



THE UNIVERSITY OF CHICAGO

Department of Statistics

DISSERTATION PROPOSAL

YUNFAN TANG

Department of Statistics
The University of Chicago

Statistical Methods on Microbiome Data

THURSDAY, November 17, 2016, at 1:00 PM
Jones 304, 5747 S. Ellis Avenue

ABSTRACT

Microbiome refers to the full collection of genes of all microbes in a community. The advent of next generation sequencing technologies, such as 454 pyrosequencing and Illumina Solexa, has allowed researchers to investigate the microbiome communities at an unprecedented level of quantification. Traditional Dirichlet-Multinomial (DM) models ignore species relatedness, leading to loss in efficiency and to results that are difficult to interpret. We develop a DM extension, named PhyloDM, that solves these issues by replacing the global model with a cascade of independent local DMs on the internal nodes of the phylogenetic tree. Since distributional differences tend to occur in clusters along evolutionary lineages, we design a scan statistic over the phylogenetic tree to allow nodes to borrow signal strength from their parents and children. In the end, we will also discuss prediction methods that take advantage of these properties of microbiome data.

For information about building access for persons with disabilities, please contact Laura Rigazzi at 773.702-0541 or send an email to lrigazzi@galton.uchicago.edu. If you wish to subscribe to our email list, please visit the following web site: <https://lists.uchicago.edu/web/arc/statseminars>.