Analysis of 16S rRNA in Human Fecal Samples to Assess Stability of Microbial Population by Terminal Restriction Fragment Patterns

Jonathan C. San
Environmental Biotechnology Institute, Biological Sciences, Cal Poly State University, San Luis Obispo, CA 93407

Abstract

Terminal Restriction Fragment Length Polymorphism analysis targeting the 16S rRNA gene is a very useful tool in quantification of gastrointestinal microflora communities. During a fourteen day period, consecutive fecal samples were collected from an individual without a prior history of antibiotics consumption within a 60 day period. Through extraction and isolation of DNA, community dynamics were assessed through yielded T-RFLP patterns. The TRF data was used to construct a Bray-Curtis similarity dendrogram to determine stable and unstable microflora in studied subjects. Subjects with an average similarity below 50% are deemed unstable while those above 50% can be regarded as stable. With an average similarity of 64.4%, the microflora activity present in the subject was deemed stable.

Introduction

As an ecosystem, the human intestinal flora plays an important aspect in maintaining and preserving the health of an individual. Clinical studies have shown that therapeutic doses of antibiotics may change the stability of eubacterial species in an individual.(1,2,3) Often, doses of antimicrobial drugs may lead to the selection of resistant bacteria and disruption of the colonization barrier (or barrier effect) of resident intestinal flora. Colonization barrier or barrier effect is known as the inability of the normal flora to colonize the gastrointestinal tract.(4) Therefore, the reduced function in the barrier effect will allow exogenous and indigenous non-communal pathogens to compete for space and nutrients, leading to and permitting colonization within the host's gastrointestinal tract. Intestinal flora also enables both the innate and adaptive immunity to respond to pathogenic microorganisms in a timely manner by priming the immune system.

The most common intestinal flora in the human intestinal tract mainly consists of bacteria belonging to the genus *Bifidobacterium* and *Eubacterium.*(5) Mainly, Bifidobacteria are known for their maintenance of human health.(6) The presence of gastrointestinal flora is of obvious importance to our gastrointestinal system and in maintaining our body as a whole.

Currently, there are numerous ways to identify bacterial cells within the intestine but the most effective and least error prone method is through the isolation of the 16S rRNA gene for identification and comparison via database. The 16S rRNA gene sequence is about 1,542 base pairs (bp) long; its most important feature is the fact that it contains highly conserved regions. To compare unknown strains, universal primers are usually chosen as complementary bases to the conserved regions at the beginning of the gene and at either the 540-bp region or at the end of the whole sequence.(7,8) The 16S rRNA gene sequence has been determined for a large number of strains. Since the 16S rRNA gene is universal among bacteria, relationships can be measured among all bacteria.(9,10) In general, the comparison of the 16S rRNA gene sequences allows differentiation between organisms at the genus level across all major phyla of bacteria, in addition to classifying strains at multiple levels, including what we now call the species and subspecies level. However, the identification of bacterial species is not as eminent in this study. Instead, changes of bacterial species over time will be focused upon through analysis of the 16S rRNA gene by means of Terminal Restriction Length Polymorphism (T-RFLP) analysis.

T-RFLP has been proven to be an effective and useful tool in describing changes in intestinal microflora. This particular technique allows quick analysis of bacterial species without cultivation. Terminal Restriction Fragment (TRF) patterns will be generated from isolated DNA, targeting the conserved 16S rRNA gene. Prior to TRF patterns generation Polymerase Chain Reaction (PCR), Gel Electrophoresis (GE), and restriction enzyme digestion will amplify and fluorescently label the 16S rRNA gene. Similar techniques were used in a previous experiment to attempt to determine the stability of microflora community in individuals who were not consuming antibiotics during

the time of the study. However, the experiment was unsound due to the restricted amount of subjects available, insubstantial data and liberal sample storage procedures. Subjects with an average similarity >50% are labeled as stable, while those below this threshold are considered unstable (Figure 1). The data collected in the previous study deemed 40% of the studied subjects unstable. This study aims to accrue further data to generate a statistically reliable study. The procedures in this study will employ similar techniques as previous studies to gather and analyze data.(11)

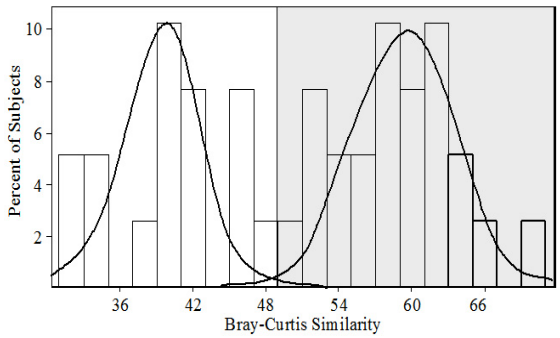


Figure 1. Average Bray-Curtis similarity of TRF patterns created using sample days 1, 7, and 14. Stability/unstability was determined to be between the two distributions, estimated to be 50%. (Figure and data taken from Engelbrektson)

Materials and Methods

Experimental Approach

To quantitatively represent the amount of microflora present in a subject, the highly conserved 16S rRNA gene will be exploited and utilized. DNA will be extracted from fecal samples and used as a template for amplification in PCR. The products will then be subjected to gel electrophoresis to determine whether the reaction was successful. Amplicons are then digested with HaeIII, ethanol precipitated, and then analyzed with the Beckman Coulter CEQ 8000X DNA analysis system.

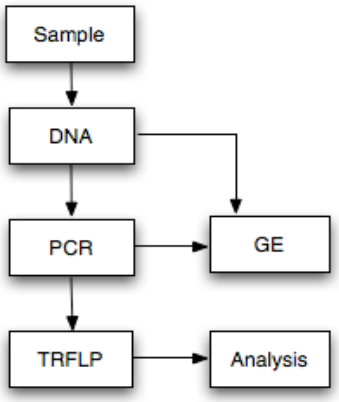


Figure 2. Outline of the molecular approaches used to analyze the fecal bacterial communities present in samples.

Recovery and Storage of Fecal Samples

Fourteen consecutive fecal samples were collected from a healthy individual.‡ There were no dietary restrictions; in addition, antibiotics were not used 60 days prior to or during the study. The samples were voided from both urine and toilet water through the use of a catching device. To ensure a homogenized sample, manual mixing of the sample was required. Samples were then subsequently stored in a microcentrifuge tube and placed in the freezer (0°C) for no more than five days.

DNA Extraction

Fecal samples were extracted using MoBio Powersoil® DNA kit. Each sample was extracted in triplicates adding the fecal sample (0.1 g) to each 2 mL Power Bead tubes. The tubes were subjected to the vortex briefly. Solution C1 (60 µl) was added to each tube with a brief vortex afterwards. The samples were fast prepped at 4.5 m/s for 30 sec. Subsequently, the supernatant was transferred to a clean microcentrifuge tube. To the supernatant, Solution C2 (250 mL) was added. The mixture was then vortexed for 5 sec and then incubated (0°C) for 10-15 min. Solution C2 precipitates any non-DNA organic and inorganic materials. After the incubation period, the samples were centrifuged at 10,000 x kg for 1 minute. Approximately, 550 µL of the supernatant was transferred to a clean microcentrifuge tube. Solution C4 (1.2 mL) was added to each tube and was vortex for 5 sec. The supernatant was loaded in 675 µL into a spin filter and centrifuged at 10,000 x kg for 1 min. The flow through was discarded; this process was repeated until all replicates for a sample day was filter through a single filter. Solution C5 (500 µL) was added and centrifuged at 10,000 x kg for 30 sec. The flow through was discarded and the spin filter was centrifuged at 10,000 x kg for 1 min a second time to remove residual C5 solution. The spin filter was carefully transferred to prevent any contamination from any remaining C5 solution into a sterile microcentrifuge tube. Solution C6 (100 µL) was added to the center of the filter membrane and was left to absorb for 15 min. The spin filter was centrifuged at 10,000 x kg for 30 seconds. Lastly, the spin filter was discarded and the DNA was stored at -20°C. Agarose gel electrophoresis 1.0% (wt/vol) was utilized to determine the success of each extraction.

PCR Amplification

PCR was performed with fluorescently labeled forward primer 46-Ba2F [5' GCY TAA CAC ATG CAA GTC GA 3'] and reverse primer 536-K2R [5' GTA TTA CCG CTG CTG G 3'] was used to amplify the bacterial 16S rRNA conserved gene. The cycling parameters for the reaction was performed under the following conditions: 94°C for 10 min; 30 cycles of 94°C for 1 min, 46.5°C for 1 min, 72°C

‡ Campanale, J., B. Bunyea-Baron

for 2 min; and finally 72°C for 10 min with a 4°C soak. Each sample reaction contained 1 µL of the extracted DNA, 5 µL of 10X buffer GOLD ®, 3 µL of dNTPs (10 µM), 1 µL of K2R (10 µM), 29.7 µL of PCR grade water, and 0.3 µL of taqGOLD ® (5 U/ µL). PCR products were verified by post-stained electrophoresis on a 1.0% (wt/vol) agarose gel with a closed control and an *E. coli* standard acting as the positive control. All reactions were performed in triplicate. Each sample triplicate was then combined during PCR clean up using MoBio PCR Ultra Clean kit following the manufacture's protocol. Quantitation of the PCR products was determined by using the Bio-Tek Fluorometer by measuring the Cy5 fluorescent label.

Enzyme Digest of Amplicons

The amount of DNA template used was dependent upon the concentration in ng/µL determined by the Bio-Tek Fluorometer. Each digestion contained 0.5 µL of HaeIII restriction endonuclease, 2 µL of buffer, DNA template (resolved by dividing the concentration (ng/µL) by 75 ng), and PCR grade water was determined by subtracting 20 µL from the total volume (DNA template, HaeIII, and buffer). Digestion was performed via the GeneAmp Thermal Cycler under the following parameters: 37°C for 4 hours; inactivation at 80°C for 20 min; followed by a 4°C soak.

Ethanol Precipitation/ CEQ Sample Prep

Cold 95% ethanol (100 µL), 2 µL of 3M NaAc (pH4.6), and 1 µL of glycogen (20mg/mL) were added to each digest. The tubes were inverted a couple of times to ensure complete homogenization of the solutions. The mixtures were precipitated at -20°C for 30 min. After removal from the freezer, the samples were centrifuged at 53000 RPM for 15 min. The pellet was washed with 100 µL of 70% cold ethanol and subsequently centrifuged for an additional 5 min at 5300 RPM. Finally, the tubes were inverted and centrifuged at 700 RPM for 1 min to dry the pellet. The pellet was then resuspended in 20 µL of formamide and 0.25 µL of 600 base pair standard. One drop of mineral oil was added to each well to prevent evaporation. The Beckman Coulter CEQ 8000X DNA analysis system was initiated and utilized to obtain terminal restriction fragment patterns.

Statistical Analysis

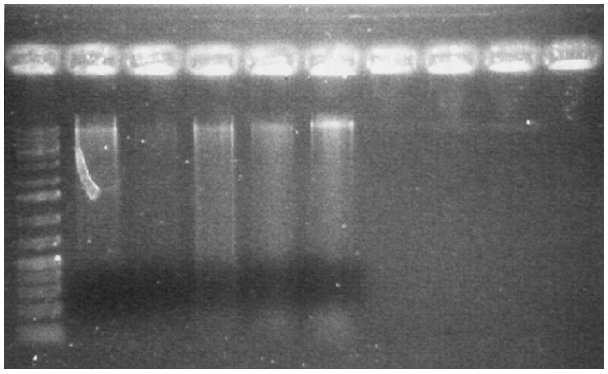
Terminal restriction fragment length polymorphism (TRFLP) in nucleotides, and TRF peak area were exported from the CEQ8000 program into Microsoft Excel. TRF peaks representing less than 0.5% of the total DNA present in the sample were excluded from analysis. TRF peak area was converted to relative abundance (percent of total area) manipulated by square-root to deemphasize the larger peaks, and analyzed with a Bray-Curtis similarity matrix. Statistical analyses were performed using PrimerV and Excel.

Results

DNA Extraction and GE

Nucleic acids were extracted from fecal samples by a mechanical procedure that has been deemed effective via MoBio Powersoil® DNA kit. The bead-beating, used to lyse bacterial cells, optimized the amount of DNA to be extracted. The integrity of the extraction was then analyzed by agarose gel electrophoresis 1.0% (wt/vol). Extraction was completed successfully for each sample day; Figure 3 (below) represents a successful extraction.

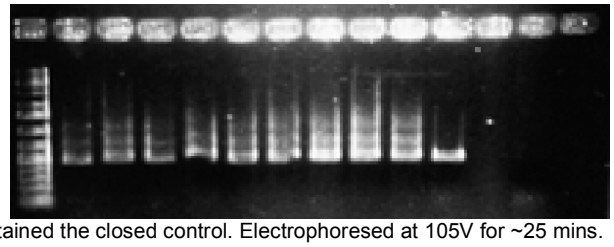
Figure 3. Nucleic acid extraction from sample days 10-14. Lanes 2-6 represent samples 10-14 respectively. Lane 1 (left) contained the 1 Kb ladder. Electrophoresed for ~25 minutes at 100V.



PCR and GE

Amplification of the bacterial conserved region was carried out using forward primer 46-Ba2F [5' GCY TAA CAC ATG CAA GTC GA 3'] and reverse primer 536-K2R [5' GTA TTA CCG CTG CTG G 3']. The faint band present in the closed control lane (Lane 13) can be contributed to primer-dimers; therefore the reaction was not contaminated. All samples were amplified successfully.

Figure 4. GE of 16S rRNA PCR with primers 46-Ba2F and 536-K2R. Lane 1 (left) contained the 1Kb ladder. Lanes 2-10 contained samples 1-9, respectively. Lane 10 contained the E. coli control and lane 13 contained the closed control. Electrophoresed at 105V for ~25 mins.



Bio-Tek Fluorometer

The fluorescently labeled forward primer 46-Ba2F allowed effortless quantification of the PCR products through the use of the A260 UV fluorometer. The following results were yielded:

Sample #	Concentration (ng/µL)
Sample 1	27.367
Sample 2	27.270
Sample 3	30.152
Sample 4	32.591
Sample 5	26.614
Sample 6	36.984
Sample 7	41.191
Sample 8*	49.075
Sample 9	28.584
Sample 10	28.354
Sample 11	32.787
Sample 12	27.758
Sample 13	38.190
Sample 14	40.152
Blank	49.962

Table 1. Data received from the fluorometer with the sensitivity set to 107 (s=107) and the curve set to 75 ng/µL.*Sample 8 was diluted to a 1:10 dilution.

Beckman Coulter CEQ 8000x

Digested Amplicons were subjected to electrophoresis within the Beckman Coulter CEQ 8000x. During this process only fluorescently labeled terminal restriction fragments were visualized. The TRF patterns provided (Figure 5) displays the two most dissimilar samples. On the abscissa, the intensity of a dye signal signifies the number of labeled PCR primers crossing the CEQ detector, verses the length, in nucleotides, of the crossing DNA fragments represented on the ordinate. Digested 16S rRNA fragments will often fall between a range of 50-300 nucleotides; undigested 16S rRNA genes are observed between 450-500 nucleotides possibly belonging to the species *Bifidobacterium*. Raw data yielded from the CEQ was then transferred and transposed into Excel for statistical analysis.

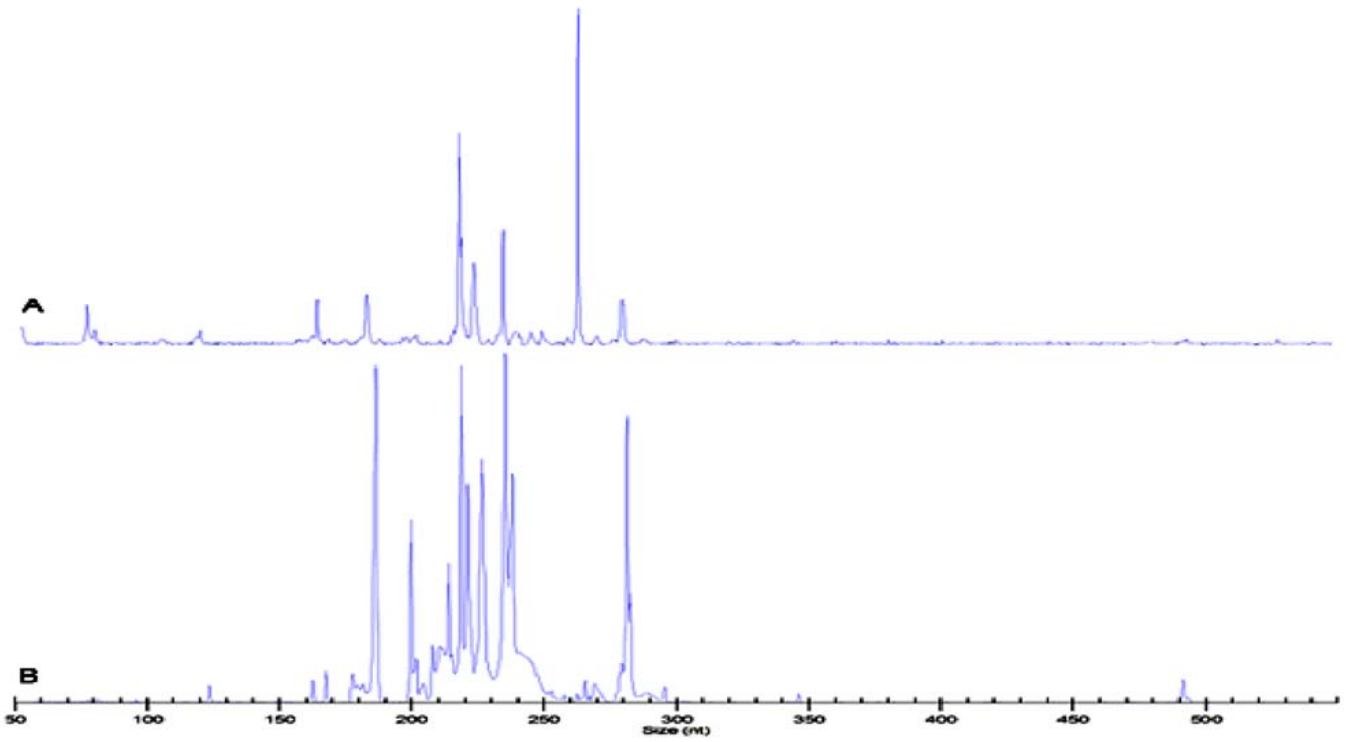


Figure 5. T-RFLP profile of bacterial communities present in fecal samples. Terminal fragments were generated from an HaeIII digestion of 16S rRNA using a fluorescently tagged 46-Ba2F forward primer. These patterns represent the two most dissimilar samples (A: Sample Day 1; B: Sample Day 11).

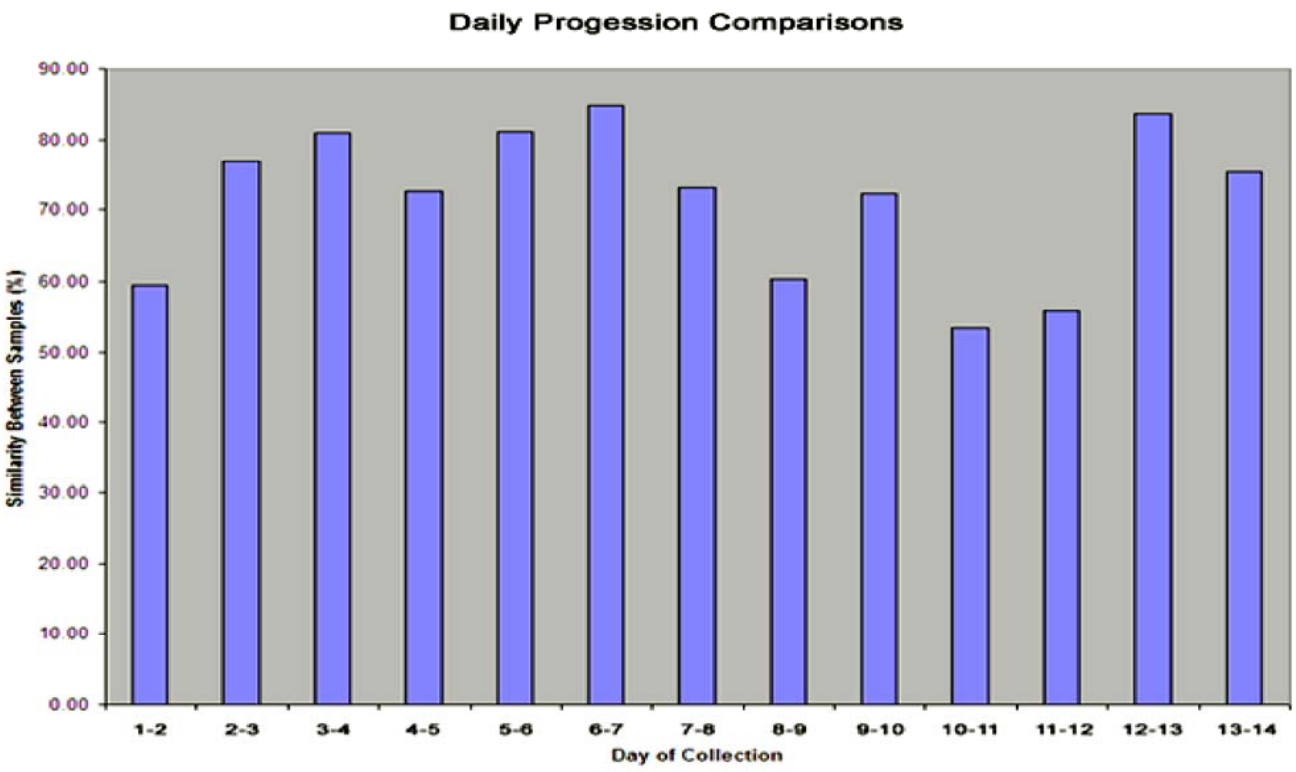


Figure 6. Day-to-day comparison of samples collected during the study

Bray-Curtis

Percent similarity of each TRF pattern was calculated in PrimerV, in which each pattern was compared to the other thirteen patterns. The results were used to create a Bray-Curtis similarity matrix which afterward was used to create a daily progression bar graph indicating the percent similarity of each consecutive TRF pattern to the next (Figure 6). The least similar comparison between two consecutive samples was between samples on day 10 and 11 at 53.3% similarity. Overall the average similarity between all samples was 64.4%. In these two cases percent similarity was above the 50% cutoff threshold and thus are deemed stable. However, random fluctuations of microflora could possible lead to a conclusion where one's microflora is unstable based on Bray-Curtis similarity. Upon averaging the three most dissimilar samples a similarity of 46.0% was yielded. These three samples, days 1, 9, and 11*, could result in the subject being regarded as unstable leading one to conclude that random fluctuations in gastrointestinal microflora can significantly affect the outcome of the study.

Bray-Curtis Similarity Dendrogram

Using PrimerV and data from the Bray-Curtis similarity data, a dendrogram was created to visually aid in the comparison of similarity between samples (Figure 7, pp. 6). Samples that are most similar are clustered together and are separated from those that are least similar. Samples that are clustered near the base of the dendrogram represent the most similar samples within the study.

Individual TRF Peak Persistence and Abundance

TRF peak persistence is based upon the number of samples in which a specific peak is observed. Peaks are then grouped based on their persistence. Once grouped, the area of each peak in a particular group was averaged (Table 2, pp. 6). In conjugation with the number of TRF peaks and persistence, the yielded average peak area from each persistence group was used to determine if a relationship exists between the size of certain bacterial populations and its peak frequency. To assist in analyzing the data, a bubble chart was created (Figure 8, pp. 6) with peak persistence graphed against the number of TRF peaks. The graph chart displayed two very distinct populations that existed; some peaks occurred with high TRF peak persistence but displayed low average peak areas while the other group showed a greater persistence with large average area. An intermediate group was present which both infrequently occurred and contained a small average peak area. This pattern was common among previous experiments.[12,13,14,15,16] This observation of populations with small average peak area and low

* Similarity of samples: 49.9%, 37.2%, and 51.0%

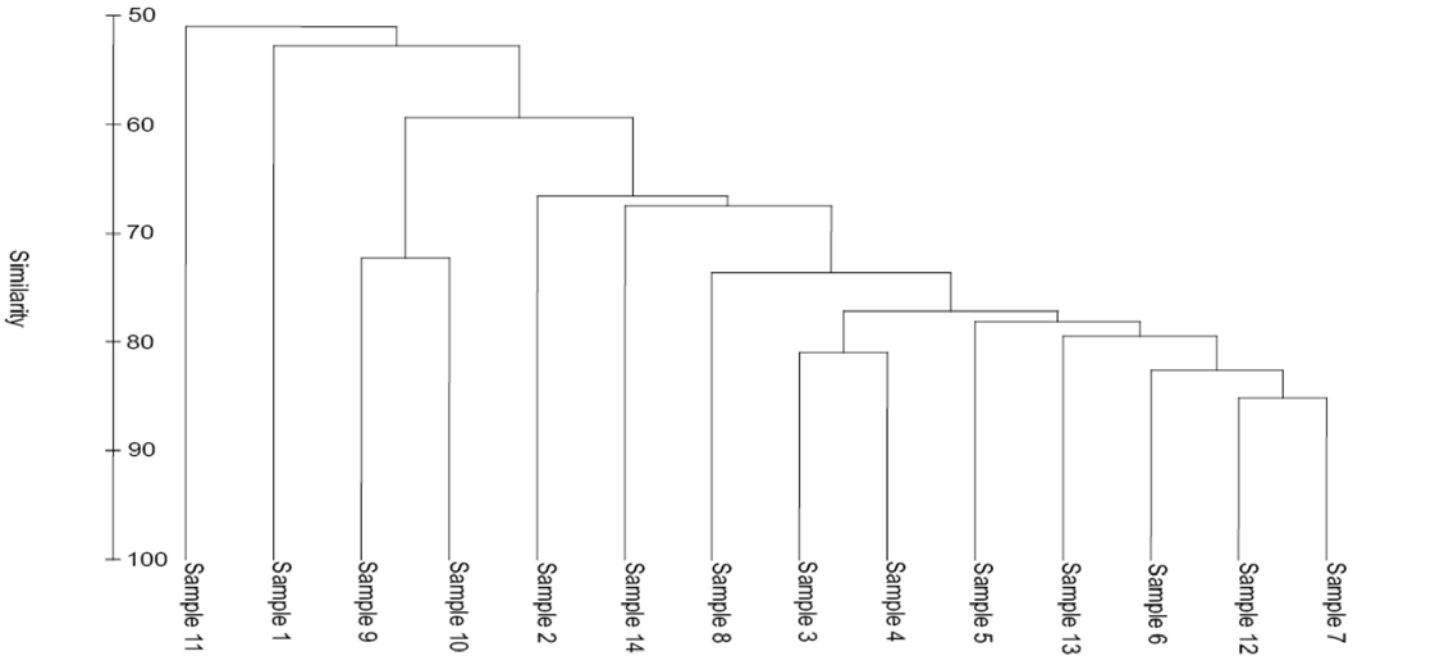


Figure 7. Bray-Curtis similarity between analyzed samples. Such a graph was created to help visually aid in the determination of sample similarity.

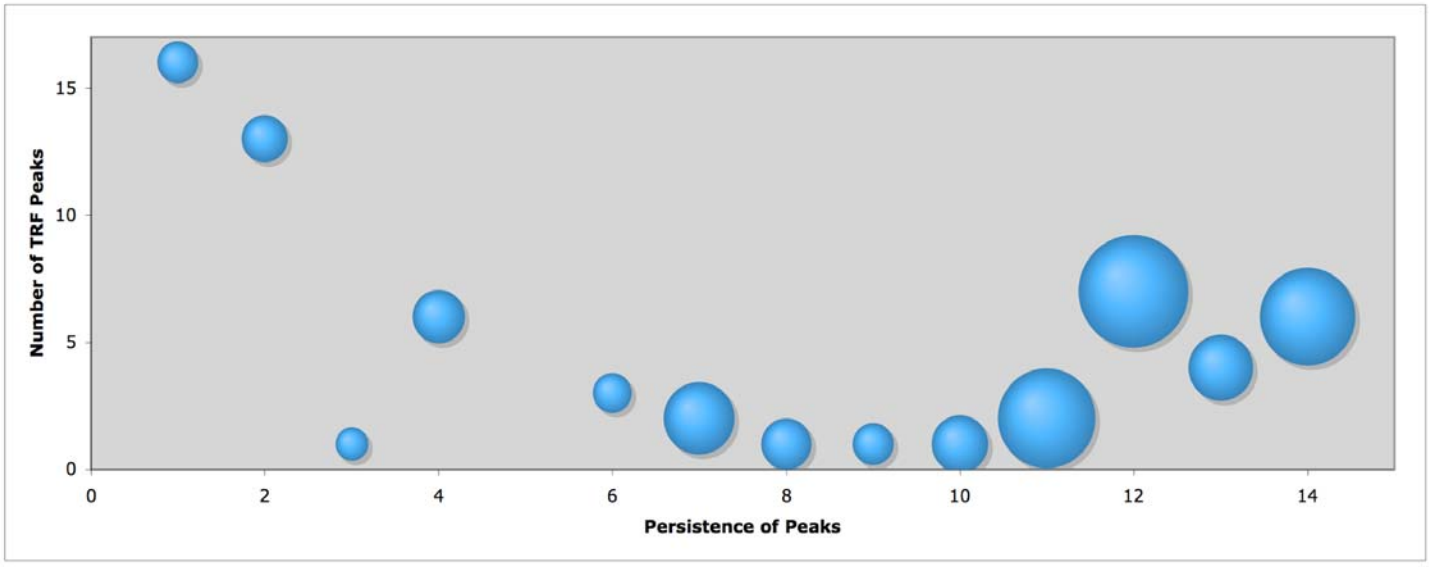


Figure 8. TRFLP peaks were grouped by persistence according to the regularity of presence in nearly all samples (bubbles at the right) or regularity of absence in nearly all samples (bubbles to the left). The average peak area is represented by the diameter of each bubble. Data was obtained from table 2, pp. 7.

Sample Number	Total Number of Peaks	Average of Individual Peak Area (%)
1	16	0.93
2	13	1.21
3	1	0.62
4	6	1.53
5	0	0.00
6	3	0.86
7	2	2.78
8	1	1.41
9	1	0.94
10	1	1.76
11	2	5.26
12	7	6.59
13	4	2.28
14	6	5.02

Table 2. The data present in the above table was transposed onto a graph (Figure 8, pp. 6) for a lucid representation of the data for analysis.

persistence suggests that a group of bacterial populations that exists transiently within the gastrointestinal tract. Thus, frequently occurring populations, those that are represented by a larger average peak area and greater persistence, are more likely to be present in the gastrointestinal tract in greater numbers and persist for a much longer period.

Discussion

The gastrointestinal tract is a habitation for a large and diverse bacterial community, which provide various functions contributing to the overall health of the host. Here, in this study, the stability of bacterial communities was assessed to help determine if a stable bacterial communities exist within the population analyzed. A subject is deemed stable if the samples collected provided TRFLP patterns with a similarity >50%, otherwise anything lower than 50% is deemed unstable. (11) The isolation and analysis of 16S rRNA helped determine that the subject studied in this experiment can be deemed stable, with an average similarity of 64.4%. However, if three random samples were taken and those happen to be the 3 most dissimilar samples*, a percent similarity of 46.0% could have been yielded leading one to believe that the host is unstable. Previous experiments† conducted with the same protocols provided a range of 75.8 - 58.5%, with a mean of 68.2%.

The dendrogram that was created using Primer5 revealed that similarities between the 14 TRF patterns did not exhibit any recognizable order or structure. That is, samples that were collected consecutively were not any more similar than two samples collected at random. Rather, the similarities between samples seem to fluctuate and display a dynamic character. These fluctuations can be a result of various external factors such as diet and stress. Even though a food diary and stress level chart was maintained, this experiment has not compiled enough data or support to help analyze the effects of external factors such as diet and stress.

Various precautions were taken to help obtain tenable and plausible data. In an attempt to alleviate biases in TRF patterns analysis, samples were analyzed in replicates. To exclude artifacts and produce reproducible peaks, samples were manipulated through dilutions to obtain patterns that surpassed the minimum dye signal threshold of 10K.

Although the subject studied was deemed stable, a much more comprehensive analysis of a larger sample size should be conducted to help accrue further data. It is possible that with enough data and studies, the notion of a stable or unstable intestinal flora can be extrapolated over a large population.

References

1. **Berg, R. D.** 1981. Promotion of the translocation of enteric bacteria from the gastroititestinal tracts of mice by oral treatment with penicillin, clindamycin, or metronidazole. Infect. Immun. 33:854-861.

2. **Bohnhoff, M., B. L. Drake, and C. P. Miller.** 1954. Effects of streptomycin on susceptibility of intestinal tract to experimental Salmonella infection. Proc. Soc. Exp. Biol. Med. 86:132-137.

3. **Clark, J. D.** 1971. Influence of antibiotics or certain intestinal bacteria on orally administered Candida albicans in germ-free and conventional mice. Infect. Immun. 4:731-737.

4. **Guarner, F., and Malagelada, J.** 2003. Gut flora in health and disease. The Lancet. 360: 512-519.

5. **Finegold, S. M., V. L. Sutter, and G. E. Mathisen**. 1983. Normal indigenous intestinal flora, p. 3–31. *In* D. J. Hentges (ed.), Human intestinal microflora in health and disease. Academic Press, Inc., New York, N.Y.

6. **Blay, G., Blaut, M., and Schwiertz, A.** 2000. Quantification of different *eubacterium spp.* In human fecal samples with species-specific 16S rRNA- targeted oligonucleotide probes. AEM, 66:375-382.

7. **Chen, K., H. Neimark, P. Rumore, and C. R. Steinman**. 1989. Broad-range DNA probes for detecting

* Sample days 1, 9, and 11

† (12,13,14,15,16)

and amplifying eubacterial nucleic acids. FEMS Microbiol. Lett. 57:19–24.

8. **Relman, D. A.** 1999. The search for unrecognized pathogens. Science 284:1308–1310.

9. **Woese, C. R.** 1987. Bacterial evolution. Microbiol. Rev. 51:221–271.

10. **Woese, C. R., E. Stackebrandt, T. J. Macke, and G. E. Fox.** 1985. A phylogenetic definition of the major eubacterial taxa. Syst. Appl. Microbiol. 6:143–151.

11. **Engelbrektson, A.** 2005. Analysis of treatment effects on the microbial ecology of the gastrointestinal tract. California Polytechnic State University of San Luis Obispo. Biological Sciences Department.

12. **Campanale, J., B. Bunyea-Baron.** 2004.16S Ribosomal RNA Gene Terminal Restriction Fragment Pattern Analysis Used to Assess the Stability of Bacterial Communities Found in the Human Gastrointestinal Tract. California Polytechnic State University of San Luis Obispo. Biological Sciences Department.

13. **Dinh, A.** 2005. Analysis of microflora in human gastrointestinal tract by using a PCR-base method called terminal restriction fragment (TRF) length polymorphism. California Polytechnic State University of San Luis Obispo. Biological Sciences Department.

14. **Karolys, M., Tang, E., and Schlegel K.** 2005. Investigation of microfloral stability within the human intestinal tract using terminal restriction fragment (TRF) length polymorphism analysis. California Polytechnic State University of San Luis Obispo. Biological Sciences Department.

15. **Sanchez, E.** 2006. Examination of the stability of baterial communities within the human gastrointestinal tract through terminal restriction fragment length polymorphism (TRFLP). California Polytechnic State University of San Luis Obispo. Biological Sciences Department.

16. **Cherepakho, V.** 2005. An investigation of microbial flora stability of the human gastronintestinal tract using terminal restiction fragment polymorphism analysis of the 16S ribosomal RNA gene. California Polytechnic State University of San Luis Obispo. Biological Sciences Department.