

BINF*6210 Software tools

Microbiome

Saira Asif

Content

- **Intro to microbiome**
 - Why study the microbiome
 - Papers to review
- **Microbiome analysis**
 - Sample collection
 - DNA extraction
 - Sequencing
 - Data preprocessing
 - Analysis

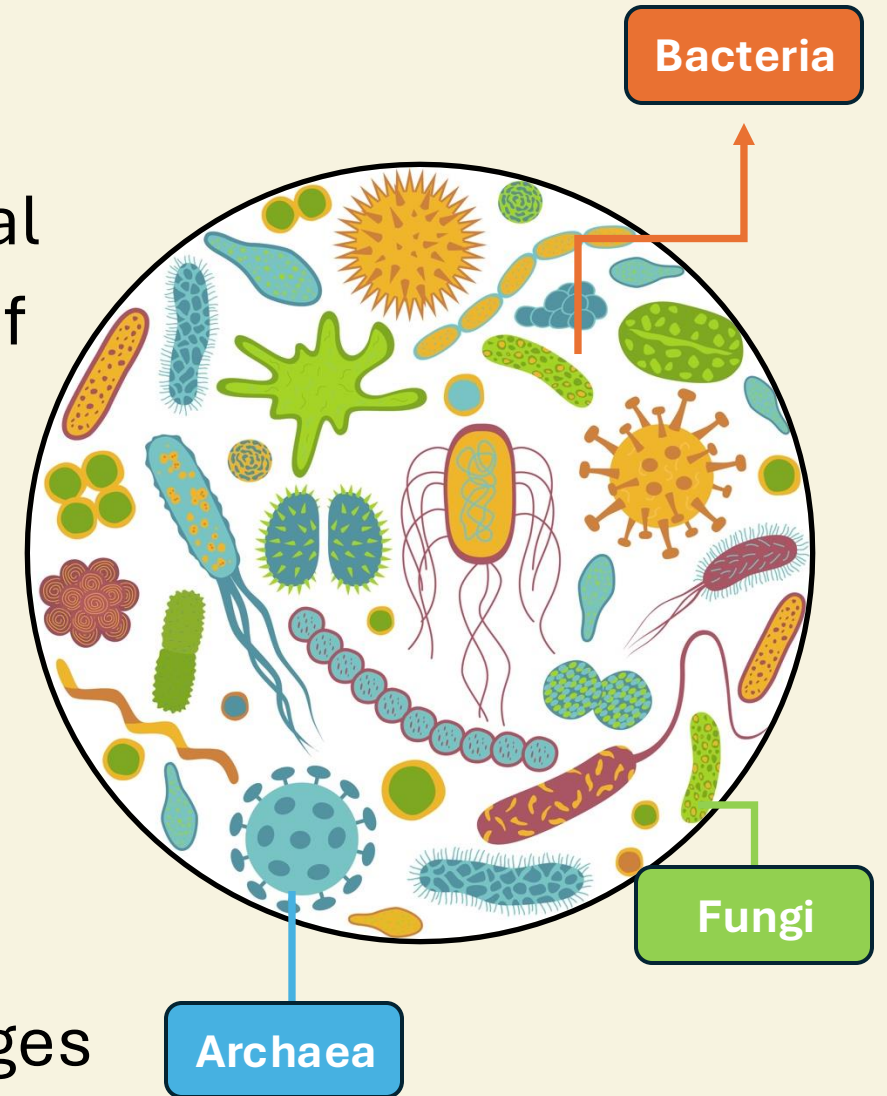
Introduction to the Microbiome

Microbiome comprise all the genetic material within a microbiota. Refers to a community of microorganisms that exist in an environment.

- **Bacteria** (90 - 95%)
- **Viruses** (5 - 10 %)
- **Fungi** (< 1%)

Found in **gut, skin, soil, oceans**

- Important research area at U. Guelph, spanning multiple departments and colleges



Why study the microbiome?



Digestion

Breaks down food and absorbing nutrients



Immunity

Trains immune cells, prevent over-inflammation

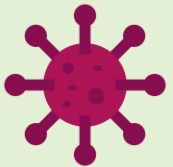
- Commensal vs pathogenic bacteria



Mental health

Gut-brain axis regulates mental health

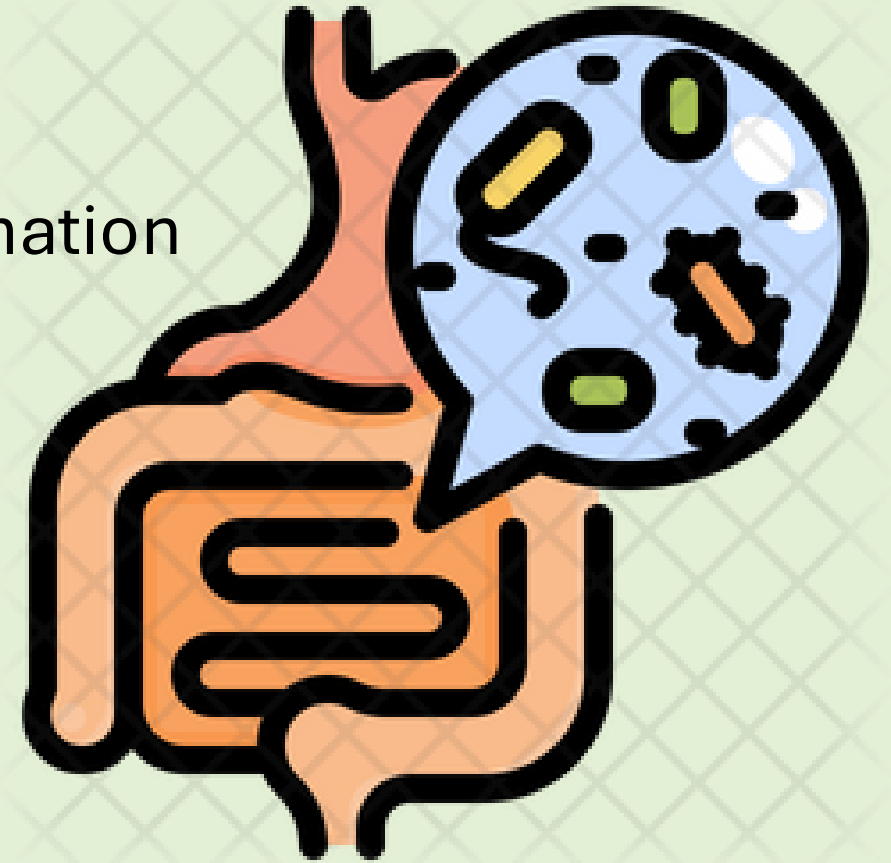
- Influences, mood, motivations, drives



Links to diseases

Obesity, IBD, diabetes etc.

- Nutrient absorption, hormones, appetite etc.
- Bacterial shifts (gut dysbiosis)



Microbiome composition

Microbiome = **dynamic** ecosystem that **changes** with:

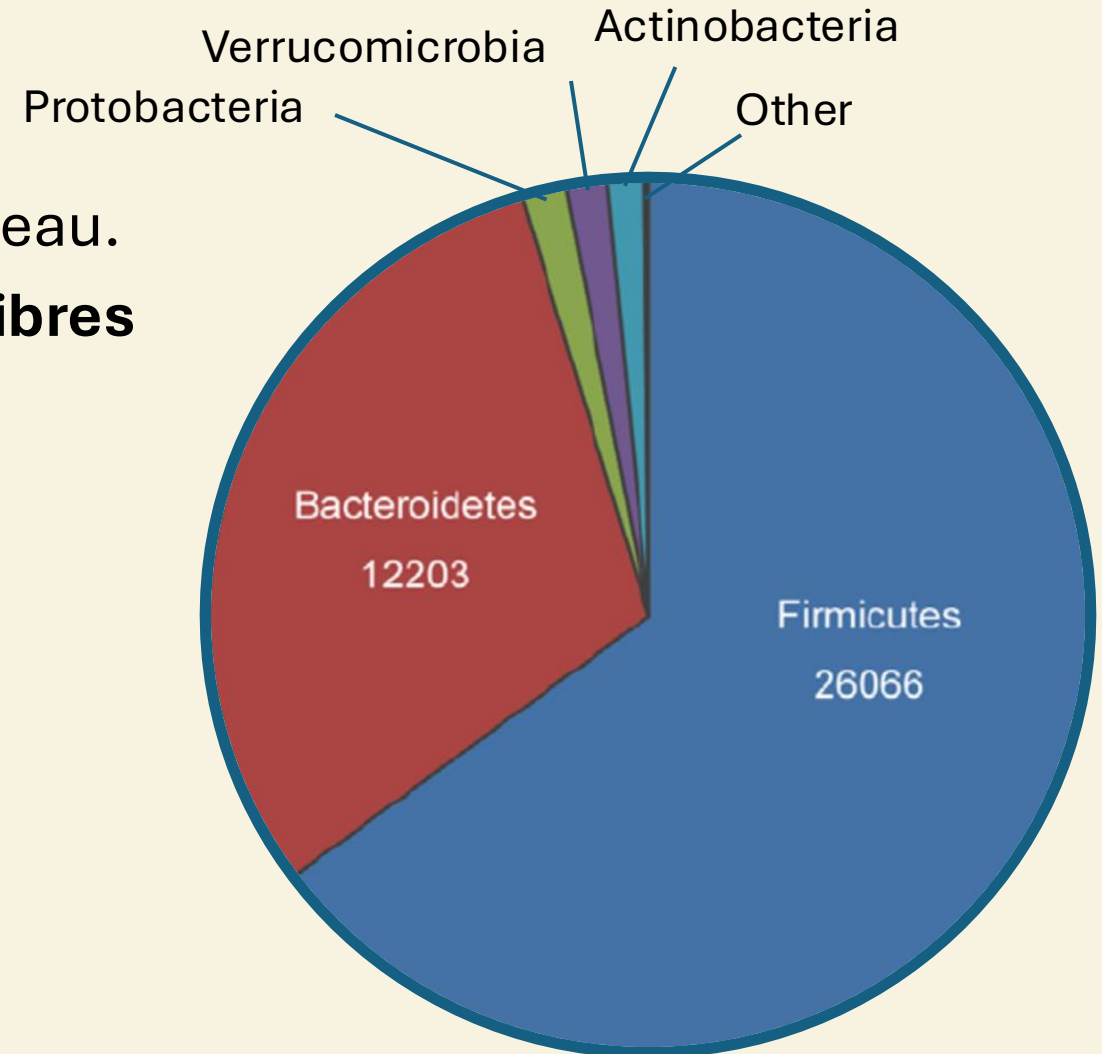
- Age: High diversity at birth, followed by plateau.
- Diet: Firmicutes = **carbs**, Bacteroidetes = **fibres**
- Antibiotics: Declines in diversity
- Disease: IBD, Chrones, Colitis

majority of your gut bacteria gets established with vaginal contact during birth

Healthy human gut

- ~**1000** species of bacteria
- **Dominant phyla:** Firmicutes and Bacteroidetes

throwing off this ratio makes it easier for pathogens to enter the system - it becomes a feedback loop where it continues to get worse.






throwing off this ratio provides an opportunity for p;

ADDENDUM



Links between environment, diet, and the hunter-gatherer microbiome

Gabriela K. Fragiadakis ^a, Samuel A. Smits^a, Erica D. Sonnenburg^a, William Van Treuren^a, Gregor Reid ^b, Rob Knight^c, Alphaxard Manjurano^d, John Chagalucha^e, Maria Gloria Dominguez-Bello^f, Jeff Leach ^{g,h}, and Justin L. Sonnenburg^{a,i}

The study of traditional populations provides a view of human-associated microbes unperturbed by industrialization, as well as a window into the microbiota that co-evolved with humans. Here we discuss our recent work characterizing the microbiota from the Hadza hunter-gatherers of Tanzania. We found **seasonal shifts in bacterial taxa, diversity, and carbohydrate utilization by the microbiota**. When compared to the microbiota composition from other populations around the world, **the Hadza microbiota shares bacterial families with other traditional societies that are rare or absent from microbiotas of industrialized nations**. We present additional observations from the Hadza microbiota and their lifestyle and environment, including microbes detected on hands, water, and animal sources, how the microbiota varies with sex and age, and the short-term effects of introducing agricultural products into the diet. In the context of our previously published findings and of these additional observations, we discuss a path forward for future work.

A microbiome-dependent gut–brain pathway regulates motivation for exercise

Exercise exerts a wide range of beneficial effects for healthy physiology. However, the mechanisms regulating an individual's motivation to engage in physical activity remain incompletely understood. An important factor stimulating the engagement in both competitive and recreational exercise is the motivating pleasure derived from prolonged physical activity, which is triggered by exercise-induced neurochemical changes in the brain. Here, we report on the discovery of **a gut–brain connection in mice that enhances exercise performance by augmenting dopamine signalling during physical activity**. We find that microbiome-dependent production of endocannabinoid metabolites in the gut stimulates the activity of TRPV1-expressing sensory neurons and thereby elevates dopamine levels in the ventral striatum during exercise. Stimulation of this pathway improves running performance, whereas microbiome depletion, peripheral endocannabinoid receptor inhibition, ablation of spinal afferent neurons or dopamine blockade abrogate exercise capacity. These findings indicate that the rewarding properties of exercise are influenced by gut-derived interoceptive circuits and provide a microbiome-dependent explanation for interindividual variability in exercise performance. Our study also suggests that interoceptomimetic molecules that stimulate the transmission of gut-derived signals to the brain may enhance the motivation for exercise.

THE MICROBIOME IN AUTISM SPECTRUM DISORDER

The human gut microbiota with reference to autism spectrum disorder: considering the whole as more than a sum of its parts

Michael C. Toh and Emma Allen-Vercoe*

Department of Molecular and Cellular Biology, University of Guelph, Guelph, Ontario, Canada

The human gut microbiota is a complex microbial ecosystem that contributes an important component towards the health of its host. This highly complex ecosystem has been underestimated in its importance until recently, when a realization of the enormous scope of gut microbiota function has been (and continues to be) revealed. One of the more striking of these discoveries is the finding that **the gut microbiota and the brain are connected**, and thus there is potential for the microbiota in the gut to **influence behavior and mental health**. In this short review, we outline the link between brain and gut microbiota and urge the reader to consider the gut microbiota as an ecosystem 'organ' rather than just as a collection of microbes filling a niche, using the hypothesized role of the gut microbiota in autism spectrum disorder to illustrate the concept

Overview of Microbiome analysis

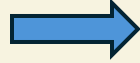


**Sample
collection**

Swabs,
soil/water etc.

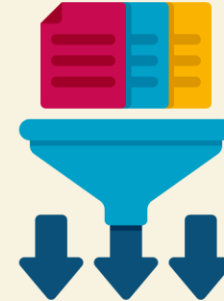
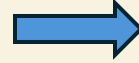


**DNA
extraction**



Sequencing

16S, shotgun
metagenomic etc.



**Data
preprocessing**

QC, filtering, host
DNA removal,
taxonomic/functional
classification



Analysis

Taxonomic composition
Diversity analysis
Functional analysis

Sample collection

Type of Sample

- Gut: Stool sample
- Oral: Saliva or swabs
- Skin: Swabs
- Environmental: Soil/water

Collection guidelines

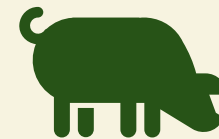
- Avoid contamination
- Store on ice after collection, freeze at -80 C for storage.

Farzan lab project



Farms (N = 20)

- Outbreaks of diarrhea (PWD)



Pigs (n = 300)

- 10 diarrheic
- 5 non-diarrheic



Samples

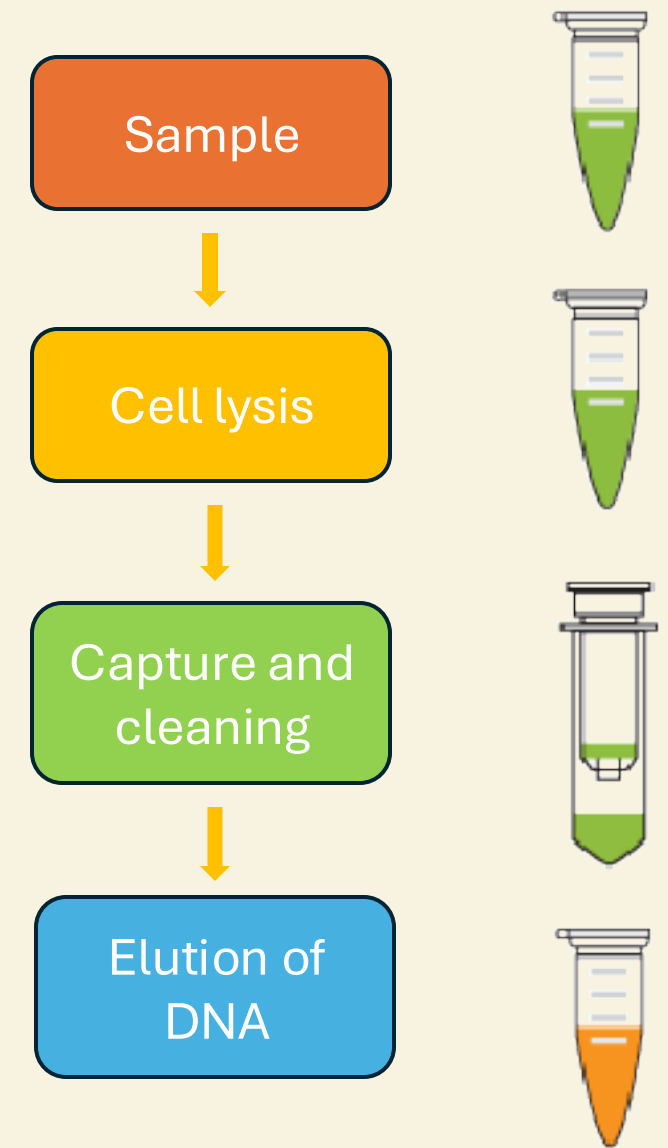
- Fecal swabs
- Blood samples

DNA extraction

Purpose: Isolate high-quality microbial DNA for sequencing.

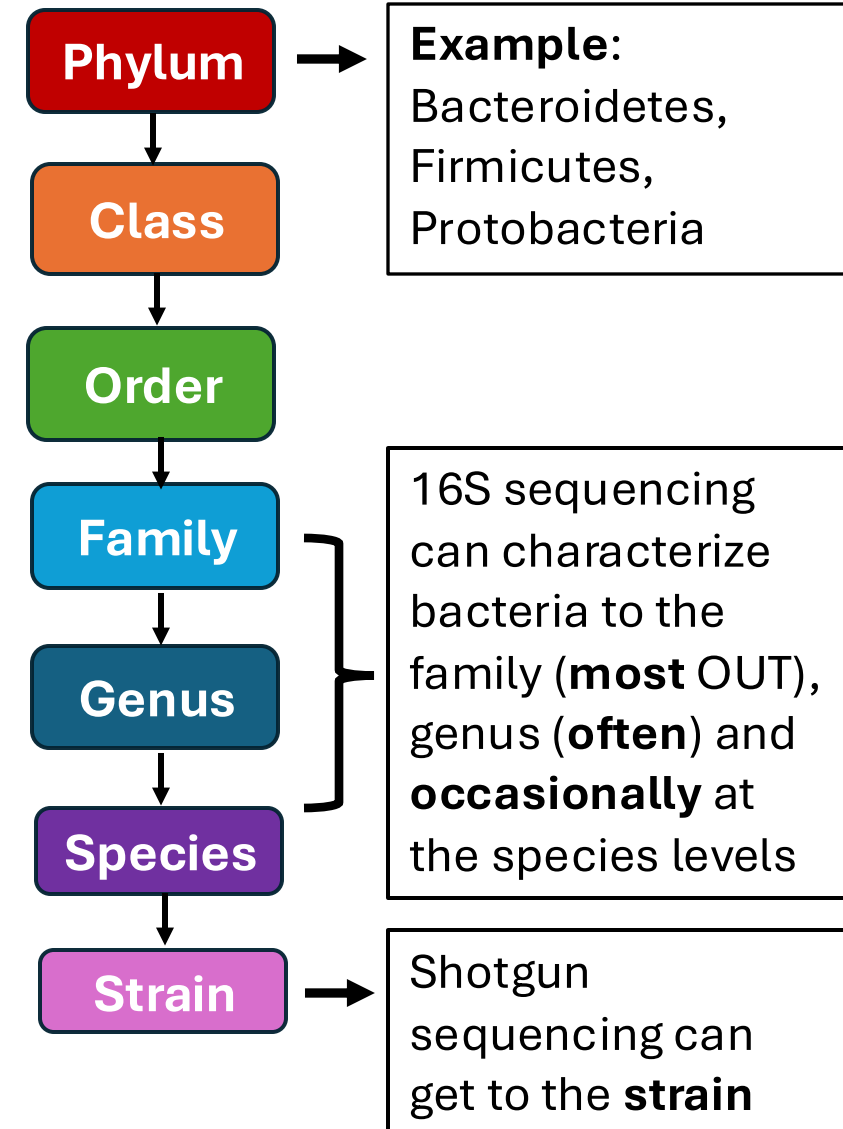
Extraction kits: ZymoBIOMICS MagBead DNA

- Feces, soil, water, biofilms, swabs, saliva, and body fluids



Sequencing

Method	Target	Pros	Cons	Software/packages
16S rRNA amplicon sequencing	Bacteria or archaea	Cost-effective, established pipelines	Lower resolution	DADA2, QUIIME, Mother
Shotgun metagenomic	All DNA	Taxonomic + functional insights	Higher costs	metagenomeSeq, phyloseq, Microbiome
Meta transcriptomics	RNA	Functional activity	RNA is more unstable	DESeq2, edgeR, phyloseq



Data pre-processing

1

Quality assessment

- QC score
- GC content

FastQC

2

Trimming & adapter removal

- Low-quality bases
- adapters

Trimmomatic

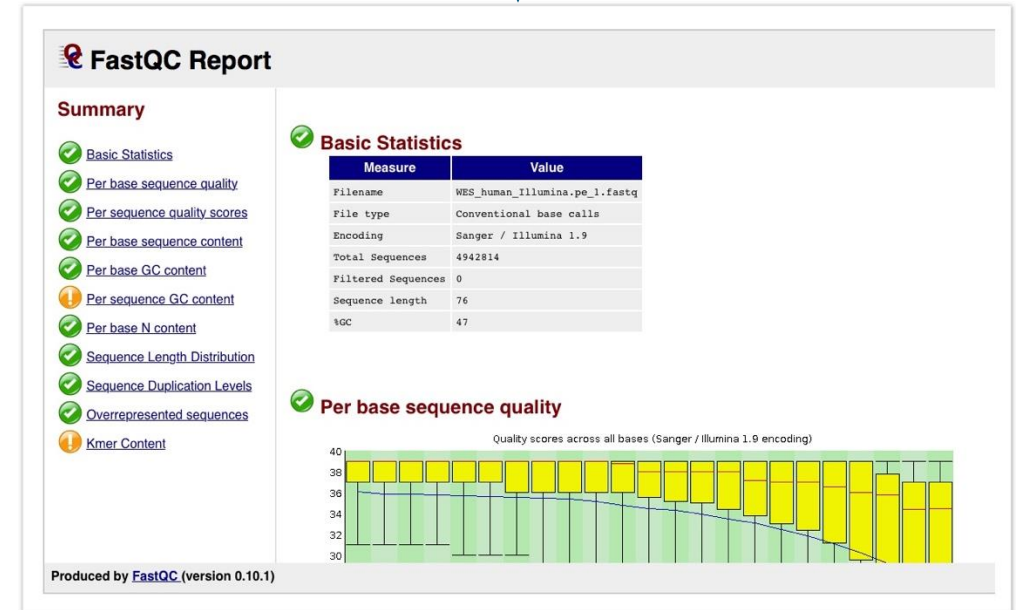
3

Host DNA removal

- Align reads to host genome

Bowtie2

Identifier — | @HWI-EAS209_0006_FC706VJ:5
Sequence — | TTAATTGGTAAATAAATCTCCTAATA
+ sign & identifier — | +HWI-EAS209_0006_FC706VJ:5
Quality scores — | efcfffffcfeefffcfffffdddf`



Bowtie2



Alignment to
host Genome

Data preprocessing cont ...

Taxonomic classification

MetaPhlAn

- Marker based profiling

Kraken2

- K-mer based classifier

species	reads	abundance	pig
s__Anaeromassilibacillus_sp_An172	1965570	10.95139	100
s__GGB9059_SGB13976	NA	6.73851	100
s__GGB82109_SGB113545	NA	5.64422	100
s__GGB9634_SGB15095	NA	4.18550	100
s__GGB9244_SGB14187	NA	3.21764	100
s__Pseudoflavonifractor_phocaeensis	1870988	3.10728	100



Functional annotation

HUMAnN

- UniRef
- MetaCyc

PICRUSt

- 16S sequences

	100	1
1CMET2-PWY: folate transformations III (E. coli) g__Actinobacillus.s__Actinobacillus_minor	0	
1CMET2-PWY: folate transformations III (E. coli) g__Akkermansia.s__Akkermansia_muciniphila	0	
1CMET2-PWY: folate transformations III (E. coli) g__Alistipes.s__Alistipes_finegoldii	0	
1CMET2-PWY: folate transformations III (E. coli) g__Alistipes.s__Alistipes_putredinis	0	
1CMET2-PWY: folate transformations III (E. coli) g__Anaerostipes.s__Anaerostipes_hadrus	0	
1CMET2-PWY: folate transformations III (E. coli) g__Bacteroides.s__Bacteroides_caccae	0	

Analysis – shotgun sequencing

adonis - permanova (better for microbiome research than regular ANOVA)

Diversity analysis

A Alpha diversity

- Shannon
- Simpson

B Beta diversity

- Brays-Curtis
- Jaccard's
- UniFrac

Differential abundance

ANCOM-BC

- Bias correction
- Covariate Adjustment
- Correlation analysis

Maaslin2

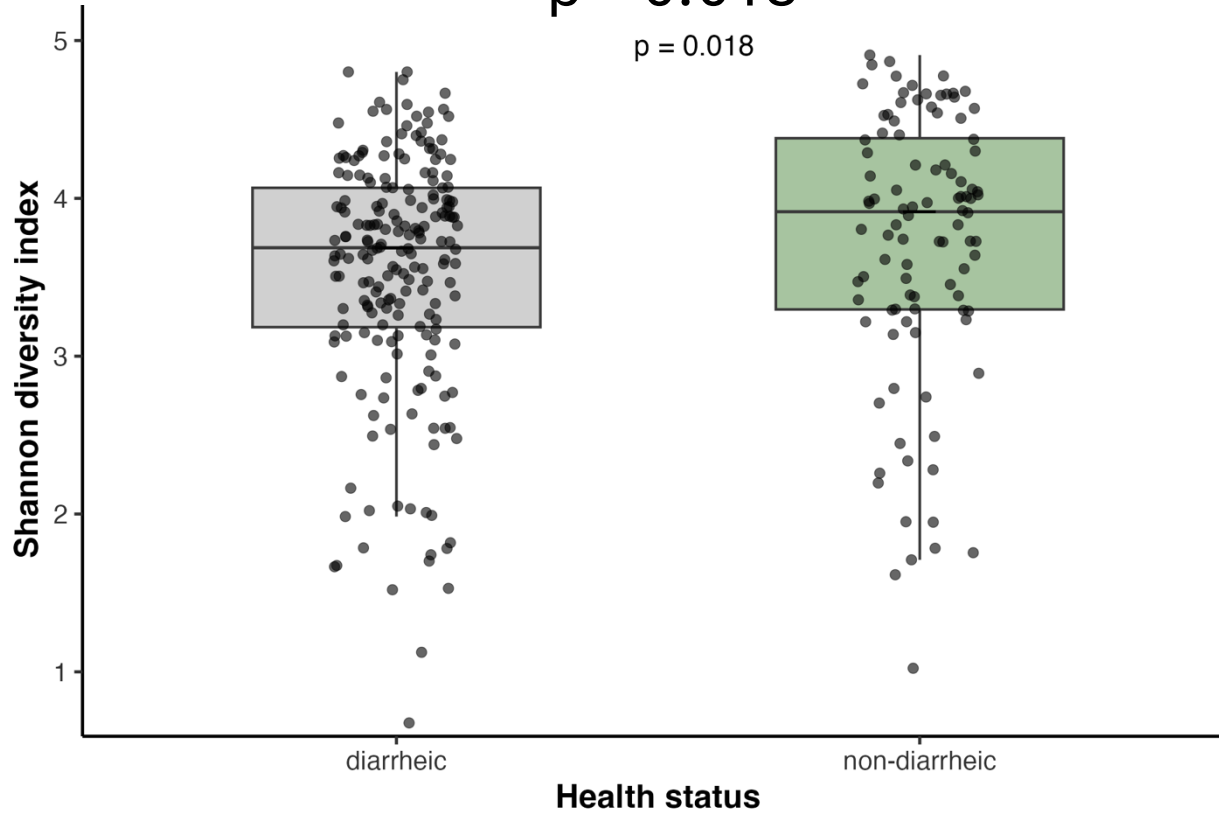
- Multivariable associations
- Linear models

- ADONIS (vegan)
- estimate_richness (Phyloseq)
- Ancomb (ANCOMBC)
- Masslin2

Alpha diversity varied across piglets' **health status** and **weaning age**

Health status

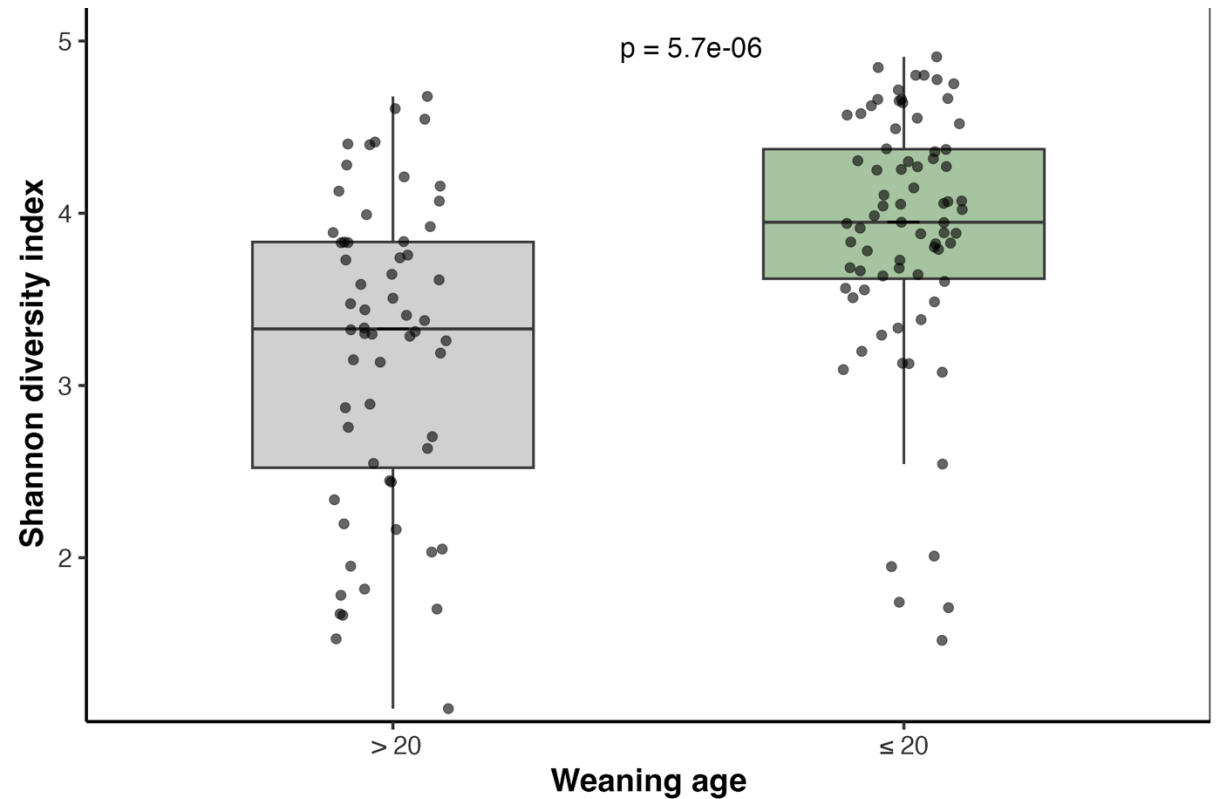
$p = 0.018$



Diarrheic piglets had ***lower*** Shannon diversity

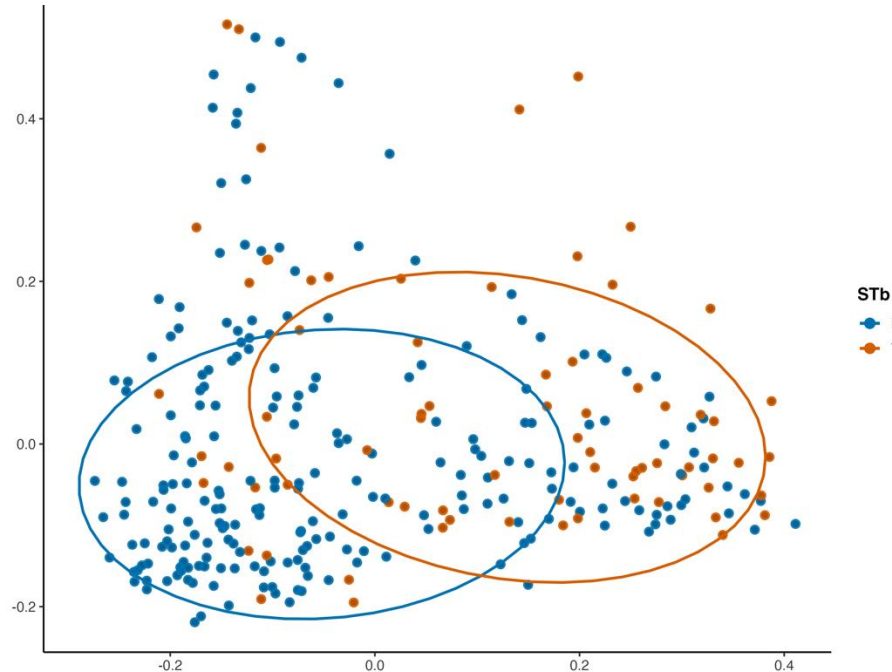
Weaning age

$p = 5.6e-06$



Early weaning led ***higher*** Shannon diversity

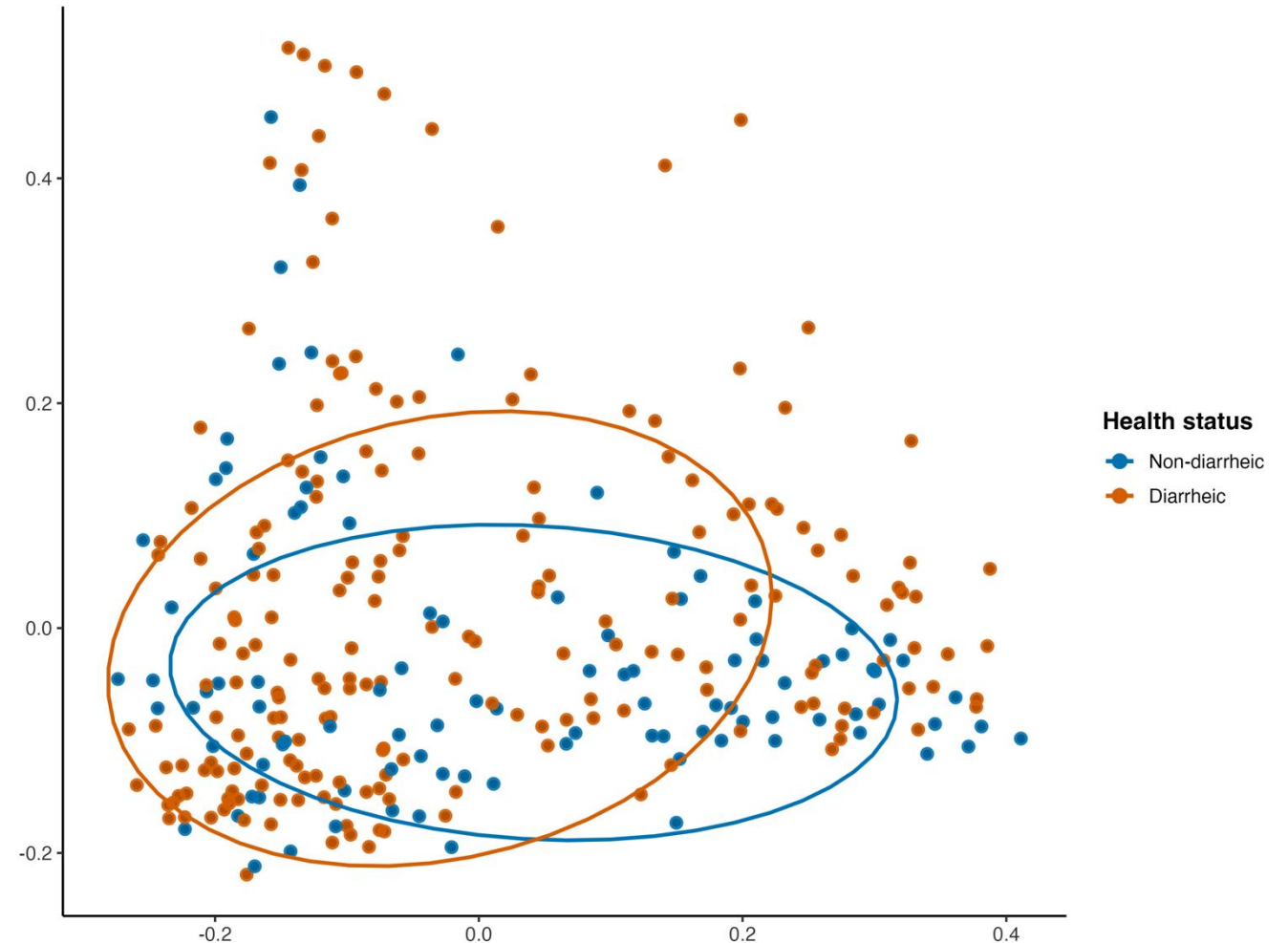
ADONIS confirms significant differences in microbiota composition due to **health status** and **STb shedding**.



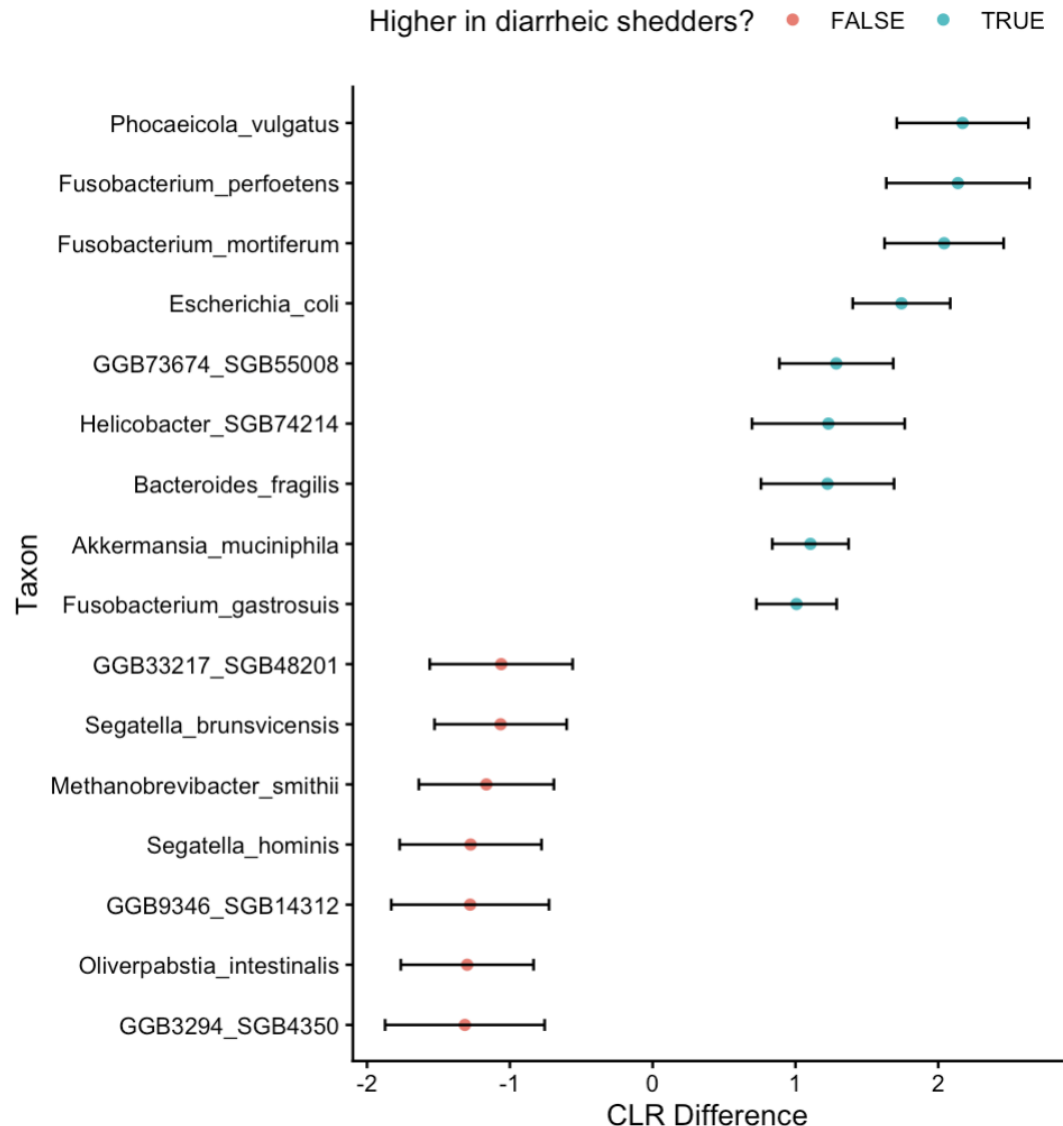
Variable	R2	F	Pr(>F)
----------	----	---	--------

Heath status	0.009	2.71	0.001 *
--------------	-------	------	---------

STb shedding	0.019	5.78	0.03 *
--------------	-------	------	--------



E. coli, *Halicobacter* sp., *Fusobacterium* species are **overrepresented** in diarrheic piglets.



Species **overrepresented** in diarrheic pigs may play a role in **gut dysbiosis**.

Species **overrepresented** in non-diarrheic pig may be associated with gut stability.

- Segatella sp.
- Oliverpabstia intestinalis

For today's vignette

10



DADA2: High-resolution sample inference from Illumina amplicon data

Benjamin J Callahan¹, Paul J McMurdie²,
Michael J Rosen³, Andrew W Han², Amy Jo A Johnson² &
Susan P Holmes¹

We present the open-source software package DADA2 for modeling and correcting Illumina-sequenced amplicon errors (<https://github.com/benjjneb/dada2>). DADA2 infers sample sequences exactly and resolves differences of as little as 1 nucleotide. In several mock communities, DADA2 identified more real variants and output fewer spurious sequences than other methods. We applied DADA2 to vaginal samples from a cohort of pregnant women, revealing a diversity of previously undetected *Lactobacillus crispatus* variants.