TJHSST Computer Systems Lab Senior Research Project
Statistical Analysis of Mouse Gut Microbiota 2009-2010

Alex Tran

January 28, 2010

Abstract

The mouse gut microbiotal community is the population of bacteria, which inhabit the digestive track of a given population of mice. The ability to understand the microbiotal community has implications, which extend beyond mice and into factors of human obesity and the dietary needs of the human body. Computational genomics is an emerging field, which blends both biology and computer science, to analyze genes. This study seeks to utilize several methods emerging within computational genomics to analyze the gut microbiome of mice to observe the effects of varying diets on the gut microbiotal community, as well as create a universal and user friendly tool for researchers to utilize when studying any gut microbiome. The applications of this can extend past studying the gut microbiome, but can also be applied to any taxonomic group being counted for analysis, however several parts of the application are exclusively beneficial to the analysis of gut microbiota specifically those found in mice.

Keywords: Computational Genomics, Genomics, Gut Microbiota, Statistics, Microbiota

1 Introduction

1.1 Scope of Study

To use genomics to observe and document the gut microbiotal community in the fecal matter of mice during all stages of analysis. From the sequencing of genetic material to the processing of the microbial community. To create an all encompassing suite to streamline the analysis of any microbiotal and community.

1.2 Expected results

To be able to present data from the microbial community in a way that is user friendly, efficient, and able to represent the data in an accurate way. Implemented Gui utilization that will display statistical information on each run, which can be used as visuals in a report, and to be able to do a full run from the ground up given 16sRNA sequencing data that can accurately and efficiently identify individual microbiota within the commmunity. The end result of which will be an all comprehensive applictaion that will be able to take generated data, and give the user a full evaluation of these data.

2 Literature Review

Other research has observed that general imbalances inside the gut microbiome can lead to differences in tendencies towards Obesity in mice(classified as OB/OB see figure.1)(Turnbaugh, Gordon).It has also observed that within various populations of mice the mirobiome appears to be so diverse that there is little known correlation between an individual and the population of their microbiome, however it was also shown that though the population in the microbiome is diverse there are tendencies based on the diet of an individual mouse for an increased representation of certain strains of bacteria to be present within the microbiotal community (Turnbaugh, et. al.)

Figure 1: Figure.1 Image of an OB/OB mouse

Though General imbalances have been observed in the gut microbiome of obese micethere is still not enough information or analysis to say pinpoint what specific species of bacteria can lead to the tendency for mice to be obese.(Turnbaugh, Gordon) This lack of information is largely due to the inefficiency and inaccurate nature of current computational 16s sequencing techniques (Turnbaugh et. al.)

3 Development

The program's end result consists of several parts, which can be created individually in any order. Due to possible problems with time constraints it was been written in reverse depth from the most superficial part to the most abstract. This will allow for user friendly verification as well as a gauranteed presentable result if time had become a problem. It was primarily programmed in Python. Coding began by generating a

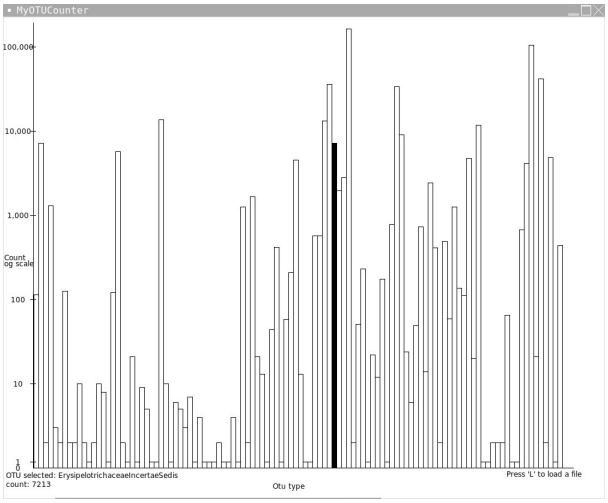

Figure 2: Figure.2 Screenshot of a loaded file, with a data point selected

simple taxonomic tree, and filling it with data points. It then moved onto generating a histogram of these data using Tkinter as a base. After this simple histogram was generated specific file loading, and a variety of taxonomic trees were made available to the user to allow the ability of the application to extend past only gut microbiota in mice to any accepted use of taxonomic OTU counting. Programming then moved on to allow for mouse movement to allow for more user friendliness, and to aid between two varying samples. To test and verify results several premade computational genomics tools that each perform a part of what the finalized project does as one universal tool was be used.

4 Quality assessment

Input data and 16sRNA sequences are to be received from the Gordon Lab at Washington University in St. Louis. The main goal of this project was to provide auser friendly OTU analysis suite. To do this a large portion of development and assessment was done in minor debugging stage to presnt the user with the easiest to use interface possible. When a user has inputted a file they are presented with the sceren seen in Figure.2

5 Expected Results

To be able to present data from the microbial community in a way that is user friendly, efficient, and able to represent the data in an accurate way. To achieve this a gui was utilized that allows the display of statistical information on each run in a way that allows the user to more easily gather information from a sample then has ever before been possible. As a result of this analysis of data gathered from the gut microbiome can be done vastly faster than other archaic forms of analysis previously used. It cuts down on the amount of fact checking, and manual calculation that was previously necessary when doing analysis of this nature.

The results from this study are potentially very helpful to the field of genomics, because tools for analysis of gut microbiotal communities are a field that is in its early stages of development, and at this point most tools which are planned to be implemented

into the end result application are only available separately, and very expensive.

6 Discussion

7 Bibliography

Peter J. Turnbaugh and Jeffrey I. Gordon, "The core gut microbiome, energy balance, and obesity." Center for Genome Sciences, Washington University School of Medicine, St. Louis MO, 2009.
Peter J. Turnbaugh, Ruth E. Ley, Micah Hamady, Claire M. Fraser-Liggett, Rob Knight Jeffrey I. Gordon, "The Human Microbiome Project."