

Microbiome Bioinformatics

2017 Spring Lecture Series for High School Students

Bioinformatics and Research Computing
Whitehead Institute

<http://jura.wi.mit.edu/bio/education/>



WHITEHEAD INSTITUTE

What is microbiome bioinformatics?

- Microbiome
 - the microorganisms in a particular environment
 - can be in or on the body or somewhere else
- Bioinformatics
 - the application of computational methods to the field of molecular biology
 - also called *computational biology*
- Metagenome
 - the combined genetic material of the microorganisms in a particular environment
- Bioinformatics is essential for studying and interpreting the microbiome and its metagenome.



WHITEHEAD INSTITUTE

2



Bioinformatics & Research Computing

Consultation and collaboration, training and education, and software in the areas of Bioinformatics and Graphics.

at Whitehead Institute



George
Bell

Bingbing
Yuan

Prat
Thiru

Inma
Barrasa



WHITEHEAD INSTITUTE



Why study the microbiome?

- Determine what microorganisms are living in an environment
- For the human microbiome, part of what we want to learn is
 - how the microbiome differs in and on healthy compared to sick people
 - if the microbiome can be used to diagnose disease
 - if changing one's microbiome can make him/her sick or healthy



WHITEHEAD INSTITUTE

4

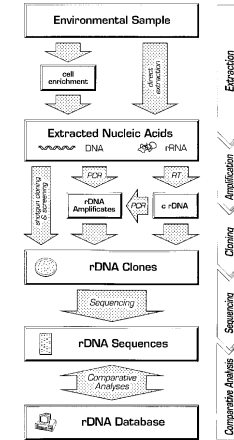


Old-fashioned microbe identification

- Grow the microbes in culture dishes and run classification tests on what grows
- Problems
 - We can only identify major microbe groups
 - Most microbes don't grow in culture dishes
- This gave us interesting information, but we missed most microbes

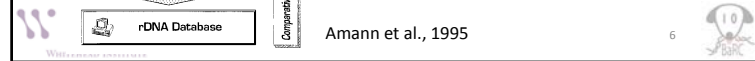


Ribosomal RNA microbe identification



- Identify each microbe
 - by amplifying and sequencing its rRNA
 - Comparing the rRNA sequence to a database of many rRNA sequences
- 16S rRNA (~1500 nt) is most common

Amann et al., 1995



Sample 16S rRNA alignment

	70	80	90	100	110	120	130
<i>E. coli</i>	CGGTARACGGAGAGCTTGCCTTCTTTCTGACGAGTGGCCGACGGGTGAGTAACTGCTGGGA						
<i>M. iranicum M05</i>	CGG-----A--CCCTTT--GGGGTT--AGTGGCCGACGGGTGAGTAACTGCTGGGT						
<i>M. iranicum HHT1087</i>	CGG-----A--CCCTTT--GGGGTT--AGTGGCCGACGGGTGAGTAACTGCTGGGT						
<i>M. iranicum GN19803</i>	CGG-----A--CCCTTT--GGGGTT--AGTGGCCGACGGGTGAGTAACTGCTGGGT						
<i>M. iranicum NJH</i>	CGG-----A--CCCTTT--GGGGTT--AGTGGCCGACGGGTGAGTAACTGCTGGGT						
<i>M. iranicum NLA001001296</i>	CGG-----A--CCCTTT--GGGGTT--AGTGGCCGACGGGTGAGTAACTGCTGGGT						
<i>M. iranicum FJ05196</i>	CGG-----A--CCCTTT--GGGGTT--AGTGGCCGACGGGTGAGTAACTGCTGGGT						
<i>M. iranicum OPB912013762</i>	CGG-----A--CCCTTT--GGGGTT--AGTGGCCGACGGGTGAGTAACTGCTGGGT						
<i>M. iranicum CCUG52297</i>	CGG-----A--CCCTTT--GGGGTT--AGTGGCCGACGGGTGAGTAACTGCTGGGT						
<i>M. iranicum UMTJL</i>	CGG-----A--CC--TT--CGGGTT--AGTGGCCGACGGGTGAGTAACTGCTGGGT						
<i>M. tuberculosis</i>	CGG-----A--RARGCCCTTC--GGGGTACTCGAGTGGCCGACGGGTGAGTAACTGCTGGGT						
<i>M. aubagnense</i>	CGG-----A--RARGCCCTTC--GGGGTACTCGAGTGGCCGACGGGTGAGTAACTGCTGGGT						
<i>M. fallax</i>	CGG-----A--RARGCCCTTC--GGGGTACTCGAGTGGCCGACGGGTGAGTAACTGCTGGGT						
<i>M. mageritense</i>	CGG-----A--RARGCCCTTC--GGGGTACTCGAGTGGCCGACGGGTGAGTAACTGCTGGGT						
<i>M. senegalense</i>	CGG-----A--RARGCCCTTC--GGGGTACTCGAGTGGCCGACGGGTGAGTAACTGCTGGGT						
<i>M. farcinogenes</i>	CGG-----A--RARGCCCTTC--GGGGTACTCGAGTGGCCGACGGGTGAGTAACTGCTGGGT						
<i>M. neworleansense</i>	CGG-----A--RARGCCCTTC--GGGGTACTCGAGTGGCCGACGGGTGAGTAACTGCTGGGT						
<i>M. fortuitum subsp. acetamidolyticum</i>	CGG-----A--RARGCCCTTC--GGGGTACTCGAGTGGCCGACGGGTGAGTAACTGCTGGGT						
<i>M. wolinskyi</i>	CGG-----A--RARGCCCTTC--GGGGTACTCGAGTGGCCGACGGGTGAGTAACTGCTGGGT						
<i>M. phocaeum</i>	CGG-----A--RARGCCCTTCGGGGTACTCGAGTGGCCGACGGGTGAGTAACTGCTGGGT						
<i>M. conceptionense</i>	CGG-----A--RARGCCCTTC--GGGGTACTCGAGTGGCCGACGGGTGAGTAACTGCTGGGT						
<i>M. porcinum</i>	CGG-----A--RARGCCCTTC--GGGGTACTCGAGTGGCCGACGGGTGAGTAACTGCTGGGT						
<i>M. septicum</i>	CGG-----A--RARGCCCTTC--GGGGTACTCGAGTGGCCGACGGGTGAGTAACTGCTGGGT						
<i>M. rhodesiae</i>	CGG-----A--RARGCCCTTC--GGGGTACTCGAGTGGCCGACGGGTGAGTAACTGCTGGGT						
<i>M. sphagni</i>	CGG-----A--RARGCCCTTCGGGGTACTCGAGTGGCCGACGGGTGAGTAACTGCTGGGT						
<i>M. austroafricanum</i>	CGG-----A--RARGCCCTTC--GGGGTACTCGAGTGGCCGACGGGTGAGTAACTGCTGGGT						
<i>M. novocastriense</i>	CGG-----A--RARGCCCTTC--GGGGTACTCGAGTGGCCGACGGGTGAGTAACTGCTGGGT						
<i>M. morioakaense</i>	CGG-----A--RARGCCCTTC--GGGGTACTCGAGTGGCCGACGGGTGAGTAACTGCTGGGT						
<i>M. komossense</i>	CGG-----A--RARGCCCTTC--GGGGTACTCGAGTGGCCGACGGGTGAGTAACTGCTGGGT						
<i>M. aurum</i>	CGG-----A--RARGCCCTTC--GGGGTACTCGAGTGGCCGACGGGTGAGTAACTGCTGGGT						
<i>M. parafortuitum</i>	CGG-----A--RARGCCCTTCGGGGTACTCGAGTGGCCGACGGGTGAGTAACTGCTGGGT						
<i>M. gilvum Spv1</i>	CGG-----A--RARGCCCTTC--GGGGTACTCGAGTGGCCGACGGGTGAGTAACTGCTGGGT						
<i>M. obuense</i>	CGA-----A--TGARGCCTTC--GGGTGGATTAGTGGCCGACGGGTGAGTAACTGCTGGGT						
<i>M. brumae</i>	CGG-----A--RARGCTCTTCGGAGTACTCGAGTGGCCGACGGGTGAGTAACTGCTGGGT						



Metagenomic sequencing

- An attempt to sequence all the DNA in a sample
- Requires a reference database of microbe genomes for comparison and identification
- Can be used to identify genes and type of organism (and predict what the genes do)
- The DNA will typically also contain DNA from the host (ex: the person)



The Human Microbiome Project

- Goal: characterize the human microbiome and its role in human health and disease
- Assayed 300 people at 15-18 body sites
- Used 16S rRNA gene sequencing
- Lots of interesting findings
- Q: Does the microbiome depend more on the person or on the body site?

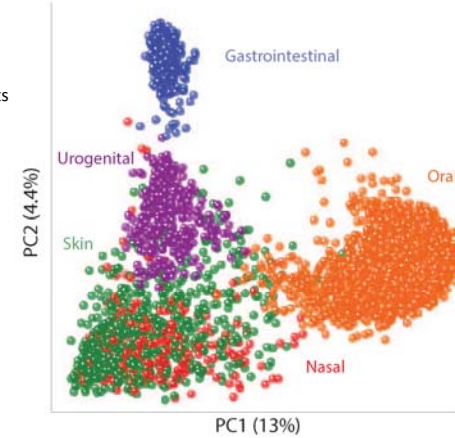


9



Human variation between 5 main body areas

Principal components plot

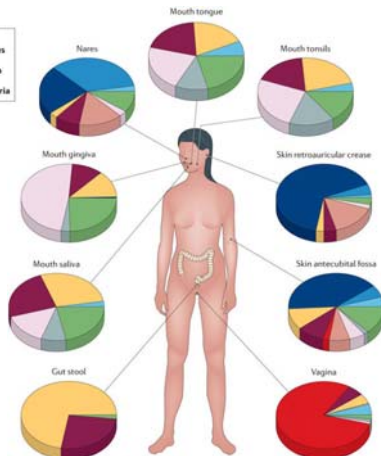
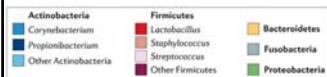


The Human Microbiome Project Consortium, 2012.

10



Bacterial phyla at different body sites

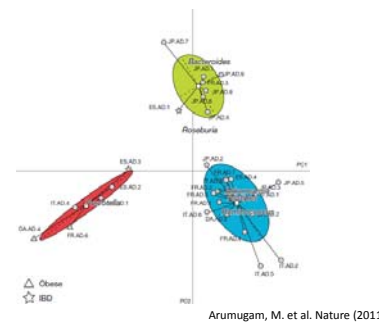


Summary from the Human Microbiome Project

Lasken and McLean, 2014

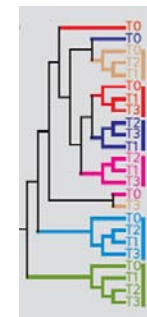


Each person has a unique gut microbiome



Arumugam, M. et al. Nature (2011)

Differences within the same enterotypes between individuals



Ley, R.E., et al. Nature (2006)

Clustering of gut flora is by individual (same color bars) over time (approx 3, 6, and 12 months)



12



Medical conditions associated with microbiome changes

- Dental cavities
- Diabetes
- Obesity
- Inflammatory bowel disease (Crohn’s disease, etc.)
- Acute gastroenteritis
- Some skin diseases
- Some food allergies



13



What influences what?

- Observation: obese people tend to have a different gut microbiome than healthy people
- Why? Many possible explanations
 1. Obesity influences the microbiome
 2. The microbiome influences one’s chance of becoming obese
 3. One’s diet leads to obesity and a change in the microbiome
 4. Others?
- How could you figure out which explanation(s) are correct?
- To remember: “Correlation does not imply causation”

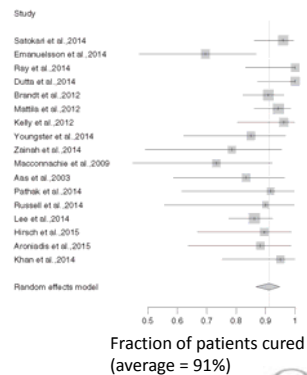


14



Fecal transplantation (part 1)

- Can a “good” microbiome cure a disease?
- Experiment:
 - Take stool from a healthy person
 - Put it into a sick person’s gut
 - Does the sick person get better?
- Example studies (at right):
 - diarrhea caused by *Clostridium difficile*
- Stool source: OpenBiome project (\$40 per deposit)

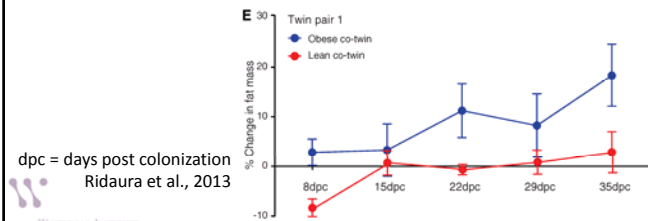


15



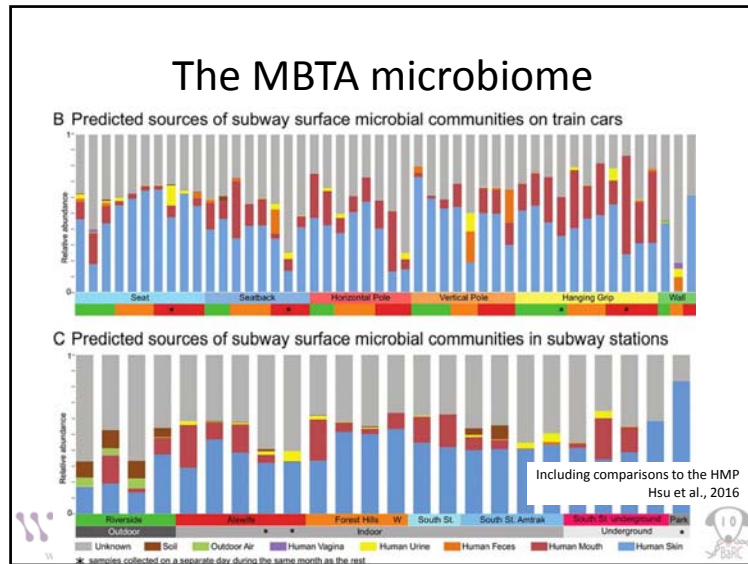
Fecal transplantation (part 2)

- Does the gut microbiome influence body fat?
- Experiment:
 - Raise germ-free baby mice
 - Do fecal transplantation from lean and obese **human** twin donors
 - Does this influence the mice as they grow? Yes!



16





Microbiome Bioinformatics: exercises

- How can gut microbiome rRNA help diagnose gastrointestinal problems?
- How do bacterial groups differ between the microbiomes of different body sites?
- [How] does the gut microbiome change after fecal transplantation?
- etc.

