

HEB1410 Gut Microbiome and Human Health

Computation Lab Section

Yijia Liow

2023-11-16

Overview

- **Slowing down**
- Experimental design overview
- Hypotheses and predictions formulation
- More on alpha- and beta-diversity
- Strategy to answering your research questions



Diet Dominates Host Genotype in Shaping the Murine Gut Microbiota

Rachel N. Carmody,^{1,2,5} Georg K. Gerber,^{3,5} Jesus M. Luevano, Jr.,¹ Daniel M. Gatti,⁴ Lisa Somes,⁴ Karen L. Svenson,⁴ and Peter J. Turnbaugh^{1,2,*}

¹FAS Center for Systems Biology, Harvard University, 52 Oxford Street, Cambridge, MA 02138, USA

²Department of Microbiology and Immunology, Hooper Foundation, University of California, San Francisco, 513 Parnassus Avenue, San Francisco, CA 94143, USA

³Center for Clinical and Translational Metagenomics, Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, 221 Longwood Avenue, Boston, MA 02115, USA

⁴The Jackson Laboratory, 610 Main Street, Bar Harbor, ME 04609, USA

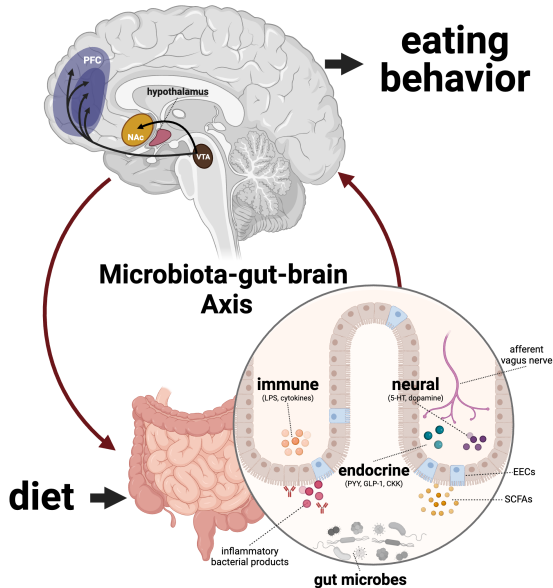
⁵Co-first author

*Correspondence: peter.turnbaugh@ucsf.edu

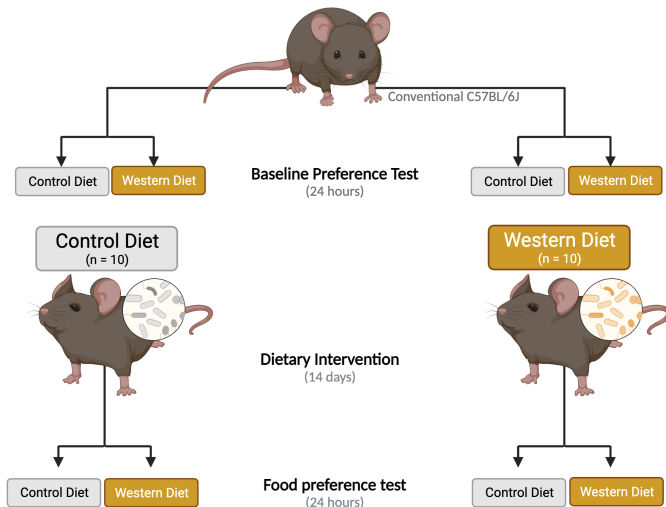
<http://dx.doi.org/10.1016/j.chom.2014.11.010>

- phyloseq
- taxonomic composition, alpha-diversity, beta-diversity
- PERMANOVA

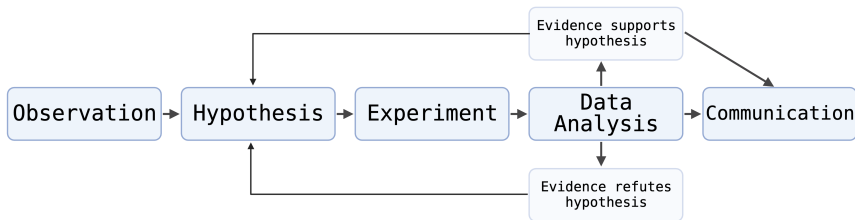
My hypothesis



How to test my hypothesis?



The scientific method



Hypotheses and predictions formulation

- **Hypothesis:** An idea that can be tested through experimentation
 - ▶ statement about the relationship between two or more variables
 - ▶ often takes the form of a cause-and-effect statement
 - ▶ changes in one variable (independent variable) cause changes in another variable (dependent variable)
- **Prediction:** A statement about expected outcomes
 - ▶ more specific
 - ▶ provides detailed expectations about experimental results

Hypotheses and predictions (gut microbiome)

- Mice fed different types of diet will have different gut microbiome composition.

Hypotheses and predictions (gut microbiome)

- Mice fed different types of diet will have different gut microbiome composition.
- A Western diet will reduce the overall microbial diversity compared to a control diet.
- Mice fed with a Western diet will show an increased abundance of the phylum Firmicutes and a decrease in Bacteroidetes.
- A higher relative abundance of pro-inflammatory bacterial genera in Western group mice.

Hypotheses and predictions (phenotypic measures)

- Dietary composition influences mice body weight and composition.

Hypotheses and predictions (phenotypic measures)

- Dietary composition influences mice body weight and composition.
- Western group mice will gain more weight over the intervention period compared to control group mice.
- Western group mice will have a higher percentage of body fat compared to control group mice.
- For every gram consumed, Western group mice will gain more weight compared to control group mice.

Hypotheses and predictions (behavioral)

- Mice on the Western diet will prefer Western diet relative to mice on the control diet.

Hypotheses and predictions (behavioral)

- Mice on the Western diet will prefer Western diet relative to mice on the control diet.
- Western group mice will consume a greater quantity of Western diet compared to control group mice.
- Western group mice will consume a greater percentage of their total caloric intake from Western diet.
- Western group mice will spend more time eating Western diet compared to control group mice.

Hypotheses and predictions formulation

- identify independent variable(s): type of diet (Western vs. control)
 - identify dependent variable(s):
 - ▶ gut microbiome composition
 - ▶ phenotype (body weight, feed efficiency, food preference)
 - propose a direct effect
-
- articulate specific outcomes from experimental data
 - specific results that would support or reject the hypothesis

Data at hand

- **Gut microbiome:** 16S rRNA gene sequencing data
- **Phenotypic measures:** body weight, body composition, food intake

Alpha-diversity (within sample diversity)



Alpha-diversity

- **Observed OTUs:** number of unique OTUs (species)
- **Shannon index:** measure of evenness of the distribution of OTUs
- **Faith's phylogenetic diversity:** measure of the phylogenetic diversity

Observed OTUs (species)

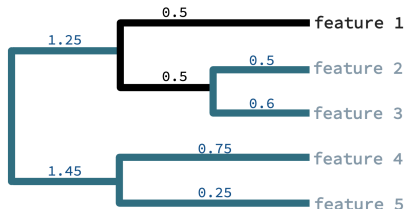
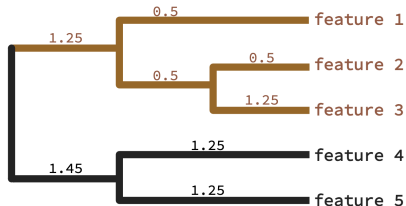
FeatureTable[Frequency]					
	feature1	feature2	feature3	feature4	feature5
4ac2	25	30	15	0	0
e375	0	17	33	25	0

FeatureTable[Frequency]					
	feature1	feature2	feature3	feature4	feature5
4ac2	1	1	1	0	0
e375	0	1	1	1	0

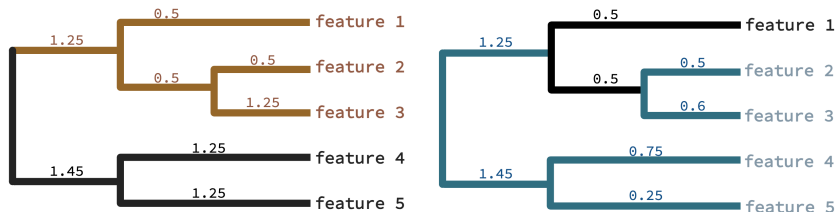
SampleData[AlphaDiversity]	Observed OTUs
4ac2	3
e375	3

Faith's phylogenetic diversity

FeatureTable[Frequency]					
	feature1	feature2	feature3	feature4	feature5
4ac2	25	30	15	0	0
e375	0	17	33	25	0



Faith's phylogenetic diversity



SampleData[AlphaDiversity]	Faith's PD
4ac2	3.35
e375	5.05

Shannon diversity: richness and evenness

$$H' = - \sum_{i=1}^S p_i \log p_i$$

FeatureTable[Frequency]					
	feature1	feature2	feature3	feature4	feature5
4ac2	25	30	15	0	0
e375	0	17	33	25	0

SampleData[AlphaDiversity]	Shannon
4ac2	1.061
e375	1.064

Beta-diversity

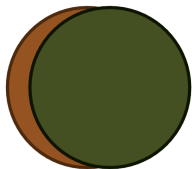


Beta-diversity

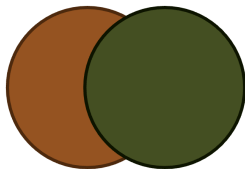
- **Jaccard distance:** presence/absence
- **Bray-Curtis dissimilarity:** abundance
- **UniFrac distance:** abundance and phylogeny

	Presence/Absence	Abundance
Without phylogeny	Jaccard	Bray-Curtis (PCoA)
With phylogeny	Unweighted UniFrac	Weighted UniFrac

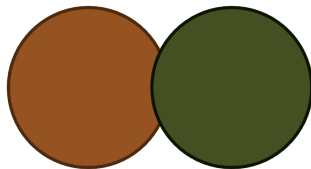
Jaccard distance



$$d_j = 0$$



$$d_j = 0.5$$



$$d_j = 1$$

Jaccard distance

FeatureTable[Frequency]					
	feature1	feature2	feature3	feature4	feature5
4ac2	42	0	37	99	1
e375	12	1	22	88	0
4qd8	25	3	23	86	0
9872	0	0	87	12	0

$$J(A, B) = 1 - \frac{|A \cap B|}{|A \cup B|}$$

DistanceMatrix				
	4ac2	e375	4qd8	9872
4ac2	0.0	0.4	0.4	0.5
e375	0.4	0.0	0.0	0.5
4qd8	0.4	0.0	0.0	0.5
9872	0.5	0.5	0.5	0.0

Bray-Curtis distance

$$BC(A, B) = \frac{\sum_i |X_{iA} - X_{iB}|}{\sum_i (X_{iA} + X_{iB})}$$

- X_{iA} is the frequency of feature i in sample A
- X_{iB} is the frequency of feature i in sample B

FeatureTable[Frequency]					
	feature1	feature2	feature3	feature4	feature5
4ac2	42	0	37	99	1
e375	12	1	22	88	0
4qd8	25	3	23	86	0
9872	0	0	87	12	0

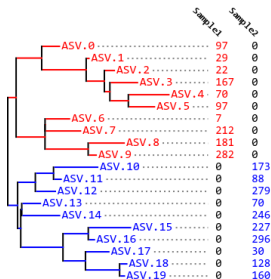
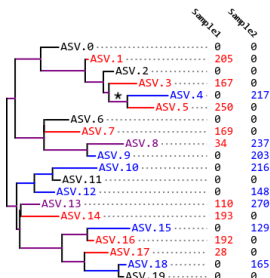
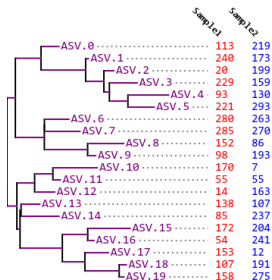
Bray-Curtis distance

FeatureTable[Frequency]					
	feature1	feature2	feature3	feature4	feature5
4ac2	42	0	37	99	1
e375	12	1	22	88	0
4qd8	25	3	23	86	0
9872	0	0	87	12	0

DistanceMatrix				
	4ac2	e375	4qd8	9872
4ac2	0.0	0.19	0.15	0.65
e375	0.19	0.0	0.07	0.60
4qd8	0.15	0.07	0.0	0.70
9872	0.65	0.69	0.70	0.0

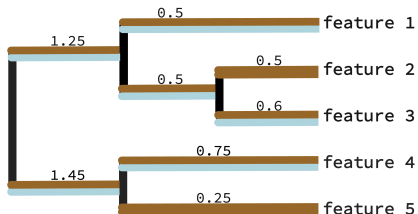
UniFrac distance: Weighted and Unweighted

$$UU(A, B) = \frac{\text{sum of unique branch length}}{\text{sum of observed branch length}}$$



Unweighted UniFrac distance

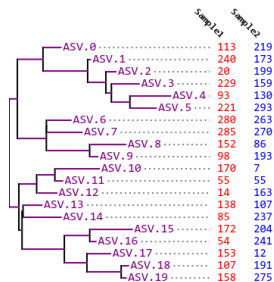
FeatureTable[Frequency]					
	feature1	feature2	feature3	feature4	feature5
4ac2	42	0	37	99	1
e375	12	1	22	88	0
4qd8	25	3	23	86	0



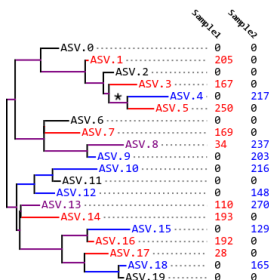
DistanceMatrix			
	4ac2	e375	4qd8
4ac2	0.0		
e375	0.13	0.0	
4qd8	0.14		0.0

Unweighted UniFrac distance

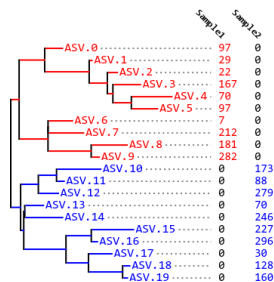
$$UU(A, B) = \frac{\text{sum of unique branch length}}{\text{sum of observed branch length}}$$



(a) 0.000000



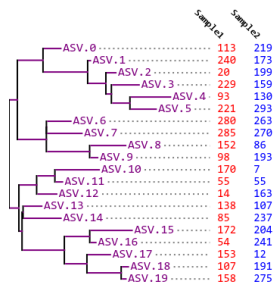
(b) 0.526659



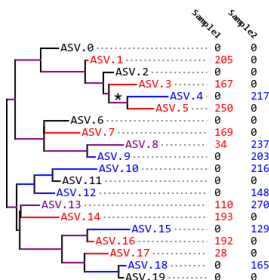
(c) 1.000000

Weighted UniFrac distance

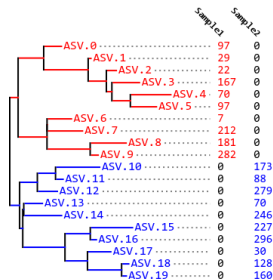
$$UU(A, B) = \frac{\text{sum of unique branch length}}{\text{sum of observed branch length}}$$



(a) 0.118186



(b) 0.560364



(c) 1.000000

Summary of distance metrics

	Presence/Absence	Abundance
Without phylogeny	Jaccard	Bray-Curtis (PCoA)
With phylogeny	Unweighted UniFrac	Weighted UniFrac

Brainstorming analysis strategies

- Formulate **hypotheses** about relationships you expect to find within the data
 - ▶ What factors do you predict will influence body weight or composition?
 - ▶ How might food preferences correlate with gut microbiome diversity?
- Make **predictions** based on your hypotheses
 - ▶ If hypothesis is correct, what patterns would you expect to see?
 - ▶ How would changes in feed intake during the intervention period reflect on body weight or gut microbiome composition?
- Decide on **analysis tools** and **approaches**:
 - ▶ Which distance metrics (e.g., Bray-Curtis, Jaccard) would be most appropriate for comparing microbiome samples?
 - ▶ Would correlation analysis help reveal associations between diet, body composition, and microbiome diversity?
 - ▶ What visual representations can illustrate the relationships you are investigating?