Homework 1

ABE 587

Kai Blumberg

I made use of the paper: Segata et al. 2012: *Composition of the adult digestive tract bacterial microbiome based on seven mouth surfaces, tonsils, throat and stool samples*

a. What was the overall purpose of the study?

The purpose of the study was to two-fold. One goal was to characterize the microbial community composition of 10 different bodily sites including buccal mucosa and stool. The second purpose was to analyze the metabolic potentials of four representative body sites which also included buccal mucosa and stool.

b. What data did they collect? (e.g. #samples, sampling site, disease state, sampling time, etc)

The authors collected data from 200 health American adults aged 18 to 40. They collected samples from 15 male and 18 female body habitats from each participant.

c. What sequencing technologies did they use? (amplicon, whole genome shotgun, or both?)

The authors made use of both 16S rRNA bacterial gene amplicons from all sample sites as well as whole genome shotgun sequencing data from selected sites.

d. What scientific questions did they address? (see Koskella et al. )

Although the study was generally just addressing the basic microbial ecology questions of “Who is there?”, and “What are they doing?” this study also in a sense addressed the question from Koskella et al., of “Comparing functional relatedness to phylogenetic relatedness to evaluate the degree to which related species are expected to have similar function and/or compete versus coexist within the microbiome.”.

e. What were the major findings of the study?

I think the main contribution of the study was the creation of a baseline understanding of the community composition and taxonomy of healthy American adults. The notable results discussed in the study included 1) the characterization of four distinctive clusters of bodily environments based on community composition. 2) The detection of Phyla previously identified from environmental samples throughout the microbiomes of human body sites. 3) Tooth-associated and oral surface microbial communities were similar (not surprisingly). Finally 4) the majority of metabolic processes were distributed widely throughout the digestive tract microbiota, and varied in terms of metagenomic abundance across different body sites.

f. Does your data overlap with this study (check the methods for accession numbers)? What scientific questions might you ask given your dataset?

We will be making use of both the 16S and WGS data collected from buccal mucosa and stool site data which was collected in the course of this study.

In terms of WGS data based on the results of this study it would be interesting to ask questions such as: “What cysteine and methionine metabolic processes, as well as central carbohydrate metabolic processes differentiate buccal mucosa and stool samples?”. I think this could actually be a neat system by which to employ a workflow which I developed in my Master’s thesis research in which I made use of term annotations from the Gene Ontology (GO) to help differentiate samples at various levels of depth within the GO hierarchy to help investigate and hypothesis putative difference in genomic potential of various environments. I can provide the figure from my Master’s thesis to further explain.