EXAMINATION OF THE STABILITY OF BACTERIAL COMMUNITIES WITHIN THE HUMAN GASTROINTESTINAL TRACT THROUGH TERMINAL RESTRICTION FRAGMENT LENGTH POLYMORPHISM (TRFLP)

By

Ellen Sanchez

Biological Sciences Department

College of Science and Mathematics

California Polytechnic State University

San Luis Obispo

2006

Examination of the stability of bacterial communities within the Human Gastrointestinal Tract (GIT) through Terminal Restriction Fragment (TRFLP) Length Polymorphism

ABSTRACT

This investigation seeks to contribute to the validity of a previous experiment conducted at the Environmental Biotechnology Institute that addressed the ecology of human microflora from the gastrointestinal tract. The study found that 17 out of 39 subjects in a clinical trial had less than 50% similarity in 3 homogenized fecal bacterial collections over a two week period. In this study, the stability of a healthy subject's fecal flora was assessed by collecting 14 consecutive fecal samples. The subject maintained a routine diet during the two week period and did not to take any antibiotics within 60 days prior to sample collection. Following fecal collection, DNA was extracted and the bacterial 16S rRNA genes were amplified by PCR, digested with *Hae*III, and used for Terminal Restriction Fragment Length Polymorphism (TRFLP) Analysis using a CEQ8000X DNA analyzer. TRFLP data was analyzed using Bray-Curtis Similarity. The average similarity of TRFLP data was greater than 50%. However, when three of the most different samples were compared, the subject displayed less than 50% similarity. The largest differences in fecal microflora occurred after a longer period of fecal retention (>1 day). This indicates that fecal microflora is usually stable on a daily basis over this short time period of 2 weeks. A more complete view of human fecal bacterial stability may emerge as this study expands to include more subjects.

INTRODUCTION

The human gastrointestinal tract is an ecosystem of diverse microbes. There are between 300-500 different species of bacteria estimated to colonize an individual's GIT including: *Acidaminococcus, Bacteroides, Bifidobacterium, Clostridium, Coprococcus, Enterobacter, Enterococcus Escherichia, Eubacterium, Fusobacterium, Klebsiella, Lactobacillus, Megamonas, Megasphaera, Peptostreptococcus, Proteus, Ruminococcus*, and *Veillonella* [1,2]. The acidity and bile in the upper GIT prevents many microbes from colonizing; however; the environment of the large intestine supports $10^{\wedge 11} – 10^{\wedge 12}$ bacteria cells/g lumen contents [1,3]. Even though it has been suggested that each person has unique strains of bacteria within his/her GIT, it is believed that the population of different bacterial species within an individual is relatively stable [1,3].

The bacterial communities within the GIT are of great physiological importance for the well-being of the human host. Gastrointestinal bacteria have metabolic, trophic, and protective functions [1]. The bacteria ferment undigested dietary products yielding short chain fatty acids which are used for energy storage [1]. Trophic effects include the promotion of intestinal endothelial proliferation and differentiation. Lastly, the bacterial microflora in the GIT have two protective functions. The interaction of the stable microflora with gut associated lymphoid tissue contributes to host immunity, and the presence of bacterial communities act as a barrier preventing the colonization of the intestinal lumen by pathogenic microbes [1]. In healthy individuals these metabolic, trophic, and protective effects are responsible for normal daily intestinal function of the host.

A previous study completed at the Environmental Biotechnology Institute found two distributions of data within the clinical control group, 60% of the subjects were stable, and 40% of the subjects were unstable (**Figure 1**) [4]. This experiment seeks to contribute to the validity of the control group from the previous investigation which analyzed TRFLP data from three

specific days (days 1, 7, and 14). Each day within the two week period was analyzed in this experiment to provided more information regarding the regularity of the presence or absence and relative abundance of TRFLP data.

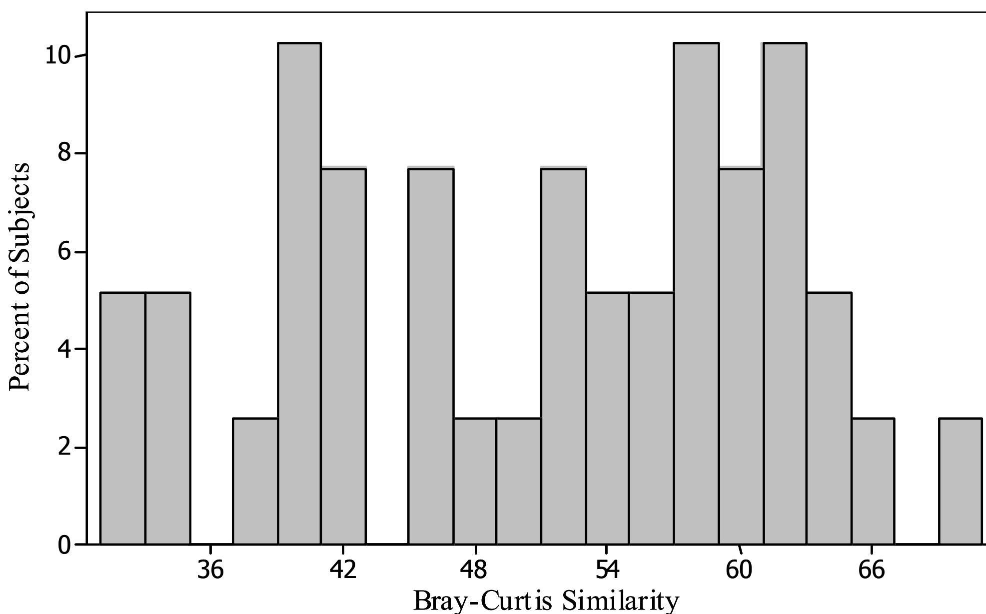

Figure 1: Average Bray-Curtis similarity of TRFLP patterns from each subject in a clinical trial across the three initial samples (days 1, 7, 14). Note the two distributions. An estimated cut-off for Stable/Unstable categories was made at 50% similarity. Subjects with average similarity >50% during the first three sampling days were deemed stable. (Figure taken from Engelbrektson [4]).

MATERIALS AND METHODS

Collection of Fecal Samples

Fecal samples were obtained from a healthy subject of 21 years of age each day over a fourteen day period. The subject was asked not to take any antibiotics 60 days prior to fecal collection. Samples were collected at the subject's home with a plastic wrap catching device designed to minimize sample contamination [5]. The subject's stress level was recorded, and samples were noted for qualitative characteristics, manually homogenized, and stored in a 1.5 mL centrifuge tube at 0°C for no more than five days when DNA was extracted.

DNA Extraction

Samples were extracted in triplicate with the MoBio Powersoil® DNA Extraction Kit according to the manufacture's protocol. The success of the extraction was checked with a 1.5% agarose gel electrophoresis.

PCR Amplification

PCR was performed using 16S rDNA primers homologous to Eubacterial conserved regions on the 16S rRNA gene. The forward primer 46-Ba2F (5'-GCY TAA CAC ATG CAA GTC GA-3'), and the reverse primer 536-K2R (5'-GTA TTA CCG CGG CTG CTG G-3'), were fluorescently labeled with phosphamide dye, and used in each reaction. Reactions were carried out using 3μL DNA, 5μL 10X buffer, 3 μL 10 mM dNTP, 2 μL 20mg/mL BSA, 7 μL 25 mM MgCl2, 1 μL 10 μM Ba2F, 1 μL 10 μM K2R, 27.7 μL PCR H2O, and 0.3 μL 5U/mL TaqGold®. Reaction temperatures and times were 94 ^{o}C for 10 minutes; 30 cycles of 94 ^{o}C for one minute, 46.5 ^{o}C for one minute, 72 ^{o}C for two minutes; and 72 ^{o}C for 10 minutes. The success of the PCR was determined through 1.5% agarose gel electrophoresis. All reactions were performed in triplicate and successful reactions were combined using a MoBio Ultraclean® PCR Clean-up Kit according to the manufacturer's protocol. The amplicon concentration was determined by A_{260} UV fluorometer following EBI protocol.

Enzyme Digest and TRF Pattern Generation

A 40-μL enzyme digest was performed on each PCR cleanup product using the New England Biolabs restriction endonuclease *Hae*III. Each digestion used 0.4 μL of enzyme, 4 μL of buffer, varying amounts of DNA and enough water to bring the final volume to 40 μL. All samples were digested for 4 h at 37 ^{o}C and deactivated at 80 ^{o}C for 20 min. Digestion products underwent ethanol precipitation and were re-suspended in 20 μL of formamide and 0.25μL of

CEQ 600 base pair standard. Terminal restriction fragment profiles were obtained using Beckman Coulter CEQ 8000X DNA analysis system.

Data Preparation and 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 [6]. TRF peak area was converted to relative abundance (percent of total area) manipulated by square-root to de-emphasize the larger peaks, and analyzed with a Bray-Curtis similarity matrix. Statistical analyses were performed using PrimerV and Excel.

RESULTS AND DISCUSSION

All TRFLP patterns from the *Hae*III digest were considered good data (**Figure 2**). A Bray-Curtis similarity matrix created in PrimerV was used to evaluate the TRFLP data. The average similarity between the TRFLP peaks for each day in relation to the other 13 days for this subject was 58.50% +/- 8.21. However, the average similarity of this subject on a daily basis (**Figure 3**) was 63.4% +/- 8.21. The subject barely met the 50% minimum stability requirement when the average similarity of all TRFLP data was considered, yet the subject exceeded the stability requirement on a daily basis. Analysis of the three least similar days (2, 7, and 12) yielded an average similarity of 45.2%. If Bray-Curtis similarity had been determined strictly from three days as in the previous experiment, this subject would have been considered unstable.

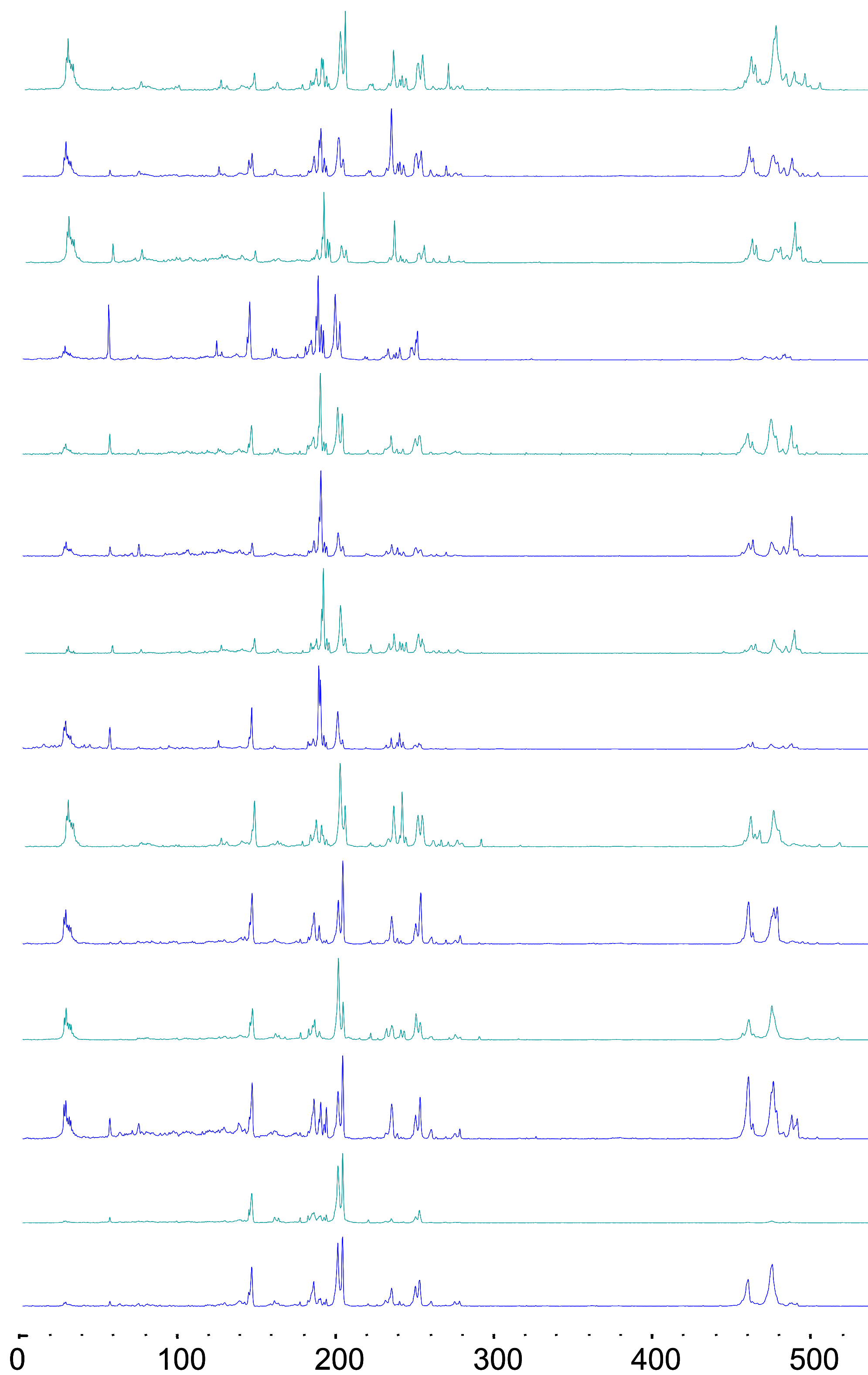

Figure 2: TRFLP data sequentially arranged with day one at the bottom through day 14 at the top. The intensity of the PCR labeled primer is represented by peak height, and the length of the *Hae*III digested 16S rRNA gene fragments in nucleotides is represented on the x-axis.

A daily comparison of data was performed using the Bray-Curtis similarity matrix. Days 2 and 3 were found to be the most divergent at 48.7% similarity, and days 5 and 6 were the most similar at 75.6% similarity (**Figure 3**).

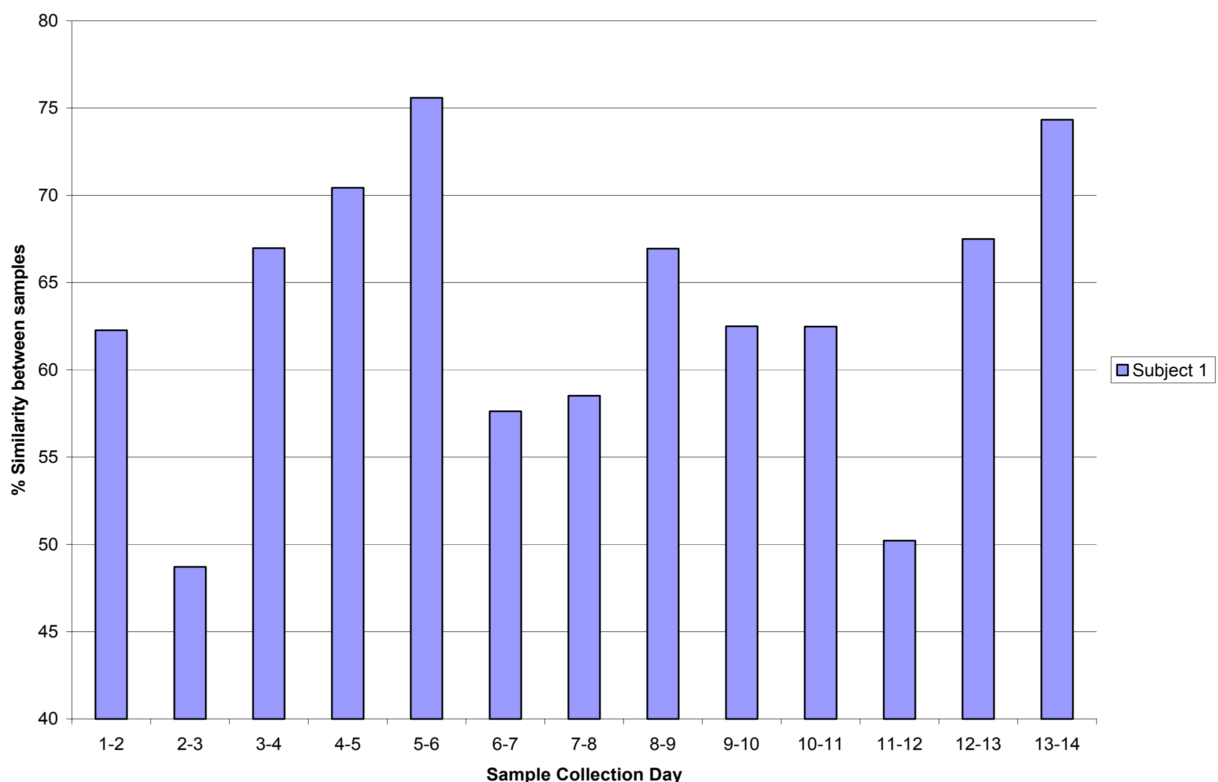

Figure 3: Comparison of Daily Progression based on Bray-Curtis Similarity matrix vs. sample collection day. Note the similarity between days 2 and 3 falls below the 50% stability requirement. The average similarity is greater than 50% indicating a stable subject.

A dendrogram was created from PrimerV to show the similarity of TRFLP data on a daily basis (**Figure 4**). The groups of samples that were most similar were clustered together and separated from other dissimilar sample groups. There was no observable relationship between the clustering and the order of days which indicated a random grouping of samples. Thus, the fecal microflora fluctuated randomly around an equilibrium on a daily basis.

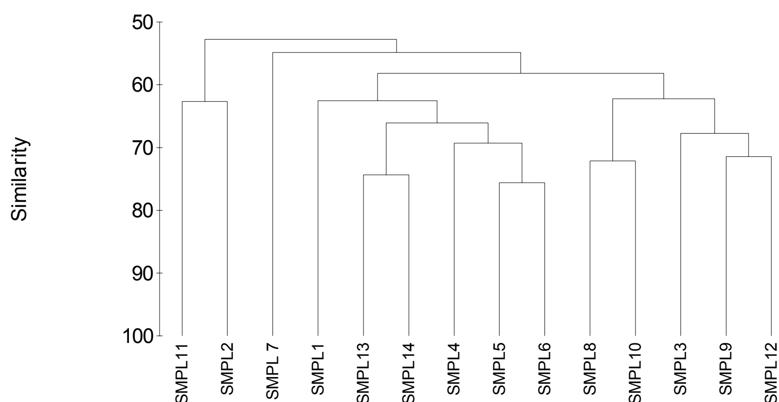

Figure 4: Dendrogram of daily samples from Bray-Curtis Similarity. The clustering of data indicates samples with greater similarity. There appears to be no obvious relationship between consecutive samples.

The peak persistence and relative size was determined to examine the structure behind the stable communities. Persistence is defined as the numbers of days in which a specific peak was observed. The peaks were then grouped according to their persistence. For example, all the peaks that appeared once were placed in persistence group one and peaks that appeared in all 14 samples were placed in persistence group 14(**Table 1**). The average peak area was calculated for each daily persistence value.

Table 1: Calculated Average Peak Area (%) from daily persistence (total number of TRFLP peaks for each group).

Persistence	Number of TRFLP peaks	Average Peak Area (%)
1	33	0.8
2	16	0.97
3	15	0.83
4	7	0.96
5	2	1.05
6	7	1.19
7	5	0.93
8	7	1.90
9	4	1.54
10	6	1.86
11	2	1.68
12	8	2.56
13	3	1.74
14	9	4.44

A bubble plot was created to display a visual relationship between peak persistence and average peak area for detailing the dynamics of bacterial fecal communities (**Figure 5**). Generally, peaks that were nearly present in all samples and peaks nearly absent in all samples had the highest number of TRFLP peaks. It appears that two distributions of bacterial communities exist: those in high abundance (large average peak area) that persist, and a fluctuating community with lower abundance.

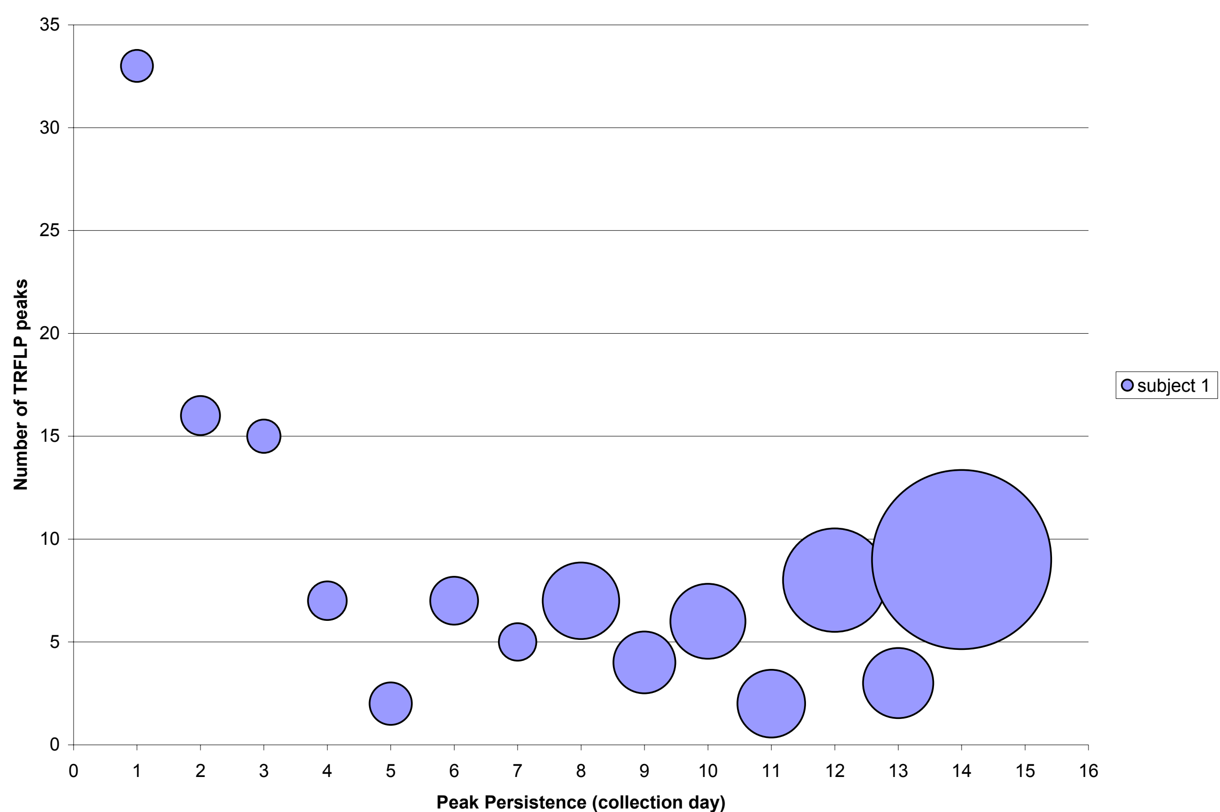

Figure 5: TRFLP peaks were grouped by persistence according to the regularity of presence in nearly all samples (bubbles at the right) or regularity of absence in nearly all samples (bubbles to the left). The average peak area is represented by the diameter of each bubble.

CONCLUSION

According to this analysis, the fecal communities of this subject are stable. Interestingly, the stability resulted from a random daily fluctuation of bacteria around an equilibrium. The bacterial communities with the largest abundance (large peak area) had the highest persistence and have a greater influence on the community stability.

In respect to the previous investigation, this subject met the 50% minimum stability requirement due to a change in the protocol. If the three least similar fecal samples (2,7, and 12) had been digested and analyzed, similar to the clinical study at the EBI, this subject would be considered unstable. It should be noted that this subject experienced a stomach ache on day six and took Tums antacid tablets. This may have affected the incubation of bacteria that were collected on day seven since the subject had routine morning bowel movements. However, there is no correlation within the subject's food diary to indicate unusual nutrition or stress levels with the low similarity between samples 2, 7, and 12. Generally, the largest differences in fecal microflora occurred after a longer period of fecal retention (>1 day). This indicated that on a daily basis, the fecal microflora is generally stable over a two week period. As this study expands to include more subjects a clearer view of the dynamics between bacterial communities within the GIT may emerge.

ACKNOWLEDGEMENTS

This investigation would not have been possible without the help, support and expert advice of many. Dr. Kitts, thank you for all of you advice and analytical contributions. A great thanks is do to Alice Hamerick who answered all questions that arose, helped with troubleshooting, and enlightened me with technique advice. Thank you to Anna Engelbrektson for being the muse for this investigation and being available both experimental and analytical advice. All of the members of the EBI deserve to be thanked for their helpful presence.

REFERENCES

1. Guarner, Francisco, and Juan-R Malagelada. 2003. "Gut flora in health and disease". The Lancet. 31: 512-519.

2. Isolauri, E., P V Kirjavainen and S Salminen. 2002. "Probiotics: a role in the treatment of intestinal infection and inflammation". Gut online. 50:54-59. 8 Jan. 2006. <http://www.gutbmjjournals.com>.

3. Holzapfel, Wilhelm H., Haberer, Petra, Snel, Johannes, Schillinger, Ulrich and Huis in't Veld, Jos H.J. 1998. "Overview of gut flora and Probiotics." International Journal of Food Microbiology. 41:85-101.

4. Engelbrektson, Anna 2005. Analysis of treatment effects on the microbial ecology of the gastrointestinal tract. California Polytechnic State University of San Luis Obispo. Biological Sciences department.

5. Bunyea-Baron, B. and Campanale, J. "16S Ribosomal RNA Gene Terminal Restriction Fragment Pattern Analysis Used to Assess the Stability of Bacterial Communities found in the Human Gastrointestinal Tract." 2004.

6. Beals, EW. 1984. Bray-Curtis ordination: an effective strategy for analysis of multivariate ecological data. pp1-49 in Advances in ecological research, v14 . A. McFayden, (Ed). Academic Press Inc., London.