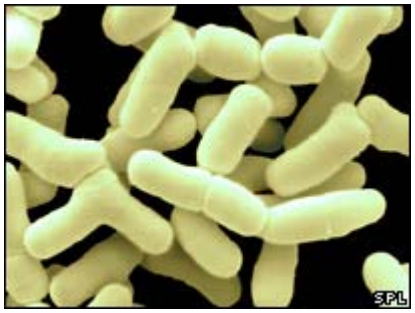

Bifidobacteria

16S Ribosomal RNA Gene Terminal Restriction Fragment Pattern Analysis: Assessing the Presence and Stability of Bacterial Communities of the Human Gastrointestinal Tract.

By Lisa Bandner

Abstract

Microbial community stability was addressed to determine the stability of microorganisms that exist in the gastrointestinal tract (GIT). Collection of 14 fecal samples occurred over a consecutive time period while the individual was eating a typical diet with no additional antibiotics. Microbial DNA was extracted from samples and the 16S rRNA gene was amplified using PCR. The primers obtained were then analyzed and observed using Terminal Restriction Fragment patterns. The patterns allowed for the visualization of fluctuations of microbial composition. The subject was found to have a day to day similarity over 50% similarity using Bray Curtis analysis. In addition, persistence of TRF peaks was used to determine the most abundant types of microorganisms, most of which were present in all samples.

Introduction

The complete gastrointestinal tract (GIT) is colonized soon after birth by a plethora of microbial species not only from the birth canal of the mother but by breast milk supplied as food[1,2]. The "adult-type" flora of the GIT gradually develops as the child begins to eat a regular diet[1,2]. These microbes are region specific within the GIT and are numerous and diverse. There are more than 400 species and subspecies that have been identified, the majority being gram-positive bacteria such as lactobaccilli, streptococci and yeast[2]. The GIT is known to harbor about 10^{14} viable bacteria, an amount that is ten times the number of eukaryotic cells found in all the tissues of the body combined[2]. There are only a few groups of colonic microflora, most of which are identified as strict anaerobes such as *Bacteroides*, *Eubacterium*, *Bifidobacterium* and *Peptostreptococcus*[2,3]. Members from the group *Bacteroides* constitute the most numerous groups of bacteria in the normal adult GIT[1,3,4].

There is a complex relationship between humans and the microbes that reside within them. These microbes are essential for overall digestive health. They are able to degrade certain food components, produce B vitamins, stimulate the immune system and produce digestive and protective enzymes for the individual[2]. The colonic mucosa is also dependent on short chain fatty acids that are supplied by the GIT flora and are a major source of useable energy and nutrients[2]. GIT microbes are able to provide a barrier to incoming pathogens and therefore discourage colonization by harmful bacteria. In exchange, they are awarded nutrients and food that enter into the human digestive system.

Recent research has found that numerous factors result in changes and variation in intestinal flora. Epidemiological studies on high and low risk colon cancer subjects indicate that the diet has some relationship to the composition of the flora[1]. An individual's intake of fiber has also been shown to possibly bind bacteria and control bacterial populations. Fluctuations in microbiota can also occur with changes or differences in host mediated factors and microbial interactions such as pH and adhesive properties[2]. There has been a discrepancy over microbial stability in the GIT and whether it fluctuates over time or if it is relatively stable for an individual. Studies suggest that even though there are fluctuations in the microbiota among different people, there should be a relative stability with small fluctuations from day to day in the same individual[3].

There are different methods of examining the composition of GIT flora, many which involve using fecal samples to quantify the presence of specific microbes as they are good indicators of colonic flora. Other methods involve the direct culture of diluted fecal samples to check for the presence of specific microbes[3]. This is a poor method for quantification since that at present less than 20% of all microorganisms can be cultured[2]. Current genetic methods use the 16S rRNA gene to study microbial populations more accurately. Polymerase chain reaction (PCR) and sequencing technology have aided in retrieving new sequences. In addition, 16S rRNA gene sequences are available in genetic databases for a multitude of bacterial species, which includes those intestinal bacteria that cannot be cultured. The gene sequence itself is highly conserved which also allows researchers to find phylogenetic relationships between bacteria[2]. The PCR products, which are suggestive of the specific microbial population's composition, can be further

processed and visualized by using denaturing gradient gel electrophoresis (DGGE), temperature gel gradient electrophoresis (TGGE), and also terminal restriction fragment (TRF) analysis[2,6,7].

The research performed builds on a previous study in the Environmental Biological Institute, at California Polytechnic State University, San Luis Obispo (EBI) sponsored by Danisco Corp. The study investigated the role of probiotics in the normalization of gut flora after treatment with antibiotics using TRF analysis. Although probiotic treatment after antibiotics showed a stabilizing effect, the population used to supply fecal samples showed a bimodal distribution of microbial similarity using Bray Curtis similarity (Figure 1).

Similarity is a concept used to determine whether the microbial populations in the fecal samples supplied were consistent between the majority of the days collected. Subjects that had an average sample similarity of 50% or greater were designated as stable whereas subjects that had an average sample similarity of less than 50% were designated as unstable. The similarity matrix created showed that 40% of the subjects were unstable while the remaining 60% of subjects were stable. The current protocol includes more complete data sets by having 14 consecutive samples rather than analyzing only days 1, 7 and 14. The same method of analysis using TRF was used in order to maintain comparable results with the previous study. Rather than focusing on specific microbial species present, the investigation focuses on the microbial communities present, their abundance, and their persistence or absence from sample to sample.

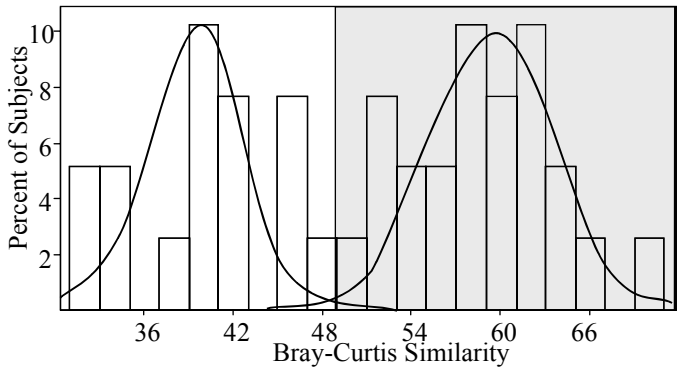

Figure 1: Average Bray Curtis similarity curve of TRF patterns for days 1, 7, and 14 for all individuals. Observe that there are two different distributions, one of which is below 50% and one of which is above 50%. This data indicates that subjects of the study with a similarity under 50% had unstable microflora. Subjects with a similarity over 50% for the first three samples were deemed stable.

Materials and Methods

Collection of fecal samples:

Samples were collected over a 14 day period at the individual's home. A collecting procedure designed to prevent contamination was used to collect samples. After collection, samples were homogenized and stored at 0 ºC for a period no longer than five days and then taken into the lab for DNA extraction.

DNA Extraction:

DNA extraction was performed for each sample in triplicate using the MoBio Powersoil® DNA kit following the manufacturer's specifications. The kit used a stationary matrix on a small filter in a 2 mL centrifuge tube for final extraction. After

DNA was obtained, agarose gel electrophoresis was used to determine the success of the extraction.

PCR Amplification:

PCR was performed by targeting a 500 base pair region of the 16s rRNA gene using primers homologous to Eubacterial conserved regions. Flourescently labeled forward primer Ba2f (5'-GCY-TAA-CAC-ATG-CAA-GTC-GA-3') and reverse primer 536-K2R (5'- GTA-TTA-CCG-CGG-CTG-CTG-G-3') was used in each sample. Each reaction was carried out using 1μL of a 1:10 dilution of DNA, 5 μL of 10x buffer gold, 3 μL 10mM dNTPs, 2 μL 20 μg/mL BSA, 7 μL 25 mM $MgCl_2$, 1 μL 10 μM Ba2F-D4, 1 μL 10 μM K2R, 0.3 μL 5U/ μL TaqGold® , and 29.7 μL water. A control sample using 3 μL E. coli instead of DNA was also used. Samples were exposed to reaction temperatures and times of 94ºC for 10 minutes; 30 cycles of 94ºC for 1 minute, 46.5 ºC for 1 minute, and 72 ºC for 2 minutes, and 72 ºC for 10 minutes. Agarose gel electrophoresis was used to measure the success of the PCR reaction. A MioBio PCR Ultra-Clean® kit was used to remove leftover salts, dNTPs, and primers using the protocol specified by the manufacturer. The Bio-Tek Fluorometer was used to determine the PCR product concentration measuring the Cy5 incorporated fluorescent label from the forward primer following EBI protocol.

Enzyme digests and TRF generation:

For each of the PCR samples, a New England Biolabs restriction endonuclease *Hae*III digest was performed using 0.4 μL enzyme, 4 μL of buffer, 75 ng of DNA, and water to

bring the sample to a total volume of 40 μL. The digest was performed for 4 hours at 37 ºC and cycled at 80 ºC for 20 minutes to deactivate the enzyme. The digest products were ethanol precipitated to remove excess salts and resuspended in 20 μL of formaldehyde and 0.25 μL of a base pair standard with a floating drop of mineral oil. Terminal restriction fragment (TRF) patterns were obtained using a CEQ 8000 DNA analysis system.

Data preparation and statistical analysis:

TRF results in nucleotides and peak area were exported from the CEQ 8000 into Microsoft Excel. Peaks that were less than 0.5% of the total DNA present were excluded from the analysis. TRF peak area was then converted to percent of the square root of the total area and analyzed with Bray-Curtis similarity. PrimerV and Excel were used to perform statistical analyses.

Results and Discussion

Fecal samples were collected over a 14 day period. These samples were thoroughly homogenized before storage to ensure that the collected sample would be representative of that day. A food diary was kept for each day where stress was also measured on a scale from 1-5. A color palate and hardness of each sample was recorded. TRF was used for the final evaluation of microbial composition of each sample by detecting the number of 16s rRNA PCR primers as well as the length of the nucleotide sequence.

Similarity between fecal samples:

Using PrimerV, a Bray Curtis similarity matrix was created. This method of analysis has suggested to be an ideal tool in the construction of similarity matrices. Bray Curtis allows for the testing of data sets to allow quantitative assessment of similarity, while removing subjectivity in comparing the data sets.

The similarity of one sample to the next was obtained by using the square root of TRF average peak area and comparing it to all consecutive samples. These data were used to create a daily progression graph (Figure 2) showing the percent similarity between samples of consecutive days. The most dissimilar comparison between two consecutive days was 56.3% between days 4 and 5. Looking back at the food diary created, I was not able to find any gastrointestinal upset or increased stress levels that would lead to a decrease in stability. A possible explanation may be a decreased intake of fruits and vegetables leading to a slower rate of turnover for samples. This could be due to the decreased intake of fiber since it plays a role in GI health by controlling

microbial populations by binding them and causing them to be passed along with other wastes in the colon.

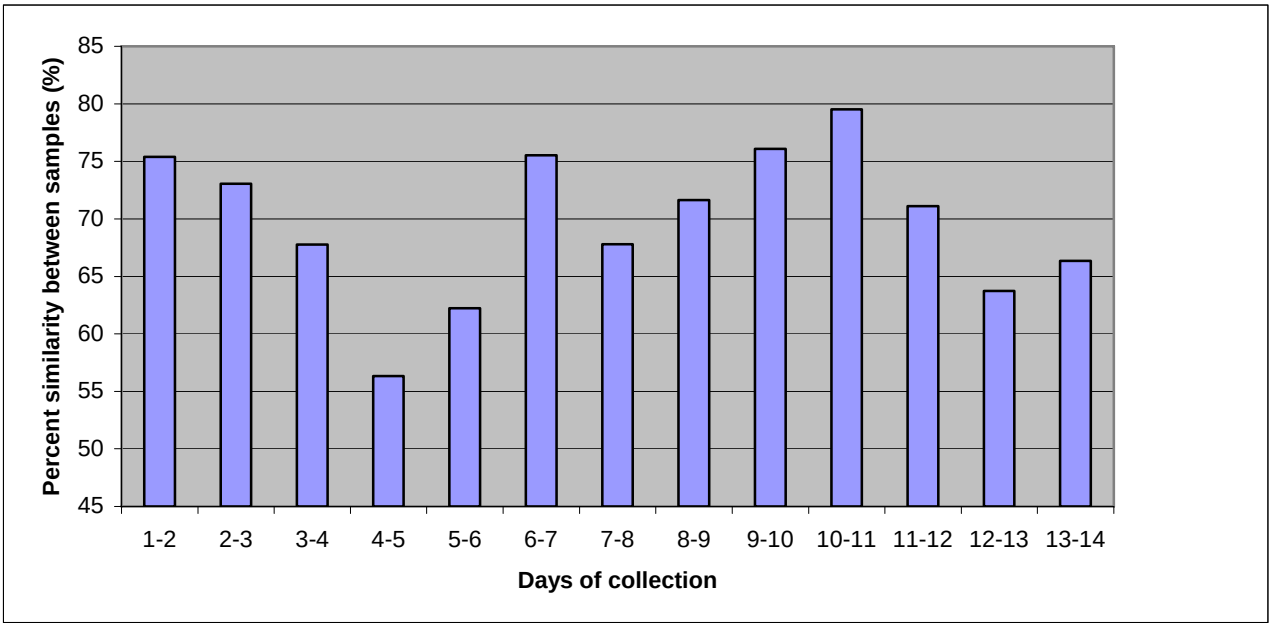

Figure 2: Daily progression bar graph using (Bray Curtis analysis) the average TRF peak area from consecutive days to depict percent similarity between samples. This type of graph allows for the comparison of two consecutive samples based on their similarity and how they compare to other adjacent samples. Notice that the similarity between days 4-5, 5-6, and 12-13 are comparatively low.

The average overall similarity was 66.2%. This value is over 50% indicating that the GIT is stable(Figure 1). In addition, to further support that the stability of the individual was not a fluke the average similarity of the three least similar days was calculated. It was found that the average similarity between these three samples was 53.9% which is still indicative of a stable GIT. This aids in showing that even the most dissimilar days appear to be stable. Therefore the subject was found to demonstrate no combination of samples that could be deemed unstable and below 50% using Bray Curtis similarity.

Bray Curtis Similarity Dendogram

In addition to the daily progression graph, a Bray Curtis similarity dendogram was created using the sample data to depict the percent similarity of the TRF patterns overall. The dendogram groups the samples into clusters to show the relationship of all the samples to one another other. This dendogram is also a way of visually depicting which samples are more closely related to others as well as those samples which are grouped out of the main cluster (Figure 3). For these results, the subject shows a similarity of 58%. In addition, these data show that samples 4 and 13 fall outside the main cluster indicating that they are the most similar to one another but also the least similar to the remaining samples. These data show that as similarity increases, the samples which are most similar to one another are random and in no particular order or sequence. This suggests that patterns in GI fluctuations are absent and instead vary randomly around a point of equilibrium.

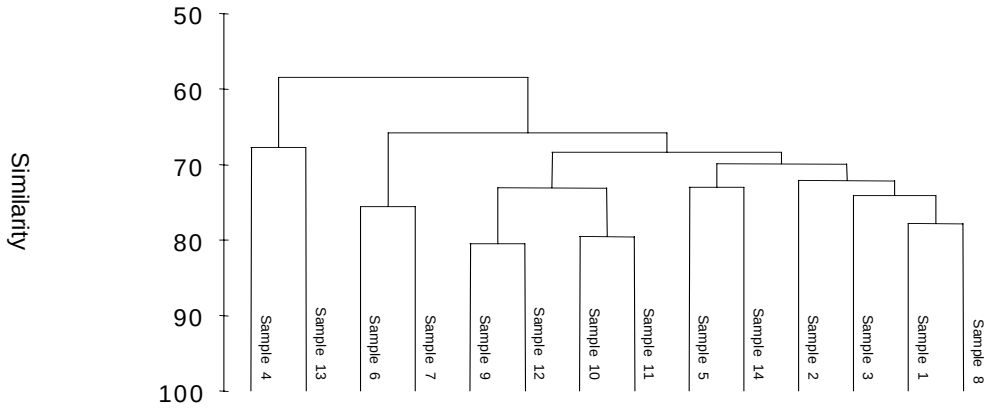

Figure 3: Dendogram depicting the relationship between all samples based on the average area of TRF peaks. Notice that there is no correlation between sample number and clustering position. Observe that there are no more than 2 consecutive samples with a similarity higher than 75%.

Individual TRF Peak and Persistence:

The persistence of TRF peaks was evaluated by observing a particular TRF peak and the number of samples in which it was present. The total number of peaks were then grouped based on their persistence. For example, if peaks showed up in one sample, these would be placed in one specific group designated 1 while those present in all samples were placed in another group designated 14. The peak areas of each group were averaged (Table 1). A bubble graph of TRF peak persistence also depicts a visual

representation of the microbial populations and fluctuations which occur in those populations of the GI tract (Figure 4).

Number of Samples a peak was present in	Total Number of Peaks	Average Peak Area (%)
1	23	0.77
2	20	0.90
3	7	0.77
4	4	0.90
5	2	0.73
6	4	1.10
7	3	1.00
8	3	2.47
9	3	1.10
10	4	1.58
11	4	2.86
12	7	1.67
13	3	2.71
14	11	4.78

Table 1: Shows the presence of specific peaks among total samples and average peak area for each group.

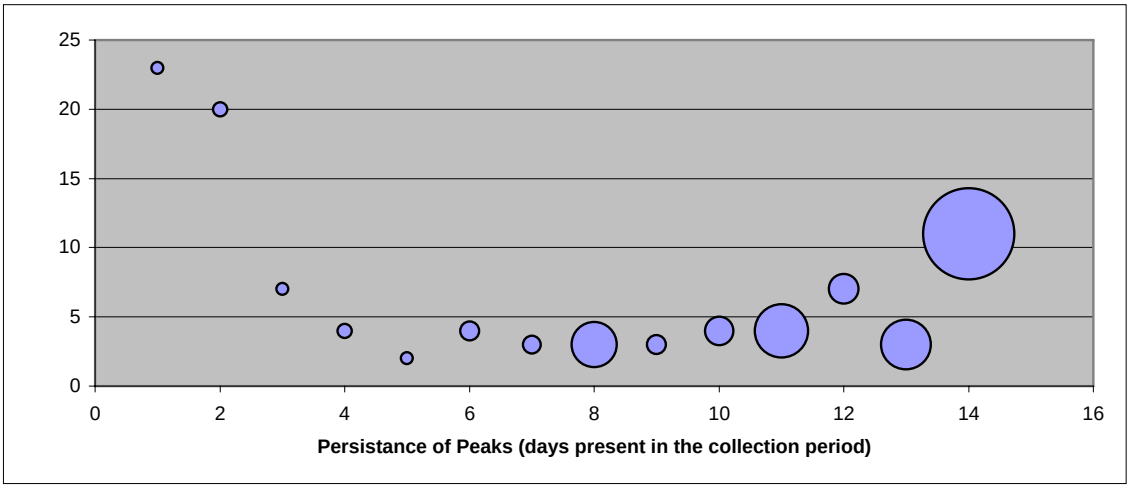

Figure 4: Bubble diagram which represents the persistence of peaks within a given number of TRF patterns and the average area of those peaks using a circle of given diameter. This visual representation shows that there were many single peaks detected in very few samples on the far right and peaks that were present in nearly all samples on the far left. The average area of each group of peaks is represented by the diameter of the circle plotted on the graph.

These data have two interesting features showing that there were two main populations observed when the number of TRF peaks was plotted against persistence: These data

show that there were multiple peaks existing in few samples (left) and multiple peaks that were present in nearly all samples (right). These two extremes had the greatest number of TRF peaks present, suggesting that the two groups having the highest abundance of peaks present represent two distinct populations of microflora. The microbial populations present in nearly all samples are most likely the stable flora which is indicated by the large average peak area. The second microbial population seems to show up sporadically which is observed by the low number of sample persistence but the high number of TRF peaks observed. The average peak area for these samples is less than 1% indicating that the increased number of TRF peaks in groups of low persistence could be an artifact of the detection limit. As indicated earlier, most bacterial populations fluctuate around some equilibrium point, but since these populations have such a low abundance and fluctuate around a 1% threshold it is possible that their detection is also sporadic. The microbial populations that are observed with high abundance do not fluctuate so close to a threshold detection value and therefore are more readily detected even though there are natural fluctuations in their populations as well. As a result of the low persistence and average area, the microbial populations present with low abundance do not directly affect the measure of similarity or stability. This can also be observed (Figure 4) by comparing the average area of the low persistence peaks with those of high abundance.

Conclusion

The final results showed that the subject had an average percent similarity of 66.2% using Bray Curtis similarity. This indicates that the subject's GIT flora is stable over time. In conclusion, the results showed (Figure 3) that the stability of flora fluctuates around a point of equilibrium. These fluctuations are due to changes in those populations that showed the greatest TRF average peak area (Figure 5). Those organisms with high abundance and persistence make up a larger percentage of the total population. Therefore fluctuations of these high abundance organisms are most likely the contributing factor to the stability of the individual's GI tract.

In comparison to the previous study, the subject had a similarity greater than 50% using Bray Curtis similarity. The results from this subject are not directly comparable to the previous study or significant due to the fact that only one subject was tested in this investigation. It is possible that this individual is consequently stable whereas others may not be. Samples were taken over 14 days and incubated at 0 ºC over a period of less than five days. In the previous study, samples were stored at -80ºC until shipment in which case they could be stored up to 45 days before DNA extraction took place. The sample collection and storage differences could have accounted for some changes or fluctuations in microflora. In addition, if improper homogenization occurred then the sample may not have been consistent throughout resulting in sampling error.

Acknowledgments

Thank you to Dr. Kitts for giving me the opportunity to gain extra experience to further my educational career. Thank you also for your patience, especially during final analysis and input in regards to finishing the actual write-up of this investigation.

Thanks to all EBI staff for allowing me to share their work space as well as for answering my questions whenever I needed help.

Special thanks to Alice Hamrick for taking the time to show me proper protocol techniques and procedures. Your patience and helpful advice were much appreciated.

References:

[1] Floch, Martin H. 1974. Human Gut Microecology. Western Jour. of Med. 5: 423-424

[2]Wilhelm H. Holzapfel, Petra Haberer, Johannes Snel, Ulrich Schillinger and Jos H. J. Huis in't Veld. 1998. Overview of Gut Flora. Inter. Journ. of Food Microbio. 41: 85-101.

[3]Zubrzycki, Leonard, and Earle H. Spaulding. 1961. Studies on the Stability of the Normal Human Fecal Flora. Department of Microbiology, Temple University School of Medicine. 83: 968-974

[4]McCartney, Anne L., Wang Whenzhi and Gerald W. Tannock. 1996. Molecular Analysis of the Composition of the Bifidiobacterial and Lactobacillus Microflora of Humans. Am. Soc. for Microbiol. 62: 4608-4613.

[5]Poupard, James A., Inistar Husain and Robert F. Norris. 1973. Biology of Bifidobacteria. American Soc. for Microbiol. 37: 136-165.

[6]Kitts, Christopher L. 2001. Terminal Restriction Fragment Patterns: A Tool for Comparing Microbial Communities and Assessing Community Dynamics. Intert. Microbiol. 2(1): 17-25

[7]Engelbrektson, Anna. 2005. Report to Danisco.

[8]Gavin N. Rees, Darren S. Baldwin, Garth O. Watson, Shane Perryman and Daryl L. Nielsen. 2005. Ordination and significance testing of microbial community composition derived from terminal restriction fragment length polymorphisms: application of multivariate statistics. Antonie van Leeuwenhoek. 86: 339-347