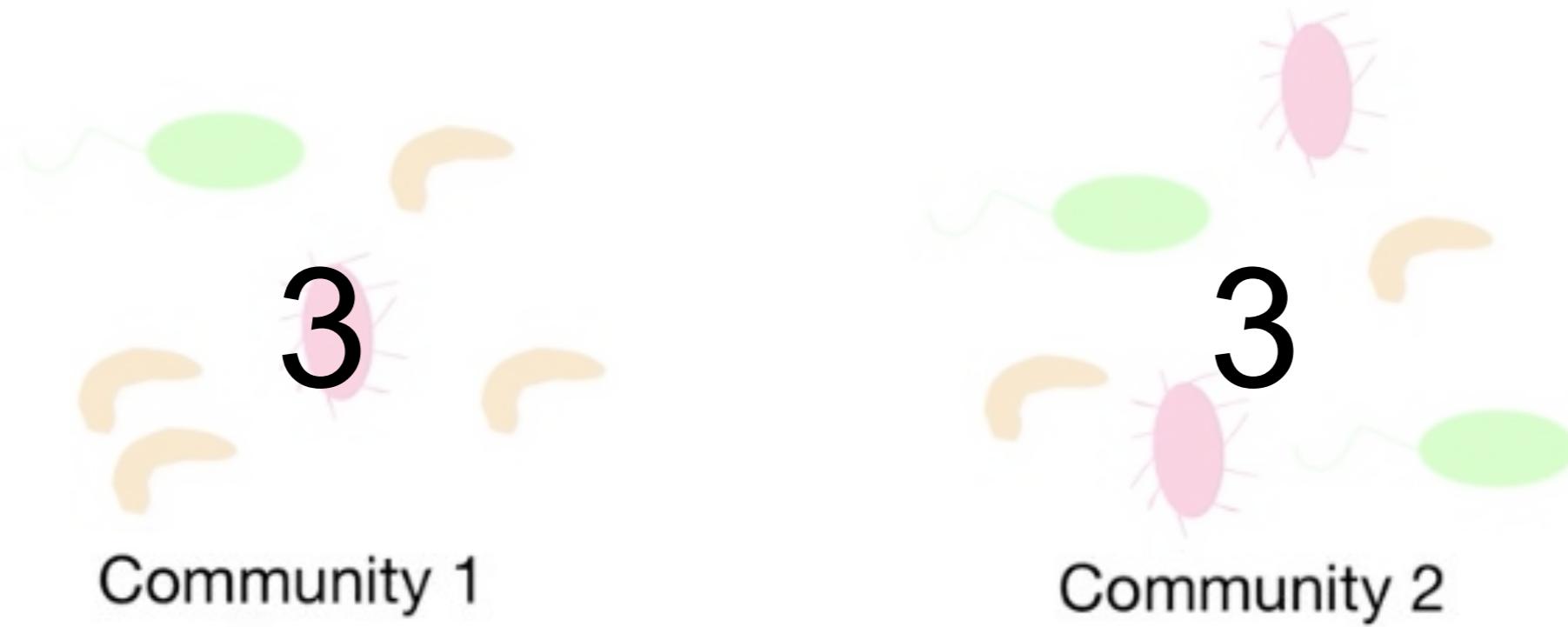
**How many kinds of microbes are there?**



© E. K. Ganda

$\alpha$  Diversity

Richness



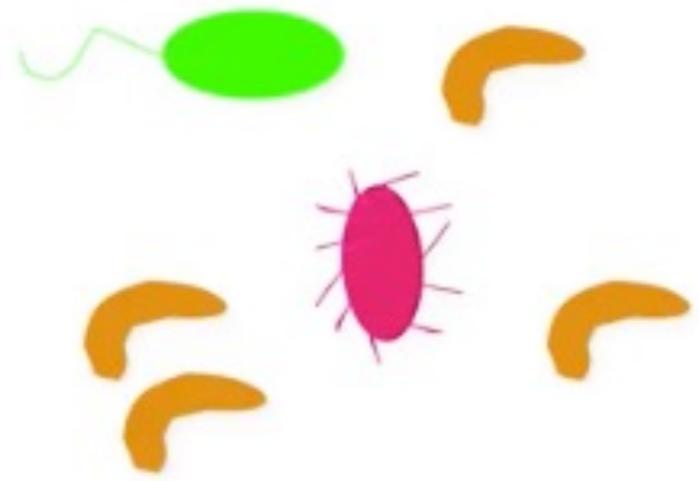
How many kinds of microbes are there?



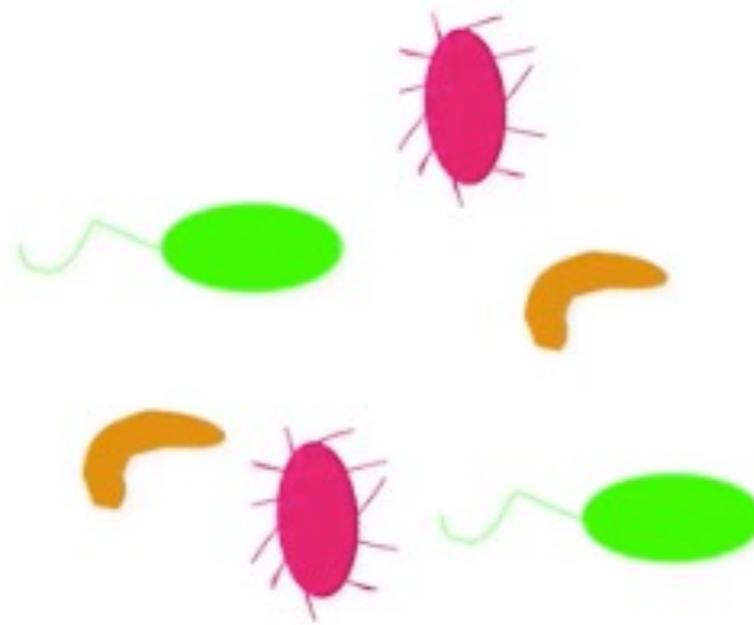
© E. K. Ganda

## $\alpha$ Diversity

## Evenness



Community 1



Community 2

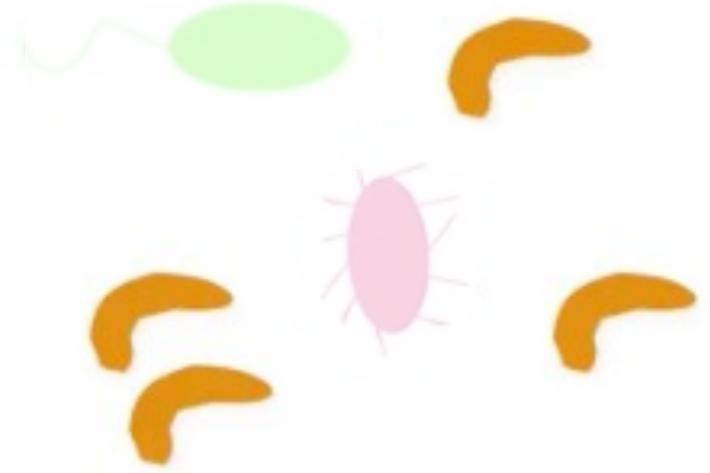
How are they distributed ?



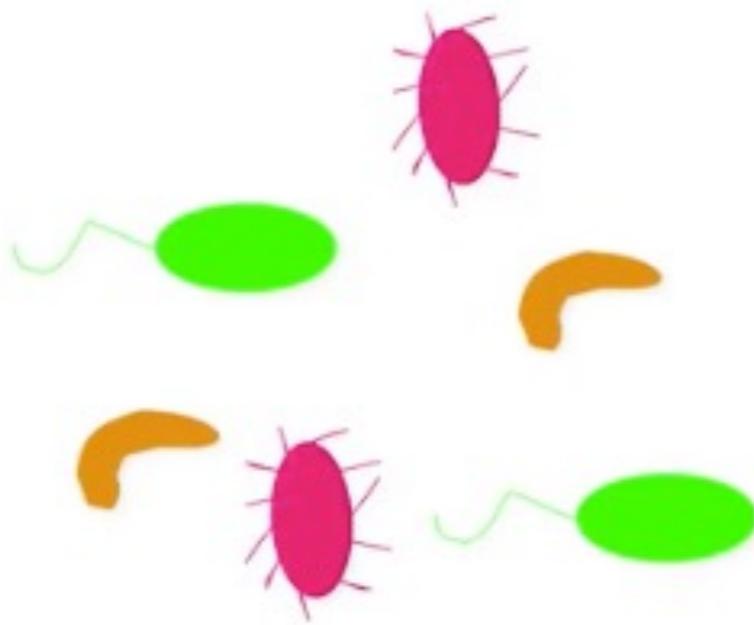
© E. K. Ganda

## $\alpha$ Diversity

## Evenness



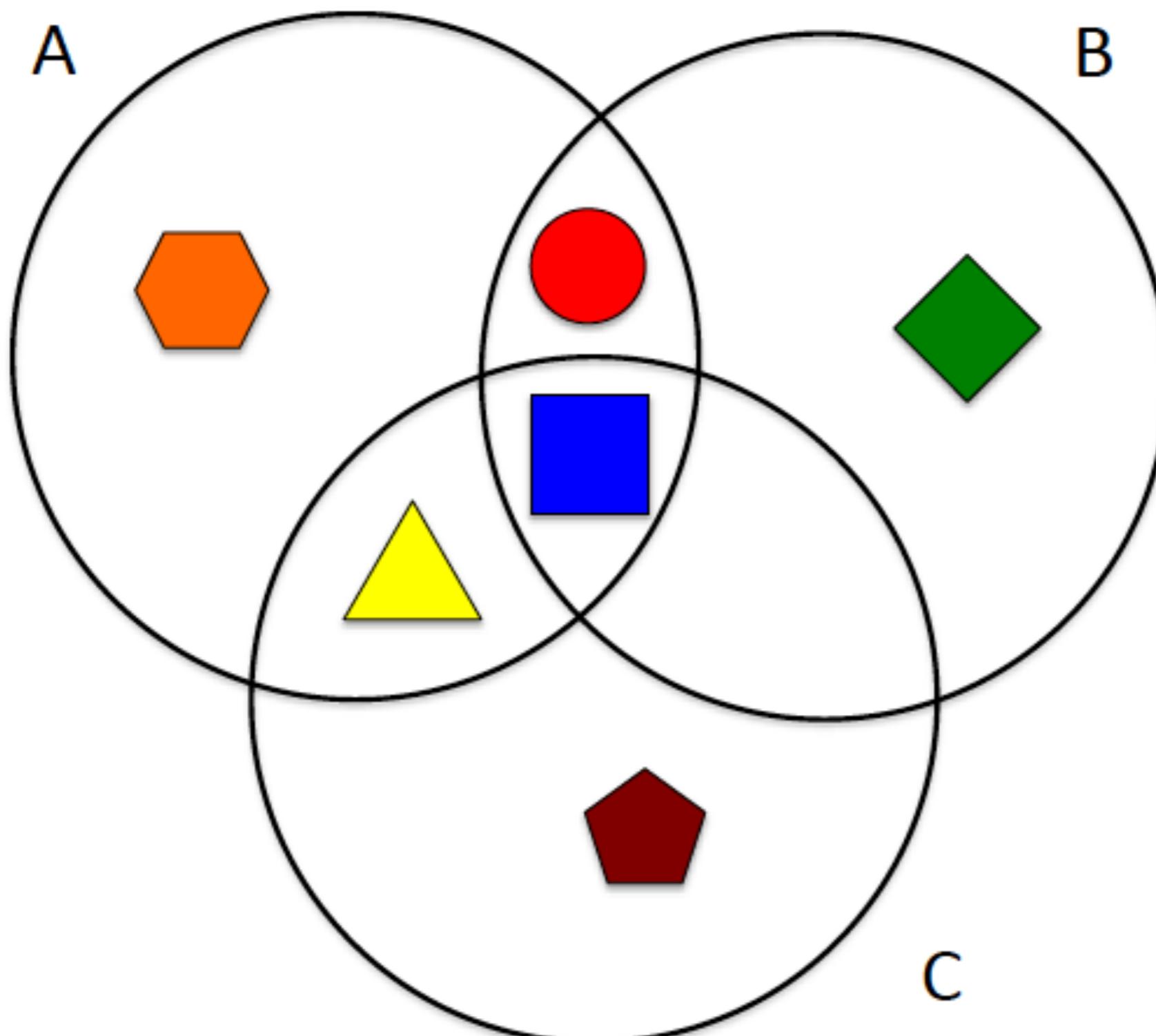
Community 1



Community 2

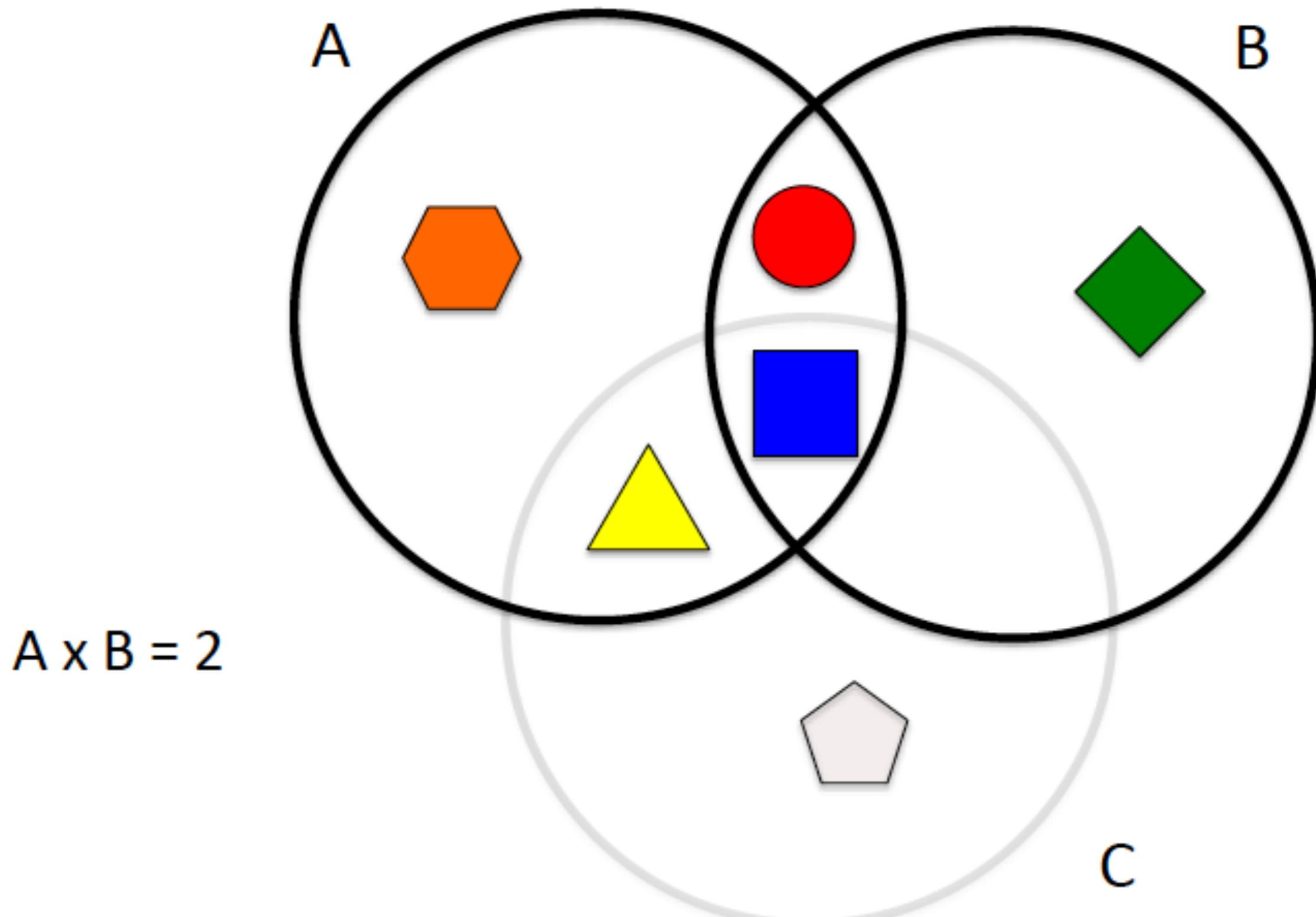
How are they distributed ?

# Beta diversity



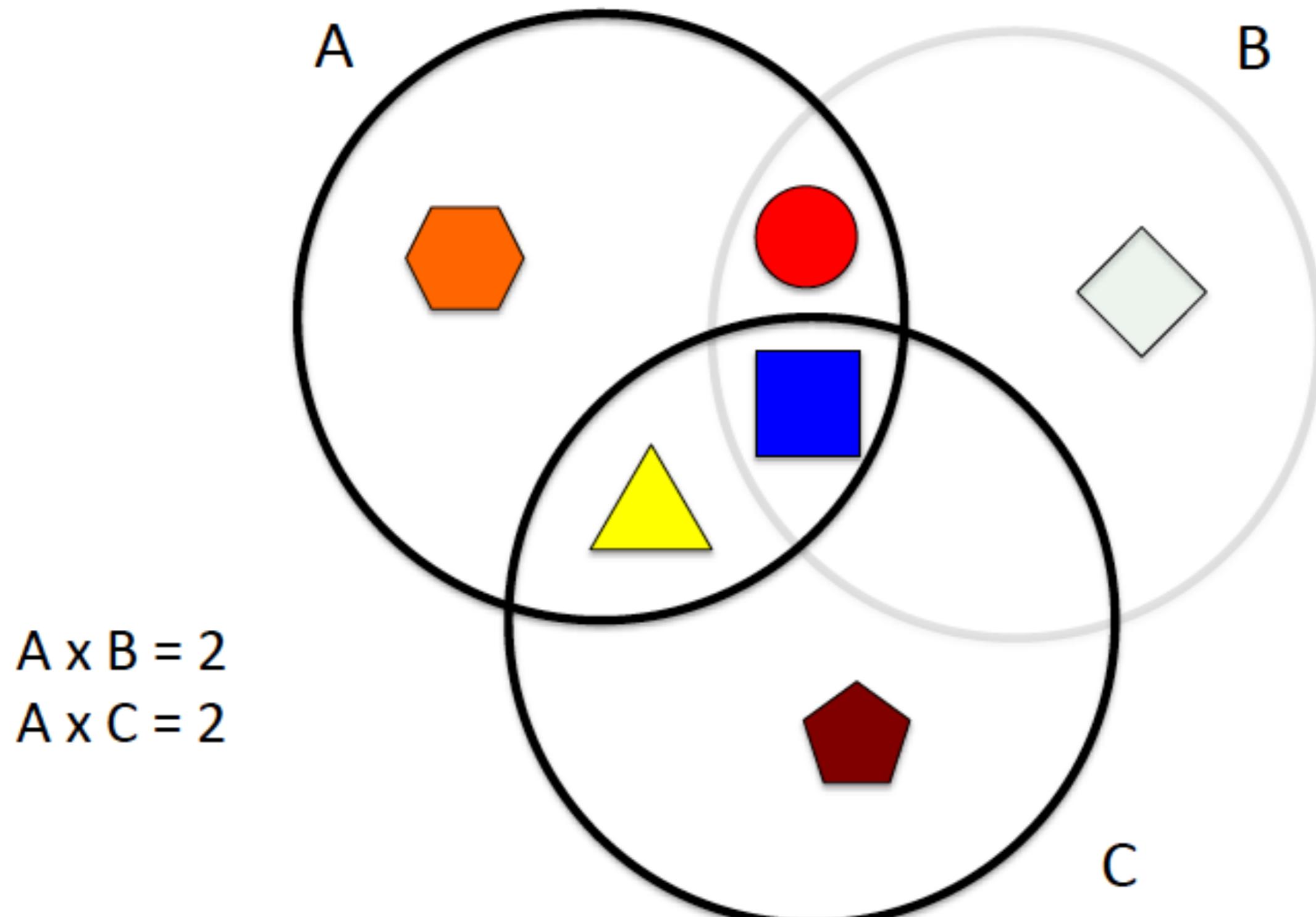
Ref: Daniel McDonald; American Gut Project

# Beta diversity



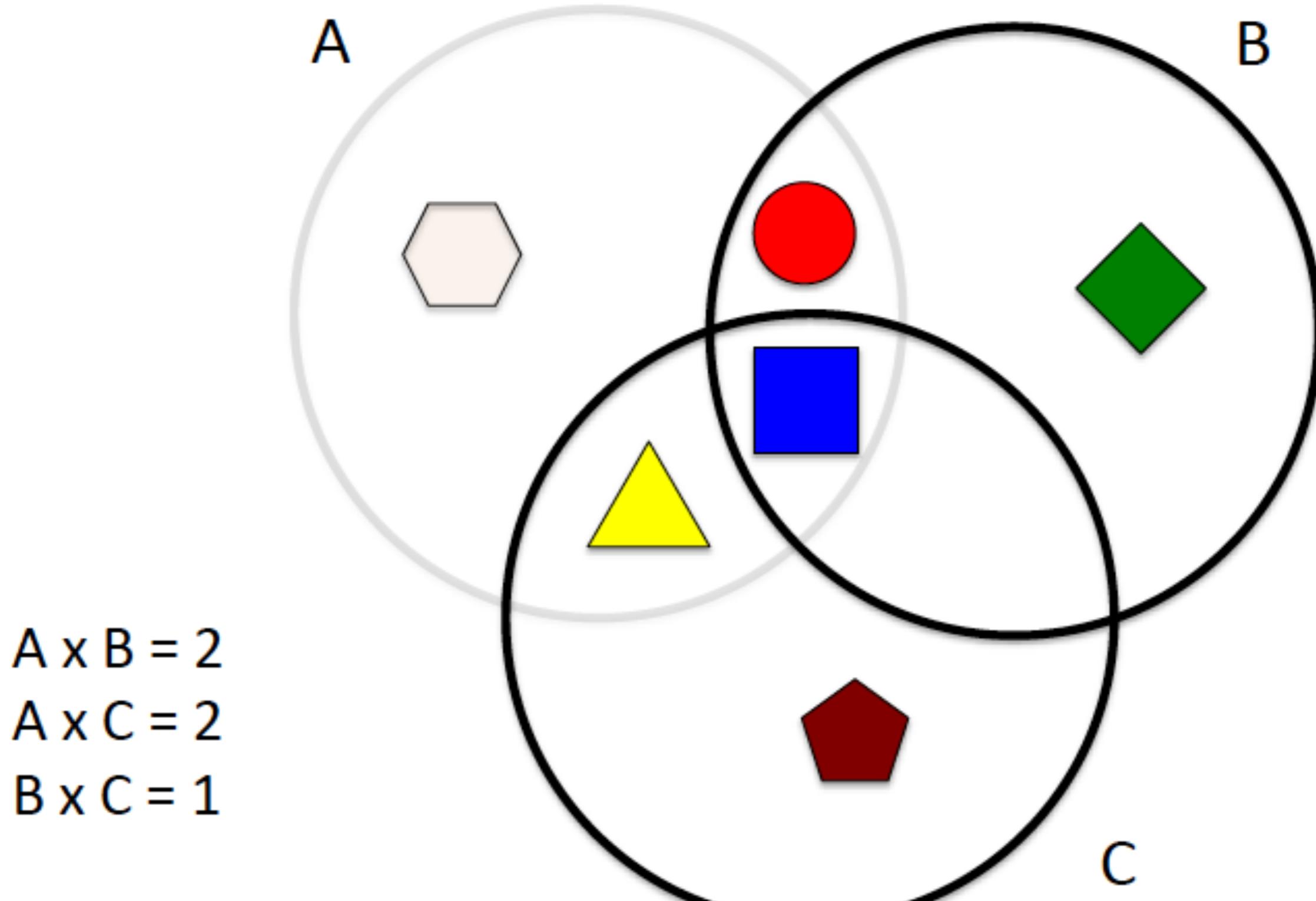
Ref: Daniel McDonald; American Gut Project

# Beta diversity

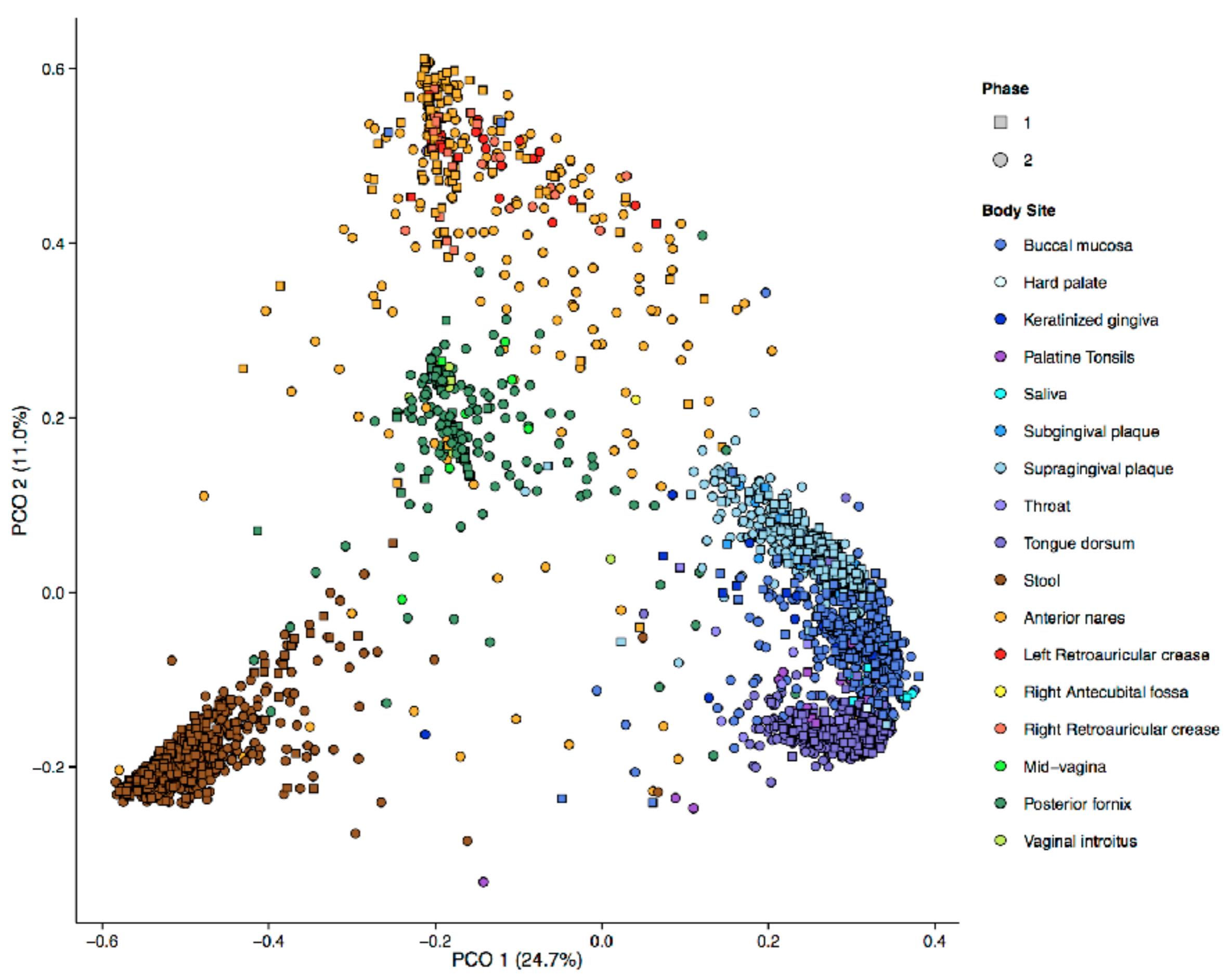


Ref: Daniel McDonald; American Gut Project

# Beta diversity



Ref: Daniel McDonald; American Gut Project



## Beta Diversity Metrics

- Shared composition (e.g., Jaccard)
- Distribution of composition (e.g., Bray Curtis)
- Phylogenetic composition (e.g., UniFrac)
- Compositionality of data (e.g., Aitchison's)



© E. K. Ganda

## $\beta$ Diversity

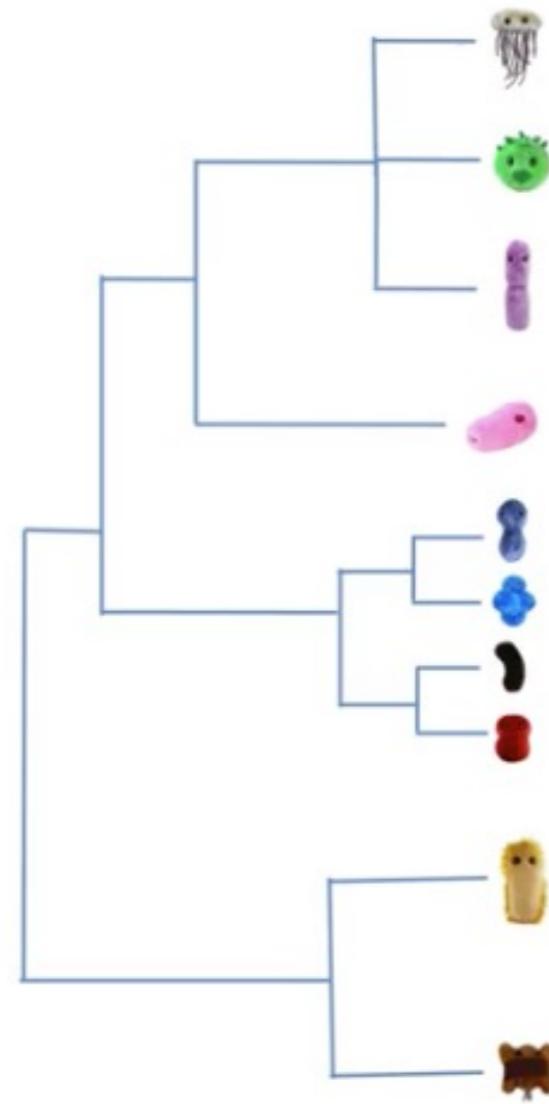
Sample 1



Sample 2



Sample 3



How are the microbes related?



© E. K. Ganda

## $\beta$ Diversity

Sample 1



Sample 2



Sample 3



**How are the microbes related?**



## $\beta$ Diversity



Sample 2



© E. K. Ganda



**How are the microbes related?**



© E. K. Ganda

## $\beta$ Diversity

Sample 1



Sample 2



APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Dec. 2005, p. 8228–8235  
0099-2240/05/\$08.00+0 doi:10.1128/AEM.71.12.8228-8235.2005  
Copyright © 2005, American Society for Microbiology. All Rights Reserved.

Vol. 71, No. 12

### UniFrac: a New Phylogenetic Method for Comparing Microbial Communities

Catherine Lozupone<sup>1</sup> and Rob Knight<sup>2\*</sup>

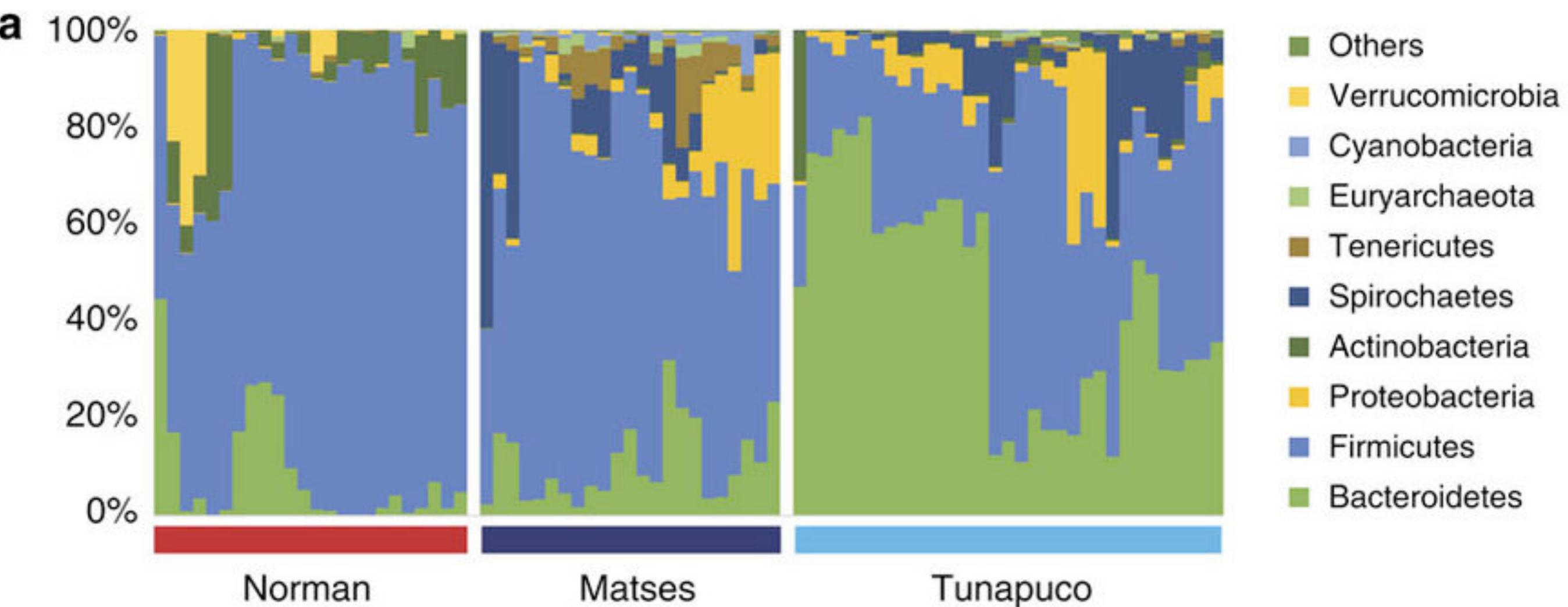
Department of Molecular, Cellular, and Developmental Biology, University of Colorado, Boulder,  
Colorado 80309,<sup>1</sup> and Department of Chemistry and Biochemistry, University of  
Colorado, Boulder, Colorado 80309<sup>2</sup>

Received 3 May 2005/Accepted 26 August 2005



# How are the microbes related?

# Summarise the taxa



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

# QIIME2 install

- <https://docs.qiime2.org/2018.11/install/native/>
  - (install miniconda first)
  - Install ‘Text Wrangler’ to record information

# QIIME2 Tutorials

- All tutorials here:
- <https://docs.qiime2.org/2018.11/tutorials/>
- Today's tutorial:
- <https://docs.qiime2.org/2018.11/tutorials/moving-pictures/>

# Questions?

1. Go to: **https://qiita.ucsd.edu/**
2. Sign up for an account  
Get info from your email  
Log In to QIITA

### **3.) Play Around**

Study -> View Studies

Search for studies of interest

What looks interesting to you?

### **4.) Select Samples for Study**

Analysis -> Create study

Search for 'Evolution'

Select these two studies by clicking on the green +:

'Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans'

Select the: (Reference phylogeny for SEPP: Greengenes\_13 BIOM: reference-hit.biom) | Trimming (length: 150) option.

## 5.) **Begin Analysis**

Click on Green file icon at top right

Create Analysis

Name your analysis

-Ants vs anteaters?

## 6.) **Examine Unique Features (DNA sequences)**

Click on (BIOM) table

What is the maximum number of sequences you find in a sample?

## **7.) Calculate alpha-diversity (within sample diversity - who has more species?)**

Click the (BIOM) circle.

Process -> Calculate alpha diversity -> Add Command  
-> Run

## 7.) Calculate alpha-diversity (within sample diversity - who has more species?)

Click the (BIOM) circle.

Process -> Calculate alpha diversity -> Add Command  
-> Run

Click (alpha\_vector)

Scroll through the Categories (e.g. 'host\_family' or  
'gut\_physiology' or 'diet')

Which species contain more microbial gut diversity?  
How might gut physiology contribute?

**7.) Calculate alpha-diversity  
(within sample diversity - who has more species?)**

**B. Correlate alpha diversity with a known numerical variable**

Select Alpha Vectors -> Process -> Calculate  
Alpha Correlation -> Add Command -> Run

Once loaded, select different categories.

Does alpha diversity correlate with specific variables?

## **8.) Calculate beta-diversity (between sample diversity; who is more similar)**

Click the (BIOM) circle.

Process -> Calculate beta diversity ->

Select 'Diversity metric' : Bray Curtis -> Add Command  
-> Run

Click on (distance\_matrix)

Process -> Perform Principle Coordinates Analysis ->

Add command -> Run

## **8.) Calculate beta-diversity (between sample diversity; who is more similar)**

Select biome table -> Process -> Select your biome table  
Select 'Diversity metric' : Bray Curtis -> Add Command  
-> Run

Click on (distance\_matrix)  
Process -> Perform Principle Coordinates Analysis ->  
Add command -> Run

Click on (ordination\_results), play with viewer

What drives the diversity in these samples?  
Are they linked to species diets?

## **8.) Calculate beta-diversity (between sample diversity; who is more similar)**

Click on (distance\_matrix)

Process -> Calculate Beta-group significance ->

Metadata: 'host\_order' or 'gut\_physiology' or 'diet'

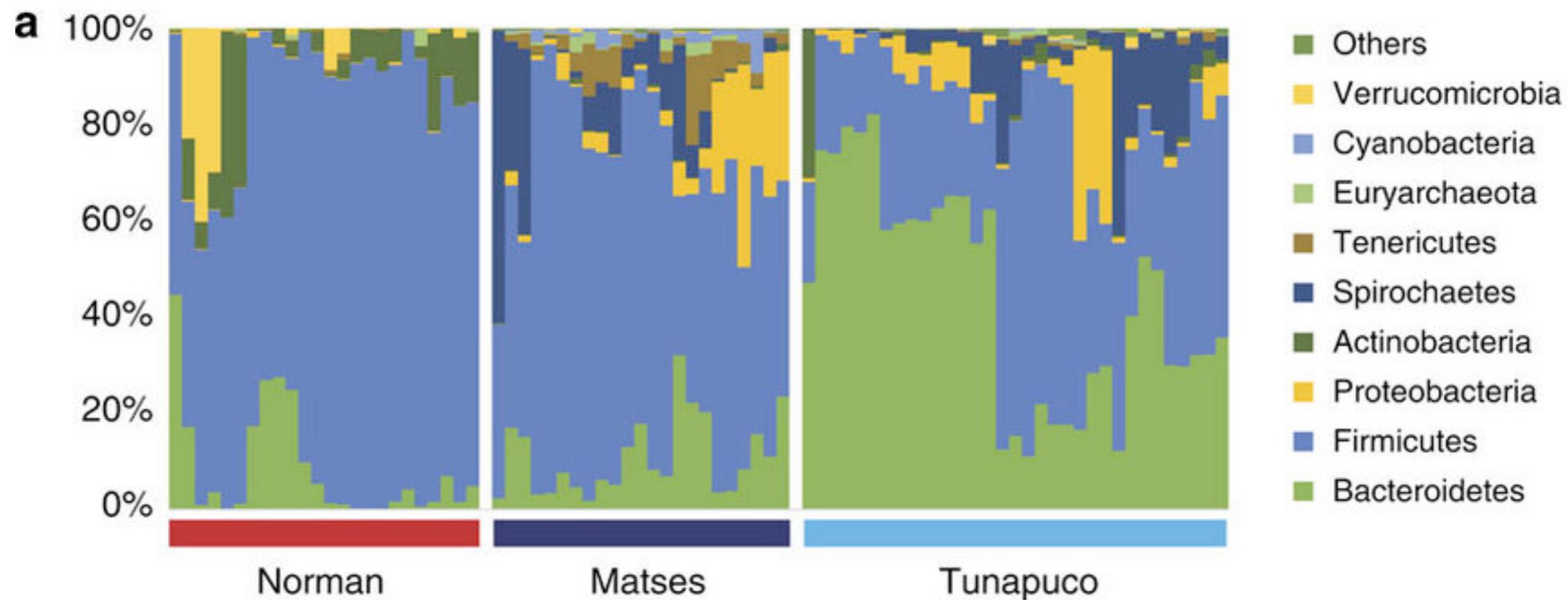
Add command -> Run

Are the factors driving diversity in the PCoA plot significantly different?

## 9.) Summarise the taxa

Click the (BIOM) circle.

Process -> Summarise Taxa -> Add Command -> Run



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

## 9.) Summarise the taxa

Click the (BIOM) circle.

Process -> Summarise Taxa -> Add Command -> Run

Taxonomic Level: Level 5

Sort Samples by: 'host\_family'

Which taxa are dominant in the Ursidae?

# Conclusions:

Diet and physiology drive microbial differences across the mammalian tree of life.

The screenshot shows the header of the Science journal website. At the top, there is a navigation bar with links for Home, News, Journals, Topics, and Careers. Below the navigation bar, there is a promotional banner for Jetstar with the text "All day, every day, low fares" and a "Book at jetstar.com" button. To the right of the banner, there is a search bar and a "Become a member" link. The main content area features the title of the article: "Diet Drives Convergence in Gut Microbiome Functions Across Mammalian Phylogeny and Within Humans" by Brian D. Muegge, Justin Kuczynski, Dan Knights, Jose C. Clemente, Antonio González, Luigi Fontana, Bernard Henrissat, Rob Knight, and Jeffrey I. Gordon. The article is dated May 20, 2011, and is from Volume 332, Issue 6032, pp. 470-474. The abstract, figures, and metrics sections are also visible.

SHARE REPORT



## Diet Drives Convergence in Gut Microbiome Functions Across Mammalian Phylogeny and Within Humans



Brian D. Muegge<sup>1</sup>, Justin Kuczynski<sup>2</sup>, Dan Knights<sup>3</sup>, Jose C. Clemente<sup>3</sup>, Antonio González<sup>3</sup>, Luigi Fontana<sup>4,5</sup>, Bernard Henrissat<sup>6</sup>, Rob Knight<sup>2,7</sup>, Jeffrey I. Gordon<sup>1,\*</sup>



<sup>1</sup>Center for Genome Sciences and Systems Biology, Washington University School of Medicine, St. Louis, MO 63110, USA.

<sup>2</sup>Department of Molecular, Cellular and Developmental Biology, University of Colorado, Boulder, CO 80309, USA.

<sup>3</sup>Department of Computer Science, University of Colorado, Boulder, CO 80309, USA.

<sup>4</sup>Department of Medicine, Washington University School of Medicine, St. Louis, MO 63110, USA.

<sup>5</sup>Division of Nutrition and Aging, Istituto Superiore di Sanità, Rome, Italy.

<sup>6</sup>Architecture et Fonction des Macromolécules Biologiques, CNRS and Aix-Marseille Université, Marseille, France.

<sup>7</sup>Howard Hughes Medical Institute, Boulder, CO, USA.

\*To whom correspondence should be addressed. E-mail: jgordon@washu.edu

- Hide authors and affiliations

Science, 20 May 2011,  
Vol. 332, Issue 6032, pp. 470-474  
DOI: 10.1126/science.1198719



Science

Vol 332, Issue 6032

20 May 2011

Table of Contents

Print Table of Contents

Advertising (PDF)

Classified (PDF)

Magazine (PDF)

### ARTICLE TOOLS



Email



Print



Alerts



Citation tools



Download Powerpoint



Save to my folders



Request Permissions



Share

Advertisement

Article

Figures & Data

Info & Metrics

eLetters

PDF

### Abstract

Coevolution of mammals and their gut microbiota has profoundly affected their radiation into myriad habitats. We used shotgun sequencing of microbial community DNA and targeted

Now - try your own analysis.

## 6.) **Examine Unique Features (DNA sequences)**

Click on (BIOM) table

What is the maximum number of sequences you find in a sample?

## 7.) **Calculate alpha-diversity**

Process -> Select your bio table -> Run

Which samples contain more diversity?

## 8.) **Calculate beta-diversity**

Process -> Calculate Beta-Diversity -> Run

Which samples are more similar to each other?

## 9.) **Summarise the taxa**

Process -> Summarise Taxa -> Run

Which taxa are shared across all samples?