2011

Utilization of a Custom-Designed Microbiota Array to Determine the Distal Gut Microbiota of Healthy Human Adults

Richard Thomas Agans
Wright State University - Main Campus

Utilization of a Custom-Designed Microbiota Array to Determine the Distal Gut Microbiota of Healthy Human Adults

A thesis submitted in partial fulfillment of the requirements for

the degree of Master of Science

By

RICHARD THOMAS AGANS

B.S. Ohio Dominican University, 2008

2011

Wright State University

Wright State University

School of Graduate Studies

Date: Feb. 07, 2011

I HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER MY SUPERVISION BY RICHARD THOMAS AGANS **ENTITLED** Utilization of a Custom-Designed Microbiota Array to Determine the Distal Gut Microbiota of Healthy Human Adults **BE ACCEPTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF** Master of Science

_____

Oleg Paliy, Ph.D.
Thesis Director

_____

Barbara Hull, Ph.D.
Program Director, Microbiology and Immunology
M.S. Graduate Program

Committee on
Final Examination

_____
Oleg Paliy, Ph.D.

_____
Mill Miller Ph.D.

_____
Nancy Bigley Ph.D.

_____
Andrew T. Hsu, Ph.D.
Dean, School of Graduate Studies

Abstract

Agans, Richard Thomas. M.S., Microbiology and Immunology M.S. Program,Wright State University, 2011. Utilization of a Custom-Designed Microbiota Array to Determine the Distal Gut Microbiota of Healthy Human Adults

The human microbiota is an essential component of human health and disease. It is involved in metabolism of dietary components and is at the forefront of the intestinal immune response. Classical techniques applied to study intestinal microbiota illustrated higher presences of aerobic and facultative-anaerobic bacteria, however; the levels of obligate anaerobes had been underrepresented. Modern technologies based on DNA and RNA analysis have circumvented previous challenges allowing researchers to gain more extensive insight into the complex intestinal environment. In this work, a recently developed Microbiota Array was used to assess intestinal microbiota of 10 healthy adults (age 22-61, ave. 34.3yrs). Class level results showed dominance of *Clostridia* in all samples, encompassing approximately 74.0% of total class signal. Little variation was seen among samples at the Class level, however, this variation increased at the Genus level. *Ruminococcus* was the most abundant genus (21.1%), followed by *Faecalibacterium* (8.7%) and *Papillibacter* (6.8%). Members of 87 genera were detected in at least one sample. Among these, most were present at low abundance levels (65 genera had average abundance less than 1%). A core microbiome of 113 phylo-species was found in every sample. Among these core phylo-species, *Ruminococcus*, *Roseburia*, and *Papillibacter* were the major contributors. Overall, similar numbers of phylo-species were detected across samples (294-385, ave. 330). The results here show that the Microbiota Array is an efficient and viable method for analyzing the intestinal microbiota. In conclusion, this work solidifies the Microbiota Array as a formidable tool in analysis and characterization of intestinal microbiota.

Table of Contents

Pages

List of Figures

Page

List of Tables

Page

Acknowledgements

The work on my thesis has been long and difficult, it is the culmination of many hours of physical labwork mixed and longer hours spent in front of the computer learning to analyze data. I would like to thank my advisor, Dr. Oleg Paliy for his continued guidance and mentorship. I would also like to thank Dr. Paliy for not giving up on me, even when it seemed to me (and perhaps him as well) that I was a lost cause. Through his tutelage I have learned quite a lot about what it takes to be a good researcher and a good scientist. He has always been available to answer my questions and address my doubts. Dr. Paliy leads by example and his hard work and dedication show through his continued pursuit of good research. I would also like to thank my labmate Laura Rigsbee, who has been working on similar projects and has been a source help and a good friend. We have had quite a few laughs, brain storms, problem solving talks, and plain old good times.

I'd like to thank my parents Tom and Linda, for never letting me take the easy way out. I want to thank them for their never ending support and for letting me practice explaining my work on them. I want to thank my older sister Krystle, who has always been my role model, and who has always given me sound guidance and advice. Lastly, I want to thank my wife Sarah. Through all my ups and downs, all the successes and defeats she has been by my side, never letting me falter, never letting doubt and sadness linger. She has helped me through more than she knows and without her this work would have taken much longer.

Dedicated to my parents Tom and Linda Agans, my sister Krystle Agans and to Sarah, my wonderful and loving wife.

1. INTRODUCTION

Microbiota

Microorganisms inhabit a plethora of environmental niches, from soil, water (salt and fresh), air, and plant surfaces, to organisms such as insects, mammals, birds, reptiles, etc. [1]. There has long been much interest in studying this minute world of life, exemplified by the vast number of studies conducted on microbiomes inhabiting environments within the human body [1].

There have been numerous studies investigating the microbiota of human body niches [2-7]. Studies examining the skin microbiota of humans have shown variations between different body niches. One such report by Costello et al. (2008) observed differences between the hands, head, legs, arms, etc.; it was determined that microbiota inhabiting right and left sides of the body tended to cluster together with the exception of index fingers which were more related to each palm. These authors found differences across niches and attribute those differences to local environmental characteristics of each skin area. These authors go on to discuss how bacterial diversity relates to body habitats of the skin, gut, auditory, and oral tissue; and found that diversity decreases from upper body skin habitats with gut, oral, lower body skin habitats and auditory tissues following in that order. Human skin is covered with prokaryotic organisms, Whitman et. al point out, "the density of prokaryotes is about 10^3-10^4 cells/cm^2, except in the groin and axilla where it is 10^6 cells/cm^2. Based on an average adult human's skin, the total number of prokaryotes on the skin of an individual is about 3 X 10^8 cells" (1998). The observed diversity in such niches could possibly be related to exposure to external environmental conditions and stimuli. The community observed by Costello's group consisted of the phyla Firmicutes, Actinobacteria, Bacteroidetes, and Proteobacteria. The oral and stomach microbiota have been shown to differ in composition compared with skin and gut microbiota, interesting

since the mouth has close proximity to the skin epithelium and stomach contents are sent through the gut. One such study aimed at determining bacterial profiles in healthy oral cavities claimed that of five sites sampled within the mouth there was a common presence of genera from the classes Bacilli and Clostridia [7]. Interestingly, it was determined by culture-independent techniques that *Helicobacter pylori*, a common gastric pathogen, is present in healthy stomachs; however, the authors were unable to concretely determine whether this presence was due to potential contamination, transient organisms, or cell remnants [6]. While these authors do indeed show that *H. pylori* was present in their samples, the inability of being able to determine the extent that potential contamination plays in this observation suggests that broad range polymerase chain reaction (PCR) based approach is not advantageous.

The intestinal microbiota is responsible for processes including nutrient digestion of dietary components, immune modulation, and pathogen resistance, among others [8]. It has been estimated that the human intestine contains upwards of 10^{13} bacterial cells, with a combined genome larger than that of the human host [1]. The data collected in these early studies indicate the intestinal microbiota is integral part of human health and disease.

The Human "Super-Organism:" Symbiosis in the Intestinal Tract

A constant bacterial presence in the intestinal tract has resulted in a relationship between the host and bacteria that is beneficial for both parties. This relationship is evidenced by the presence of many bacterial cells at any given time in the intestinal tract without significant immunological action or challenge. Therefore, there must be symbiotic relationships allowing these bacteria to be present in such numbers on a constant basis, and this symbiotic relationship is found upon understanding the roles that microbiota play in human digestion. It is well known that most of the body's metabolism and absorption of dietary components takes place in the

intestinal tract. Bacteria are able to break down pectin, hemicelluloses, starches, cellulose and other polysaccharides, the byproducts of which, can be used by the human host, along with other bacterial species for biochemical processes [9-19]. This ability to break down certain dietary products is specific to certain microbiota and is a result of complex evolution and gene transfer, allowing for such abilities to be nonobligatory for the human host.

One example of such relationships comes from Xu et al. [20] who looked at two species of intestinal *Bacteroides* and observed multiple instances of similar protein-coding genes being inherited by the microbiome, proposing that these genes were laterally transferred due to the genes differing composition from that of the bacterial genomes. The presence of genes found in the *Bacteroides* species were shown to be involved in metabolism of plant polysaccharides and proteins, as well as glycans associated with the intestinal epithelium mucous layer. The group went on to suggest that there are two main forces driving the establishment of symbiosis in the intestinal environment, a "top-down" selection via the host to produce a homogenous microbiome with a "bottom-up" selection between microbes to produce a differentiated genome. These selection processes result in a functional microbiota that has established symbiosis with the human host.

Li et al. [21] were investigating relationships between microbiota structure with metabolite Nuclear Magnetic Resonance (NMR) data. They employed denaturing-gradient gel electrophoresis (DGGE) to differentiate bacterial groups, and NMR to assess urinary tract metabolites. These authors reported correlations between groups detected through DGGE and the NMR results. They suggested that their results would lay the groundwork for a "Rosetta Stone," to understand the intestinal microbiome and relations to human hosts. Along with these works

there have recently been many others detailing symbiotic, evolutionary and, mutualistic relationships between host and bacteria as well as between bacterial species [22-25].

These works begin to illustrate the vast significance of microorganisms in the human host. To fully understand human health and disease, a proper knowledge of microbiota/host relationships (cell-cell communication, bacterial numbers, activity, etc.) is needed.

Bacteria in the Intestinal Tract: Relations with Host and Each Other

Sghir et al. state that there are on average around 10^{11} bacterial cells for every gram of fecal content. This statement is further solidified by Dethlefsen et al. who explain that microbes inhabiting the colon outnumber human cells by 10-fold, and contain a combined total of 100-fold more genes than that of humans. When breaking down the microbial community it has been shown that the majority of this population is made up of the phyla Firmicutes, Actinobacteria, and Bacteroidetes [4, 26].

Bacteria in the intestinal tract are involved in many processes and interactions with each other and the human host as well. As a result of reactions between microorganisms and the host, a person's immune system has to deal with the sheer numbers of foreign and commensal microorganisms. In response to the persistent presence of bacteria, albeit commensal, the human immune system has developed mechanisms to keep bacteria in check. The immune system maintains this balance through production of mucus, antimicrobial proteins, activation of immune cells, and clearance of bacterial cells [27]. Goblet cells lining the epithelium produce mucus, which serves two functions, 1) to act as an external layer to which bacteria bind, keeping them from invading the simple epithelium, and 2) an internal layer containing a high concentration of antimicrobial peptides, that act against bacteria upon contact [27]. Another mechanism by which

the human body recognizes bacteria (commensal/pathogenic) is through the binding of pattern recognition receptors to bacterial cell wall components, triggering a downstream response to produce antimicrobial proteins or activation of adaptive immune cells i.e. B and T cells.

The body is also proactive in the production of a certain class of antimicrobial peptides known as defensins, more specifically the alpha-defensins. These defensins are constitutively expressed and released into the lumen of the digestive tract. Upon recognition of bacteria defensins, punch holes in the membranes of microorganisms, further keeping bacterial numbers manageable [28-30]. These immune products represent only a small portion of the many mechanisms that are involved in the human innate response to bacterial presence.

Along with the body's response to bacterial stimulus, there also exists cross-talk between bacterial cells in the intestinal tract. One such mechanism of cross-talk, known as quorum sensing, allows bacterial cells to regulate expression of certain genes as a function of population density [31]. Compounds called autoinducers, which can be produced by bacterial species of the same genus or by those from an unrelated genus, are key in keeping communities fit by keeping communities from growing beyond their ability to maintain themselves. This mechanism is beneficial for the microorganism in relation to the immune response in that this sensing allows the community to grow to a point that it is able to function and thrive without initiating an immune response by growing too large [32]. The body is remarkable in the relationships it has evolved with the intestinal microbiota, and this is only a small portion of the interaction between microorganisms and the human intestinal systems; to go into more detail would be beyond the scope of this literature.

As Pai and Kang explain, the intestinal microbiota is equipped to metabolize dietary components unusable by the host until broken down into smaller constituents. These bacterial species utilize nutrient sources including plant polysaccharides such as cellulose, xylan, and pectin. They are also actively involved in absorption of lipids, recycling of nitrogen, production of vitamins, etc. [33]. This is further explained for *Bacteroides thetaiotaomicron*, describing how the bacterium has more than double the glycoside hydrolases as the human host along with eight starch utilization genes. These authors also discuss the production of short chain fatty acids and absorption by host cells; in a culminating statement the authors describe an "eating together" relationship further illustrating symbiosis in the intestinal tract. The evidence in these works further suggest that the intestinal microbiota is crucial in proper homeostasis.

Major Intestinal Phyla

Although the intestinal microbiota contains more than 100 different bacterial genera, encompassing more than 1000 different species, all these species are members of only a few phyla. There are many publications that discuss the distribution of bacteria in the intestinal tract, indicating that the phyla Firmicutes, Actinobacteria, and Bacteroidetes dominate the intestinal microbiota [4, 26]. Notable among these phyla is the ratio between the Firmicutes and Bacteroidetes, a ratio that has received attention from researchers. It appears that this ratio "evolves" in connection with a person's life. This ratio was seen to increase from infancy to adulthood and then decrease again [34]. Further investigation into this ratio may potentially lead to an alternative diagnostic approach for assessing intestinal disease.

Prominent Bacterial Genera in the Intestinal Tract

At the genus level, it has been shown that *Ruminococcus, Bacteroides,* and *Faecalibacterium* are the most prominent. Of particular interest, is the genus *Faecalibacterium*;

which has been shown to aid in immune remediation of colitis, in addition to its involvement in metabolism [19, 35]. Sokol et al. looked at *Faecalibacterium prausnitzii* in a trinitrobenzene sulfonic acid (TNBS) model of colitis and observed that both *in vitro* and *in vivo* models resulted in a decrease in colitis. This remediation was attributed to *F. prausnitzii*'s ability (through production of metabolites) to induce IL-10 (anti-inflammatory cytokine) production, and decrease production of both IFN-γ and IL-12 (pro-inflammatory cytokines). Along with the diminished effects of the colitis model, *F. prausnitzii* was shown to rectify the dysbiosis that was created as a result of said colitis [35]. This effect of *Faecalibacterium* is further solidified by reviewing literature addressing the effects of butyrate on immune function which itself produces an anti-inflammatory effect by increasing IL-10 production, along with this promotion of anti-inflammatory cytokine stimulation it has been shown that butyrate decreases levels of proinflammatory cytokines by inhibiting NFkB [36]. This organism has been shown to be a prominent producer of butyrate in the normal diet as well as a major player in utilization of acetate [19, 37].

The involvement of *Bacteroides spp.* in the intestine has been well documented [10, 11, 38-40]. The increased presence of *Bacteroides* in vegetarian diets is indicative of its ability to hydrolyze plant polysaccharides, further evidenced by the prediction of 226 glycoside hydrolases in *Bacteroides thetaiotaomicron* compared to 96 known hydrolases in the human genome [41]. This organism is also known to be efficient in its utilization of starches in the diet, containing eight starch utilization genes which allows the organism to break down dietary starches into short chain fatty acids such as acetate, butyrate and propionate; which can then be utilized by *Faecalibacterium*, other microorganisms and human colonocytes [33, 37]. Aside from *Bacteroides spp.* being a crucial component in proper nutrient extraction it is also known to be important in proper development of the post-natal immune system [42]. It has been reported that

germ free mice, which do not contain a developed immune system, can have this lack of proper immune function corrected through administration of either harvested microbiota from conventional mice or *Bacteroides thetaiotaomicron* [42]. Their results are important because they highlight the importance of microbiota members in maintaining proper intestinal health.

Another constituent of the intestinal microbiota is the genus *Ruminococcus*. *Ruminococcus spp.* are part of the Firmicutes phylum , which has been shown to be one of the major phyla in the human colon [38, 43, 44]. *Ruminococcus* species contain cellulolytic complexes that enable these organisms to degrade plant cell wall polysaccharides, which is essential for further processing of plant components [44]. A benefit of these complexes is the allowance for *Ruminococcus spp.* to attach to the particle-phase of intestinal luminal content, creating a close association between organism and nutrients further enabling propagation and community formation [9, 44]. It has been documented that the Firmicutes phylum has higher presence when characterizing particle-phase luminal content, which further suggests the role of *Ruminococcus spp.* in the human diet in the essence that *Ruminococcus spp*. make nutrients available for other bacteria which make energy sources for human colonocytes. [9, 38, 44, 45]. In fact, Abell et al. show that *Ruminococcus bromii* and related phylotypes were increased in patients receiving a diet that was high in resistant starches [46].

Initial Colonization and Structure of the Intestinal Microbiota

Bacteria and other organisms do not spontaneously appear in the human gastrointestinal tract, they must originate externally. Tlaskalova-Hogenova et al. state, "[s]tarting from the first hours after the delivery from the sterile uterine environment…the interaction of the macro-organism with micro-organisms begins: the main portal of entry of microbes is skin and mucosal surfaces of the gastrointestinal, respiratory and urogenital tracts." [47] It is known that infants are

exposed to bacteria when contacting birth canal and mothers vagina upon delivery which results in an infant exposure mainly to aerobic species [48]. Palmer et al. performed a study aimed at determining a healthy microbial community structure in infants. This study addressed two aspects of microbiota interrogation by assessing not only the microbiota of newborns, but also determining the difference in analysis through pyro-sequencing and microarray analyses. To address the latter first, they were able to show that both microarray and pyro-sequencing gave extremely similar results. These authors note that of all their obtained samples, two belonging to a pair of twins delivered through caesarian-section, had relatively low bacterial counts compared to their natural born counterparts (2007). These findings are in agreement with previous publications showing that differing modes of delivery result in altered bacterial communities [49-51]. The work done by Palmer's group focused on a long-term investigation, analyzing fecal content up to one-year post delivery. They showed that anaerobic bacteria start to take over colonization up through the first year of life, and that at this point the microbiota is strikingly different than earlier months of development, and more similar in structure to the general adult microbiota. As was previously stated these studies have illustrated the difference in infant microbiota shortly after birth and later on during infancy. It appears that the "creation" of this adult-like community is associated with an infant's consumption of solid foods [48]. Lastly, while it is documented that delivery method, infant feeding type (breast versus formula), and antibiotic use during infancy affect the microbial structure; it is agreed that young children possess an intestinal microbiota similar to that of adults consisting mainly of anaerobic groups (*Clostridia, Eubacterium, Proteobacteria*, etc.) [48-52].

Aging and Microbiota Structure

The adult microbiota is another topic of interest among researchers. In comparison studies between elderly (>=60yr) and adult (20-50yr) sample groups, there appears to be differing

opinions that *Bacteroides/Prevotella* species decrease in the elderly population, along with similar decreases in *Bifidobacteria* species as well, while other studies report opposite findings. It is possible that these contradictions are due to differing experimental methods (use of fluorescent *in situ* hybridization, sequencing, classical microbiological techniques, etc.), as well as variation associated with geography, diet, and subjects themselves. [17, 53-57] For instance Claesson et al., report an increase in *Bacteroides*, along with an observed increase in *Faecalibacterium* in elderly subjects compared to younger subjects (20-25% and 5-8% compared to 5-8% and 4-7% respectively). In comparison, Enk et al. reported differing results claiming a decrease in *Bacteroides* numbers; however this study was done using classical techniques that under represent obligate anaerobes and organisms that remain difficult to extract and cultivate. While variation does exist between subjects at a genus level, there is evidence to support the notion that at higher taxonomic levels there exist more stability [53, 57]. The majority of previous reports stop analysis at higher phylogenies, which might be explained by observed variation at lower phyolenetic levels or the inability of the techniques utilized to assess these lower levels.

External Factors Influencing Intestinal Microbiota Composition

While the intestinal microbiota has been shown to be stable long term [58], it is widely accepted that this community can be influenced to some degree by host genotype, diet, and immediate environment. These studies document the influence of diet, host genome, and surrounding environment on the intestinal microbiota.

Recently, studies assessing the microbiota related to the human genome have focused on possible differences in twins, both with respect to inflammatory bowel disease and obesity [59-62]. These studies found that monozygotic and dizygotic twin pairs shared a microbiome as well

as representative metabolic pathways that were in most cases significantly more similar to each other than unrelated individuals [59-61]. Turnbaugh et al. even went so far as to show that these twins shared greater similarity in their microbiota to their mother than to unrelated individuals (2009). Although results from inflammatory bowel disease reports do not claim any significant effect of the host genotype to susceptibility to intestinal disease, there remains a consensus that host genotype relates to microbiota [59-62]. Zoetendal et al., in 2001, examined the intestinal microbiota of monozygotic twins, marital partners, unrelated individuals, and non-human primates using denaturing-gradient gel electrophoresis, and noted a positive correlation between shared genetic makeup and microbiota. The authors observed twins had the highest degree of similarity in their microbiota composition, and with the exception of a few high similarity observations, marital partners and unrelated individuals showed no significant difference in degrees of similarity and were significantly lower than twins [63]. The results of all of these studies suggest that host genotype plays an important role in microbiota makeup; however, the degree of such affect has yet to be determined.

Another subject of interest is the relationship between diet and intestinal microbiota. This subject is an area that has been widely studied, and although there has been no conclusive argument that the intestinal microbiota is an effect of a causative diet, there is ample evidence to suggest this [9-12, 14, 16-19, 64]. Most of the available literature examines this relationship by analyzing the microbiota its relationship to extraction of nutrients from the host diet. This approach has led to the concept that the microbiota is an environmental factor; contributing to increased energy harvest from dietary contents [14, 16]. This is further documented by studies assessing the effects of transplanting the intestinal microbiota of obese mice into lean recipients and assessing dietary harvest. Turnbaugh et al. showed that the transplantation of obese microbiota into lean recipients resulted in an increase in Firmicutes while transplantation of lean

microbiota had no such effect [14]. These observations are in agreement with those of Backhed et al. who determined that conventionally raised mice (harboring natural microbiota) had an increased body mass when compared to germ free counterparts, despite consuming less chow [41]. Backhed et al. illustrated a similar effect of transplantation of the microbiota from conventionally raised mice to germ free counterparts. The germ free mice were colonized for 14 days with "unfractionated microbiota" leading to an increase in body fat content of fifty-seven percent, even with a decrease in chow consumption [41].

The fecal content in the distal colon is not uniform [38], and exists as both liquid and particle phases. Recent evidence has shown that each phase of the fecal matter harbors distinct communities [38]. Particle phase fecal matter showed increased Firmicutes with lower Bacteroidetes while liquid phase fecal matter was associated with a higher percentage of Bacteroidetes and lower Firmicutes with respect to control [38]. This was suggested to be the result of a resistancedfx. of certain carbohydrates in the diet, requiring action from multiple bacterial species for breakdown and utilization.

Interestingly, studies performed on subjects consuming vegetarian diets showed increased Bacteroides and distinct phylotypes (species) associated with such diet versus an omnivorous diet including meat [10, 11]. Hehemann et al. 2010 explained the acquisition of a porphyrinase gene common to marine Bacteroidetes, observed in the intestinal Bacteroidetes of Japanese subjects. The authors proposed that this gene was acquired by the intestinal microbiota as a result of repeated consumption of nori seaweed and the need to metabolize the plant wrap common with this type of sushi preparation. The overall suggestion of this evidence is that the microbiota can be altered after prolonged exposure to certain diets [11, 17, 64, 65].

Detection of Bacteria in the Intestinal Tract

Classical techniques for investigating intestinal microbiota in fecal content involve the culturing and counting of isolated organisms. As a result, the initial assessment of the intestinal microbiota suggested dominance by species such as *Bacteroides*, however; these results are incomplete and lack accurate representation of fastidious and strict anaerobes [66, 67]. Tissue and mucosal biopsies, and sampling of the lumen through colonic lavage and fecal collection allow researchers to gain insight into the intestinal environment [68-72]. Advances in technology have allowed researchers to bypass the need for cultivating bacterial cells, which is estimated to only encompass thirty percent of the total microbiota population, and have created more robust techniques [73].

One avenue that allows investigators to accurately identify the intestinal microbiota is through looking at ribosomes. All living organisms contain ribosomes, as they are essential for protein synthesis, and thus organisms can be related based on sequence similarity of ribosomal subunits, specifically the small 16S subunit of prokaryotic ribosomes. This 1.5kb region of the bacterial genome is very useful when it comes to determining diversity of the intestinal microbiota as it contains both conserved regions as well as species-specific variable regions. In addition to this, two of the conserved regions on the 16S ribosomal DNA happen to lie near the ends of the sequence, therefore researchers are accurately able to amplify the entire 16S sequence and then do more specific interrogation into specific variable regions for bacteria of interest [74-78].

New techniques including denaturant-gradient gel electrophoresis, 16S high-throughput sequencing, are able to determine presence of dominant species as well as detect novel species, yet lack the quantitative ability of quantitative-PCR, dot-blot hybridization, and fluorescent *in-situ* hybridization. The limitations of these latter technologies include the difficulty of applying them to large populations [74]. The outcome of studies using such advanced technology have provided a modern image of the intestinal microbiota, showing a predicted greater presence of obligate anaerobes, permitting metabolomic studies and investigations into the relationship of these newly detected organisms to human health and disease [11, 17, 64, 69, 70, 75, 77, 78].

The development of microarray technology has been extremely beneficial in the pursuit of establishing a complete picture of the intestinal microbiota. Microarrays are capable of detecting thousands of species and, encompassing large populations; they have also been shown to be accurate and reliable [74, 75, 79]. Currently, there are a few different microarray platforms, one of the more well known is the Human Intestinal Tract chip (HITchip), which was recently applied to young and elderly adults and shown to be more reliable than DGGE fingerprinting, further solidifying the microarray as a superior technique [79]. Recently Paliy et al. designed a custom Microbiota Array to be used in the analysis of human fecal content. The Microbiota Array was shown to detect accurately fifteen bacterial species tested; both when hybridized individually to the array and in mixed communities. The authors also showed the ability of the Microbiota Array to distinguish and accurately detect bacterial species when human DNA was mixed with mixed bacterial samples. They also demonstrate the sensitivity of the Microbiota Array by running 16S specific PCR up to thirty-cycles which permitted the detection limit of the Microbiota Array to drop from 4ng of bacterial genomic DNA to merely 10pg [75]. The Paliy lab also recently published a report addressing the optimization of the microarray process. One aspect of this report was the adjustments for both potential cross-hybridization and 16S gene copy

numbers possessed by different bacteria [80]. The outcome of such adjustments resulted in decreased abundance of highly abundant bacteria and increased abundance of lower abundant bacteria. This suggests that previous reports may have inadvertently over- or under represented certain bacterial groups. Overall, the development and validation of microarrays has enabled researchers to get a significantly better quantitative understanding of microbiota populations.

Thesis Overview

The following work describes the use of a Microbiota Array to accurately assess and determine the structure of the intestinal microbiota. Fecal samples were obtained from healthy adult volunteers. Bacterial genomic DNA was extracted and subjected to 16S PCR, followed by DNA fragmentation, end labeling with biotin, and hybridization onto the microarray. Microarrays were washed and stained according to Affymetrix protocol and scanned to measure fluorescence intensity, determining presences/absence calls and signal values. Presence and absence calls, along with probeset signal values were imported into a MICROSOFT EXCEL template created by Dr. Oleg Paliy; this template aided in determination of intestinal microbiota numbers. The results of this work provide quantitative information into the makeup of the human distal-gut microbiota, are a useful complement to other microarray studies, and can be combined with metabolomic, co-morbidity, proteomics, and other such investigations to continue and understand the significance of the intestinal microbiota in human health and disease.

2. Materials and Methods

Sample Collection

Fecal samples were obtained from 10 healthy adult volunteers with ages ranging from 22-61 yrs. Volunteers were asked to defecate into sterile sample collection containers; samples were frozen immediately after collection at -80°C. Aliquots of frozen stool material were taken for processing and microarray interrogation (samples were kept on dry ice to increase thawing time and keep DNA integrity intact). Table 2.1 indicates age and gender of each volunteer, except when information was not given (N/G).

Table 2.1Volunteer Age and Gender Assignment

Sample	Age	Gender
aHLT01	25	F
aHLT02	22	F
aHLT03	35	M
aHLT04	61	M
aHLT05	36	M
aHLT06	44	F
aHLT07	23	M
aHLT08	N/G	N/G
aHLT09	N/G	N/G
aHLT10	28	F

DNA Extraction

Total genomic DNA (gDNA) was extracted with the ZR Fecal DNA Kit (Zymo Research Corp), using the supplied directions:

a) 150 mg of fecal matter was added to a ZR Bashing Bead Lysis Tube followed by 750 µL of lysis buffer.

b) ZR Bashing Bead Lysis Tube was placed in a Disruptor Genie® and processed for 5 minutes.

c) Tube was centrifuged at approx. 10,000 g for 1 minute, followed by a transfer of 400 µL of supernatant to a Zymo-Spin IV Spin Filter (inside collection tube) and centrifuged at 7,000 g for 1 minute.

d) 1,200 µL of Fecal DNA Binding Buffer was added to the filtrate and 800 µL of this mixture was transferred to a Zymo-Spin IIC Column (inside collection tube) and centrifuged at 10,000 g for 1 minute.

e) Flow-through was discarded and the previous step repeated; due to a 800 µL limit in the Zymo-Spin IIC Column this process must be repeated multiple times.

f) 200 µL of Fecal DNA Pre-Wash Buffer was added to the Zymo-Spin IIC Column and centrifuged at 10,000 g for 1 minute.

g) 500 µL of Fecal DNA Wash Buffer was added to the Zymo-Spin IIC Column and centrifuged again at 10,000 g for 1 minute.

h) Total genomic DNA was eluted with 100 µL of water, after transferring the Zymo-Spin IIC Column to a clean micro-tube and adding water directly to column matrix; the micro-tube was then centrifuged at 10,000 g for 30 seconds.

i) Eluted DNA was passed through the Zymo-Spin IV-HRC Spin Filter, placed in a clean micro-tube; and spun down at 8,000 g for 1 minute.

Once gDNA was eluted, yield and quality (determined by A_{260}/A_{280} ratio) were assessed by Nanodrop 1000. Extracted gDNA was used as starting material for all replicates of downstream processing.

Bacterial 16S rDNA amplification and purification

Total gDNA, which contains eukaryotic, prokaryotic, and viral DNA was subjected to selective amplification of the 16S rRNA gene through 16S rDNA-specific PCR amplification. Amp_27F-(AGRGTTYGATYMTGGCTCAG) and Amp_1492R-(GYTACCTTGTTACGACTT) primers, designed to conserved regions of the 16S ribosomal DNA sequence [75], were used to specifically amplify the bacterial 16S rRNA gene. The PCR reaction was carried out in a 50 μL volume with 1 μL (10 pmoles) of each primer, 25 μL of Taq 2X Master Mix (New England Biolabs) which incorporates 0.2mM of each dNTP (deoxy-nucleotide triphosphate) and 25U/ml of the Taq DNA polymerase, 250 ng of total gDNA starting template (percentage 16S-0.0025% of genome), and water. Each sample was subjected to 4 PCR reactions, carried out in the MJ Thermocycler. Each reaction consisted of 25 cycles of amplification; each cycle was run at 95°C for 30 s - 55°C for 30 s - 72°C for 90 s. To decrease PCR bias, all 4 reactions were pooled together.

Table 2.2 PCR Reaction Mix

PCR Reaction	Amount
H_2O	up to 50 μL
27F_V4	1 μL
1492_R	1 μL
gDNA	250 ng
Taq 2X	25 μL
Total	50 μL

The amplified DNA product was subjected to purification using the Qiaquick PCR Purification kit (Qiagen) based on the following protocol:

a) Five volumes of PB Buffer was added to one volume of the PCR reaction, transferred to a PCR purification column and centrifuged at 17,900 g for 1 minute.

b) The flow-through was discarded and 750 μL of Buffer PE added to the column, placed back in the collection tube, and centrifuged for 1 minute at 17,900 g.

c) Flow-through was discarded, and the column spun down again to ensure removal of buffer.

d) The spin column was transferred to a new collection tube and 30 μL of RNase free water, heated to 50C, was added to the center of the column and let to sit for 2 minutes (the supplied protocol instructed using room temperature water and sitting for 1 minute, this led to poor yield, therefore the protocol was changed with the addition of warmer water and a longer sit time).

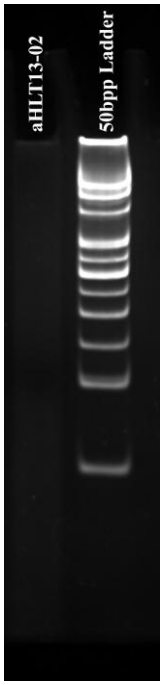
e) The column was centrifuged at 13,000 g for 1 minute to elute PCR amplified DNA.

The 16S amplification and purification was verified on a 1% agarose gel run at 75 V for 1 hour. 1 μL of amplified DNA was combined with 2 μL gel loading dye (6X) and 9 μL water, then loaded onto the gel. Positive verification of correct 16S amplification was seen with a single band at approximately 1.5kb when compared with a 2-log Ladder (NEB).

16S Fragmentation

In order to interrogate the 16S rDNA with the microarray, the amplified 16S rDNA was fragmented using Dnase I (NEB) and verified on a 10% poly-acrylamide gel. The fragmentation reaction consisted of 1800 ng of 16S rDNA amplified pool, 4 μL Dnase buffer, Dnase I (0.04 U/μg), and water to a final volume of 40 μL. The reaction was run in a MJ Thermocycler at 37°C for 10 min followed by 98°C for 10 min to deactivate the enzyme. 300 ng of fragmented 16S rDNA was removed from each reaction mix and combined with gel loading dye and water to a final volume of 6 μL, which was then added to the gel and run at 75 V for 90 minutes. A 50 bp dsDNA ladder was used to determine whether DNA fragments were within the desire 100-300 bp size.

Figure 2.1 Fragmentation PAGE Gel



Microbiota Array

The Microbiota Array is a custom designed phylogenetic array, based off the Affymetrix GeneChip design. Microbiota Array contains probes to 775 phylo-species inhabiting the distal gut. Probes contain 25 nucleotides and are grouped in probesets, ranging from 5-11 probes per set (84% of probesets contain 11 probes). Probe sets incorporate Perfect Match and Mismatch probes.

DNA Labeling and Hybridization

Fragmented DNA was end-labeled with biotin and then loaded onto the Microbiota Array. The labeling reaction combined Terminal Deoxynucleotidyl Transferase (TdT), along with buffer; Gene Label Reagent, fragmented DNA, and $CoCl_2$. This reaction mix was incubated

in the MJ-Thermocycler at 37°C for 1 hour and then stopped by the addition of 2 μL of 0.5M EDTA.

Table 2.3 Labeling Reaction Mix

Labeling Reaction	Amount
10x TDT Buffer	5 μL
CoCl	5 μL
TDT (Transferase)	2 μL
Fragmented DNA	1500 ng
Gene Label Reagent	2 μL
H2O	up to 50 μL
Total Volume	50 μL

After labeling was completed, the DNA cocktail was ready to be added to the microarray for hybridization. Affymetrix protocol combines fragmented/labeled DNA, hybridization buffer (2X), Control oligo B2, DMSO, Herring sperm DNA (10mg/ μL), and BSA (50mg/ μL).

Table 2.4 Hybridization Reaction Mix

Hybridization Mix	Amount
Frag.;Labeled cDNA	1500 ng
2X Hybridisation Buffer	up to 65 μL
Control Oligo B2	2.2 μL
100% DMSO	10.2 μL
10 mg/ml Herring Sperm DNA	1.3 μL
50 mg/ml BSA	1.3 μL
Total Volume	130 μL

Once the hybridization cocktail components were combined, the mixture was added to the microarray and then placed into the hybridization oven for 16 hours at 45°C and 60 rpms.

Washing, Staining, and Scanning of Microarray

Upon completion of hybridization the microarray was removed from the oven, then washed and stained in an Affymetrix Fluidics 450 station. Stains, along with Non-Stringent

Wash Buffer A and Stringent Wash Buffer B were prepared in accordance with Affymetrix recipes. The microarray was removed from the hybridization oven after 16 hours; the hybridization mix was removed and frozen at -20°C, replaced with 160 µL of Non-Stringent Wash Buffer A, and loaded into the Fluidics 450 along with stains. Once loaded, the array was washed and stained using the "Midi-euk_2v3_450" protocol. Upon completion of staining, the microarray was transferred to the Affymetrix GeneChip Scanner where it was scanned, fluorescence measured and recorded.

Table 2.5 Stains for Microarray

Streptavidin Vial 1	**Amount**
2x Stain Buffer	300 µL
50 mg/ml BSA	24 µL
1 mg/ml Streptavidin	6 µL
H2O	270 µL
Total Volume	600 µL

Antibody Soln. Vial 2	**Amount**
2x MES Stain Buffer	300 µL
50 mg/ml BSA	24 µL
10 mg/ml Goat IgG	6 µL
0.5 mg/ml Anti-strep	6 µL
H2O	264 µL
Total Volume	600 µL

SAPE Soln. Vial 3	**Amount**
2x MES Stain Buffer	300 µL
50 mg/ml BSA	24 µL
1 mg/ml SAPE	6 µL
H2O	270 µL
Total Volume	600 µL

Normalization and Analysis of Microarray Data

Microarray data were normalized using the CARMAweb server (https://carmaweb.genome.tugraz.at/carma) [81]. Normalization of the data was carried out using

Mas5 algorithm for background correction, VSN for normalization, Mas5 for PM correction, and Median Polish for expression.

After the normalization was completed, these values, along with presence and absence calls (calculated by GCOS software) were inserted into a MICROSOFT EXCEL template provided by Dr. Paliy; this template allowed for calculation of abundances for bacterial groups at multiple phylogenetic levels. Microarray replicates showed good concordance with one another, as viewed through correlation of probe signal values between replicates (0.89-0.94, ave. 0.92).

To assess core microbiome, presence and absence calls were converted to binary (1/0) form and loaded into Matlab. It was here that data was subjected to rarefaction, which determines how well populations are sampled and can be used to determine how many species are shared between samples. This is accomplished by doing comparisons of all possible sample combinations and assessing what is shared among all samples over each comparison.

Phylogenetic Tree Construction

Phylogenetic tree was created using BOSQUE [82]. Sequences of the 775 phylo-species on Microbiota Array were imported into BOSQUE and aligned using MUSCLE algorithm. Following alignment, Phylip method was applied for tree construction.

3. **RESULTS**

Intestinal Microbiota at the Phylum Level

This study was carried out to determine the intestinal composition of the microbiota of healthy adult individuals through the use of a custom designed Microbiota Array. Fecal samples were collected from 10 healthy adult volunteers. Adults ranged in age from 22 to 61 years of age, had not taken antibiotics within 3 months prior to donating samples, and were all in general good health. Samples were frozen upon collection to maintain DNA integrity.

Bacterial genomic DNA was extracted from fecal matter and subsequently processed for interrogation by a previously designed Microbiota Array. After washing and staining, the array was scanned, the output signals were normalized in CARMAweb and then imported into an EXCEL template to aid in analysis of microarray data.

Adjusted microarray results showed that the intestinal microbiota is dominated by the phyla Firmicutes, Actinobacteria, Bacteroidetes, and Proteobacteria at 80%, 7%, 9% and 3% abundance, respectively (Table 3.1). Seven out of ten phyla were present to some extent (Table 3.2). As average relative abundance decreased it was observed that variation among samples increased. This observed pattern in the variation of phyla among samples suggests that bacteria which are essential for proper gut health are present at similar levels in all people, whereas those bacteria which vary significantly can be thought of as disposable. Overall, certain phyla were more different in abundance across samples. This varying abundance pattern suggests that those phyla with higher variation are not as important with respect to intestinal homeostasis. An example of this, seen in Table 3.2, is Actinobacteria with abundances ranging from 0.9% to 15.9%.

TABLE 3.1 Overview of Phylum Level Abundances

Phylum	Average Abundance	St. Deviation	Coefficient of Variation
Proteobacteria	2.7%	2.25%	83.8%
Firmicutes	79.8%	7.38%	9.2%
Actinobacteria	7.2%	5.41%	75.2%
Spirochaetes	0.1%	0.08%	108.3%
Bacteroidetes	8.8%	2.97%	33.7%
Verrucomicrobia	1.0%	0.83%	84.6%
Lentisphaerae	0.4%	0.93%	251.2%

Table 1 shows the average relative abundance of intestinal bacterial phyla among ten healthy adult volunteers, along with standard deviation. There is a pattern of increasing standard deviation with respect to the abundance of each phylum. Coefficient of variation was calculated to determine the extent of variation among samples.

TABLE 3.2 Abundances of Phyla among Individual Samples

Phylum	aHLT01	aHLT02	aHLT03	aHLT04	aHLT05	aHLT06	aHLT07	aHLT08	aHLT09	aHLT10
Proteobacteria	1.9%	4.7%	0.4%	0.5%	5.2%	3.6%	0.4%	6.7%	1.2%	2.2%
Firmicutes	82.9%	65.6%	90.4%	86.8%	78.5%	72.3%	84.8%	75.7%	77.7%	83.3%
Actinobacteria	4.3%	14.9%	1.1%	0.9%	1.9%	15.9%	9.8%	5.7%	9.5%	8.1%
Spirochaetes	0.3%	0.1%	0.0%	0.0%	0.1%	0.0%	0.1%	0.1%	0.1%	0.0%
Bacteroidetes	7.8%	12.7%	7.3%	10.5%	13.4%	5.8%	4.7%	8.8%	11.0%	6.2%
Verrucomicrobia	2.8%	1.5%	0.7%	1.2%	1.0%	1.6%	0.3%	0.0%	0.5%	0.2%
Lentisphaerae	0.0%	0.6%	0.0%	0.0%	0.0%	0.1%	0.0%	3.0%	0.0%	0.0%

Table 2 illustrates relative abundances of seven phyla of human intestinal microbiota. Firmicutes was most abundant, followed by Bacteroidetes. Certain phyla showed more variation among samples than other, which suggests varied levels of importance among observed phyla. One case is Actinobacteria which had a large range of abundance (min-0.9%, max-15.9%).

Class Level Microbiota Structure

Class level results illustrated that Clostridia, a dominant member of the Firmicutes phylum, was most abundant at 74.0% with Bacteroidetes following at 8.8% (Table 3.3). Actinobacteria and Bacilli were the next most abundant classes at 7.2% and 3.7% respectively; however the ratio of standard deviation-to-abundance suggests that the variability among each sample is high and thus might point out potential outliers. Figure 3.1 shows microbiota encompassing 10 adult samples. Species that were present in at least 1 of the samples were included in the tree. The tree was created using the Bosque software [82], employing the Phylip (F84) distance method for construction. While most of the species contained in the tree belong to Clostridia, and the distribution of groups within the tree matches the numerical abundance data; there were some species that grouped closely together, despite being from different classes according to our classification (indicated by asterisk). The RDP IDs were taken and put into RDPs browser function, and came back as "unclassified_Firmicutes" indicating that recent updates to the RDP database have altered the classification of these entries.

Table 3.3 Average Abundance of Classes Comprising the Intestinal Microbiota

Class	Corresponding Phylum	Average Abundance	St. Deviation	Coefficient of Variation
Alphaproteobacteria	**Proteobacteria**	0.2%	0.8%	316.2%
Betaproteobacteria	**Proteobacteria**	1.4%	1.4%	102.1%
Gammaproteobacteria	**Proteobacteria**	0.4%	0.4%	108.0%
Deltaproteobacteria	**Proteobacteria**	0.7%	0.7%	95.0%
Epsilonproteobacteria	**Proteobacteria**	<0.1%	<0.1%	255.4%
Clostridia	**Firmicutes**	74.0%	5.6%	7.6%
Mollicutes	**Firmicutes**	2.0%	2.4%	118.8%
Bacilli	**Firmicutes**	3.7%	2.3%	61.4%
Actinobacteria	**Actinobacteria**	7.2%	5.4%	75.2%
Spirochaetes	**Spirochaetes**	0.1%	0.1%	108.3%
Bacteroidetes	**Bacteroidetes**	8.8%	3.0%	33.7%
Verrucomicrobiae	**Verrucomicrobia**	1.0%	0.8%	84.6%
Lentisphaerae	**Lentisphaerae**	0.4%	0.9%	251.2%

Table 3 depicts class level relative average abundance, standard deviation, and variation. As seen with the phylum level, there exists an increase in standard deviation related to decreased abundance. Clostridia was dominant at more than eight-times the abundance of any other class. Variation again, increases as abundance decreases, this highlights classes which are present in only a few samples overall; leading to such higher variation.

Figure 3.1 Class Makeup

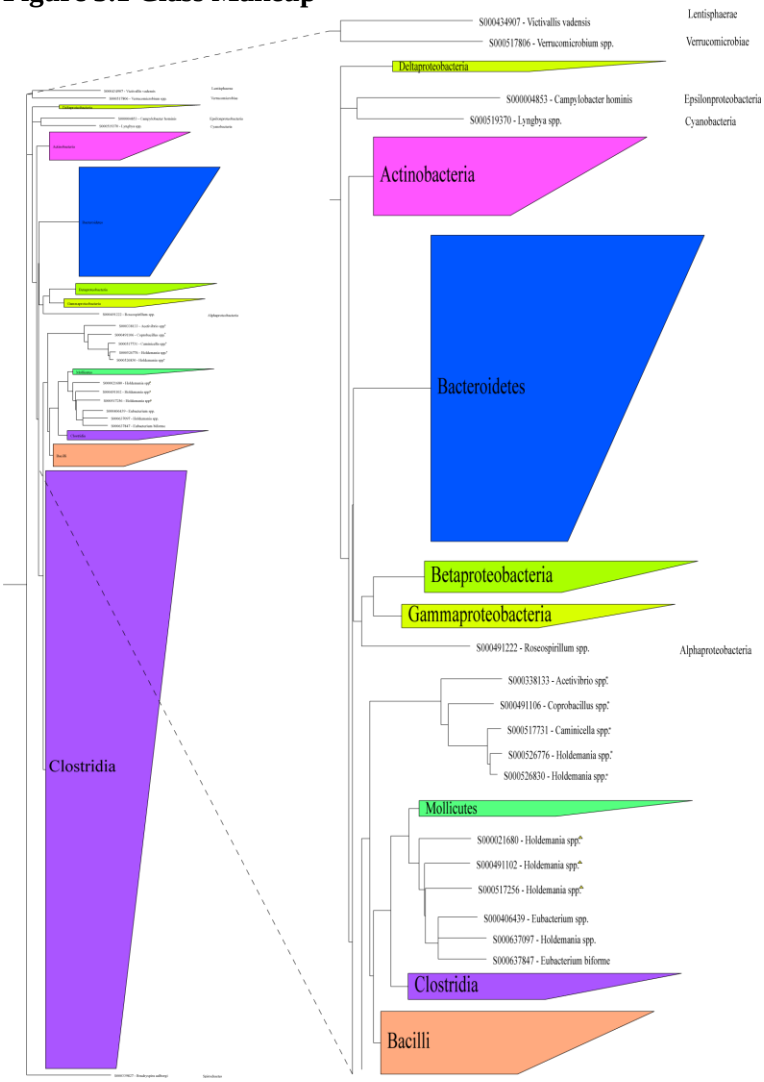


Figure 3.1 is a Cladogram illustrating class-level makeup of intestinal microbiota. Polygon size is relative to the number of phylo-species present, exceptions to this are classes where the number of present phylo-species was less than the smallest default polygon size. Species that are not part of a polygon are either the only representative of their class, or have recently been reclassified as "unclassified_Firmicutes" (asterisk) or Eubacterium (yellow triangle).

Beyond the four most abundant classes, the standard deviation was seen to be almost as high as the average observed abundance, if not higher. This pattern in the standard deviation, coupled with the variation, suggested that specific samples harbored large abundances of bacteria

compared to the same group in other samples. Upon looking at each individual sample, specific differences accounted for the variation (Table 3.4, Figure 3.2), however Alphaproteobacteria and Epsilonproteobacteria were represented in only one sample each (aHLT08 and aHLT09 respectively). This unique presence of these two classes is not enough to treat the samples as uniquely different but may potentially suggest a sort of microbial "fingerprint" when combined with other differences in the observed population structure. For instance, the presence of Alphaproteobacteria in aHLT08 at 2.4% compared to an absence in all other samples could be indicative of diet or even age (note: age for this individual not given).

Table 3.4 Abundances of Classes of the Intestinal Microbiota across Healthy Adults

Class	aHLT01	aHLT02	aHLT03	aHLT04	aHLT05	aHLT06	aHLT07	aHLT08	aHLT09	aHLT10
Alphaproteobacteria	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	2.4%	0.0%	0.0%
Betaproteobacteria	0.1%	2.7%	0.0%	0.5%	2.9%	2.7%	0.3%	3.5%	0.8%	0.2%
Gammaproteobacteria	0.0%	1.0%	0.3%	0.0%	1.0%	0.1%	0.1%	0.1%	0.4%	0.4%
Deltaproteobacteria	1.8%	0.9%	0.1%	0.0%	1.3%	0.8%	0.0%	0.8%	0.0%	1.6%
Epsilonproteobacteria	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.0%
Clostridia	76.8%	62.7%	80.0%	79.2%	68.3%	70.0%	79.7%	72.7%	75.1%	75.8%
Mollicutes	1.7%	0.8%	6.0%	6.9%	2.3%	0.6%	0.4%	0.6%	0.4%	0.7%
Bacilli	4.4%	2.1%	4.4%	0.8%	7.9%	1.8%	4.7%	2.4%	2.2%	6.8%
Actinobacteria	4.3%	14.9%	1.1%	0.9%	1.9%	15.9%	9.8%	5.7%	9.5%	8.1%
Spirochaetes	0.3%	0.1%	0.0%	0.0%	0.1%	0.0%	0.1%	0.1%	0.1%	0.0%
Bacteroidetes	7.8%	12.7%	7.3%	10.5%	13.4%	5.8%	4.7%	8.8%	11.0%	6.2%
Verrucomicrobiae	2.8%	1.5%	0.7%	1.2%	1.0%	1.6%	0.3%	<0.1%	0.5%	0.2%
Lentisphaerae	0.0%	0.6%	0.0%	0.0%	0.0%	0.1%	0.0%	3.0%	0.0%	0.0%

Table 3.4 represents individual abundances of Classes interrogated by the Microbiota Array. Of particular note was *Actinobacteria* with a range of 15% across all samples, (lowest-0.9%, highest-15.9%), interesting given the average for this class was approximately half the largest value.

Figure 3.2 Distribution of Classes of Intestinal Microbiota in Healthy Adults

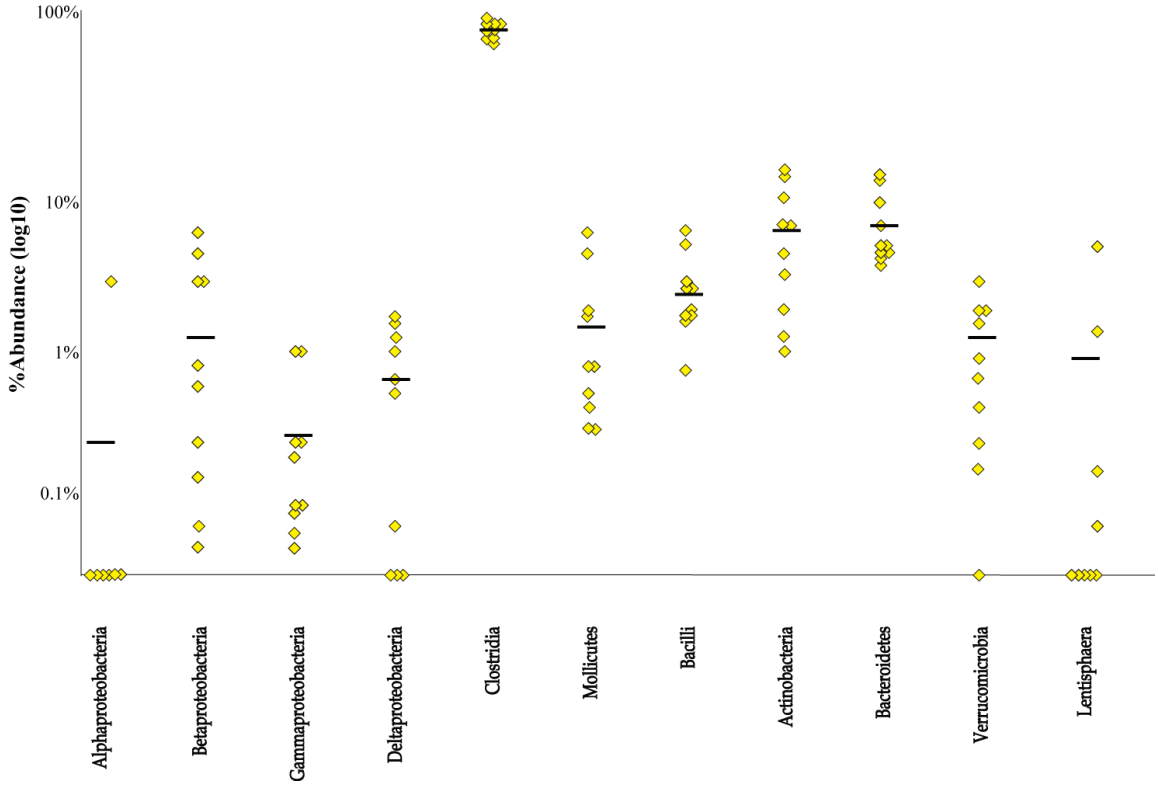


Figure 3.2 illustrates distribution of Classes across adult samples. Values were plotted on a log base 10 scale and represent Class abundances over 10 healthy adults. Black bars indicate average relative abundance for each class as shown in Table 3.

Overall, class level microbial structure is dominated by three classes, Clostridia, Bacteroidetes, and Actinobacteria, which account for 90% of the intestinal microbiota. Abundance is based upon presence of phylo-species; therefore, we assessed phylo-species detection for each class, with the purpose of determining any relationship between abundance of classes and those phylo-species that were detected for each class. Table 3.5 shows the average counts of detected phylo-species in decreasing order of present phylo-species.

Table 3.5 Phylo-species Numbers of Classes of the Intestinal Microbiota

Class	Corresponding Phylum	Average # Probesets	St. Deviation	Coefficient of Variation
Alphaproteobacteria	**Proteobacteria**	0.1	0.3	316.2%
Betaproteobacteria	**Proteobacteria**	2.3	1.3	58.2%
Gammaproteobacteria	**Proteobacteria**	2.9	1.0	34.3%
Deltaproteobacteria	**Proteobacteria**	1.0	0.9	94.3%
Epsilonproteobacteria	**Proteobacteria**	0.2	0.4	210.8%
Clostridia	**Firmicutes**	266.3	20.0	7.5%
Mollicutes	**Firmicutes**	4.8	1.7	35.1%
Bacilli	**Firmicutes**	9.7	2.5	25.3%
Actinobacteria	**Actinobacteria**	10.9	4.5	41.4%
Spirochaetes	**Spirochaetes**	0.7	0.5	69.0%
Bacteroidetes	**Bacteroidetes**	29.6	10.6	35.8%
Verrucomicrobiae	**Verrucomicrobia**	0.9	0.3	35.1%
Lentisphaerae	**Lentisphaerae**	0.4	0.5	129.1%

Table 3.5 indicates average number of probesets detected. This is relative to the number of phylo-species for each class.

The difference in pattern of detected phylo-species within each class compared to the decreasing abundance as seen in Table 3.3 suggests a potential for alternate assessment of microbiota as it relates to diet. In contrast with class level abundance the decreasing order of the number of detected phylo-species differed, mainly among proteobacteria members (Table 3.2 versus Table 3.3). The difference between detected phylo-species and abundance level among each class may suggest differences in the importance of certain phylogenetic groups. In other words, different people may harbor similar groups of bacteria, but subtle life differences could lead to different sizes of each bacterial group. This may also be explained by determining the present-phylo species per sample, as is shown in Figure 3.3 Variation among detected probesets was, at most, in agreement with variation in abundance values, in many cases the variation between probesets was much less (Table 3.2 versus Table 3.3).

Figure 3.3 Numbers of Phylo-species within Classes

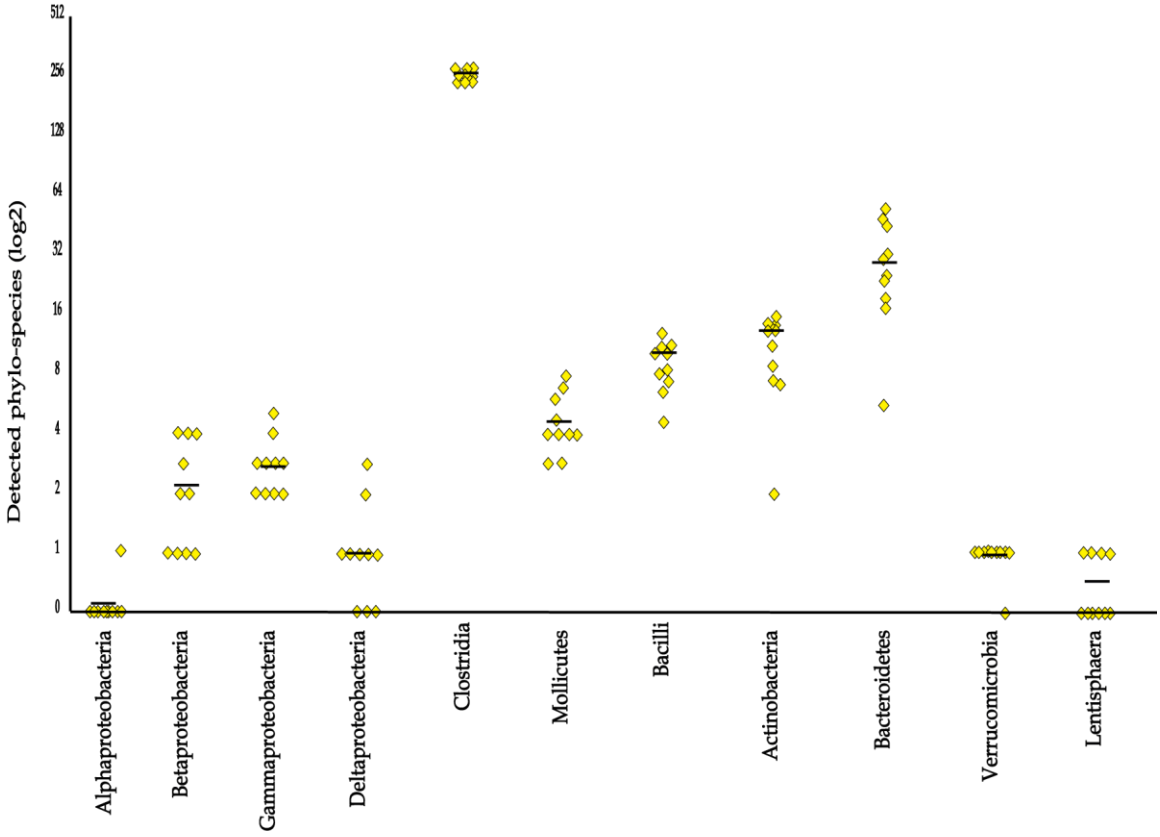


Figure 3.3 shows detected probesets for each Class across individuals. Values are plotted on a log base 2 scale.

Families of the Intestinal Microbiota

While previous studies have dealt with class level microbiota structure, this does not give sufficient evidence of specificity in the intestinal microbial community. The Microbiota Array allowed us to take a look into family groups. Results showed that out of forty-seven possible familial groups interrogated, twenty-seven were detected in at least one sample, with the remaining twenty not being detected at all (Table 3.6). Of those families that were found present, Lachnospiraceae, Clostridiaceae, and Acidaminococcaceae made up roughly 69% of total signal (36%, 24%, and 8% respectively); of note since these three families all belong to the class Clostridia. Durban et al. ([83]) report similar results with Ruminococcaceae as most abundant in fecal samples, followed by Lachnospiraceae and Bacteroidaceae.

Table 3.6 Average Relative Abundances of Family Groups

Family	Corresponding Class	Average Abundance	Standard Deviation	Coefficient of Variation
Rhodobacteraceae	Alphaproteobacteria	-	-	-
Sphingomonadaceae	Alphaproteobacteria	-	-	-
Phyllobacteriaceae	Alphaproteobacteria	-	-	-
Bradyrhizobiaceae	Alphaproteobacteria	-	-	-
Methylobacteriaceae	Alphaproteobacteria	-	-	-
Methylocystaceae	Alphaproteobacteria	-	-	-
Rhodobiaceae	Alphaproteobacteria	0.2%	0.8%	315.4%
Burkholderiaceae	Betaproteobacteria	0.2%	0.2%	109.2%
Oxalobacteraceae	Betaproteobacteria	-	0.1%	301.0%
Alcaligenaceae	Betaproteobacteria	1.1%	1.4%	125.6%
Neisseriaceae	Betaproteobacteria	-	-	-
Xanthomonadaceae	Gammaproteobacteria	-	-	-
Moraxellaceae	Gammaproteobacteria	-	-	-
Succinivibrionaceae	Gammaproteobacteria	-	-	-
Enterobacteriaceae	Gammaproteobacteria	0.1%	0.2%	163.9%
Pasteurellaceae	Gammaproteobacteria	0.2%	0.3%	142.1%
Desulfovibrionaceae	Deltaproteobacteria	0.7%	0.7%	95.3%
Campylobacteraceae	Epsilonproteobacteria	-	0.0%	220.2%
Helicobacteraceae	Epsilonproteobacteria	-	-	-
Clostridiaceae	Clostridia	24.2%	2.3%	9.4%
Lachnospiraceae	Clostridia	36.1%	4.3%	11.8%
Peptostreptococcaceae	Clostridia	-	-	-
Eubacteriaceae	Clostridia	4.9%	1.0%	20.2%
Peptococcaceae	Clostridia	0.6%	0.8%	130.8%
Acidaminococcaceae	Clostridia	8.2%	1.9%	22.9%
Thermoanaerobacteriaceae	Clostridia	0.1%	0.1%	136.7%
Erysipelotrichaceae	Mollicutes	2.0%	2.4%	118.8%
Staphylococcaceae	Bacilli	-	-	-
Turicibacteraceae	Bacilli	0.2%	0.2%	113.7%
Lactobacillaceae	Bacilli	0.1%	0.2%	268.6%
Enterococcaceae	Bacilli	0.1%	0.2%	213.4%
Streptococcaceae	Bacilli	2.8%	1.8%	66.4%
Firmicutes - IS9	Bacilli	0.6%	0.7%	111.2%
Coriobacteriaceae	Actinobacteria	1.6%	0.7%	43.9%
Actinomycetaceae	Actinobacteria	-	0.0%	91.7%

Micrococcaceae	**Actinobacteria**	-	-	-
Dermabacteraceae	**Actinobacteria**	-	-	-
Corynebacteriaceae	**Actinobacteria**	-	-	-
Bifidobacteriaceae	**Actinobacteria**	5.5%	4.8%	87.0%
Serpulinaceae	**Spirochaetes**	0.1%	0.1%	106.7%
Leptospiraceae	**Spirochaetes**	-	-	-
Bacteroidaceae	**Bacteroidetes**	5.9%	1.9%	32.0%
Rikenellaceae	**Bacteroidetes**	1.5%	0.9%	59.7%
Porphyromonadaceae	**Bacteroidetes**	0.4%	0.7%	168.0%
Prevotellaceae	**Bacteroidetes**	1.1%	2.0%	184.0%
Verrucomicrobiaceae	**Verrucomicrobiae**	1.0%	0.8%	84.1%
Victivallaceae	**Lentisphaerae**	0.4%	0.9%	250.6%

Table 6 depicts families interrogated by the Microbiota Array. Groups not detected are represented with "-".

The difference between Durban et al.'s findings and those here could be due to differing techniques (they applied cloning and sequencing), classification version (results here are based off of RDP version 4, although RDP database is updated frequently) or study design, as Durban's group did not appear to exclude subjects recently on antibiotics or who had a BMI that would indicate the subject was overweight. Of those families present it is interesting to note that multiple members of the class *Clostridia* are present at higher abundances, while only one member of *Bacteroidetes* and *Actinobacteria* are present. This pattern of abundance suggests that while some larger taxa may represent greater proportions, they may not be an essential factor in maintaining homeostasis in the intestinal tract. This shows that analysis of lower taxonomical levels will reveal more detail and allow for a better understanding of the intestinal microbiota.

Microbiota Constituents - Genus Contribution

The Microbiota Array interrogates for 115 bacterial genera, 22 of which had abundances equal or greater to 1%. The major contributors were *Ruminococcus* at 21%, followed by *Faecalibacterium* at 9% (Table 3.7). This large difference between *Ruminococcus* and the other

present genera is of particular note as it gives evidence of this particular genus' role in metabolism, as a primary degrader.

Table 3.7 Average Relative Abundance of Intestinal Genera

Genus	Corresponding Class	Average Abundance	St. Deviation	Coefficient of Variation
Sutterella	**Betaproteobacteria**	1.1%	1.4%	125.5%
Clostridium	**Clostridia**	1.9%	0.9%	47.0%
Acetivibrio	**Clostridia**	2.7%	1.2%	44.9%
Anaerotruncus	**Clostridia**	3.2%	1.5%	47.3%
Dorea	**Clostridia**	2.2%	0.8%	35.7%
Faecalibacterium	**Clostridia**	8.7%	2.3%	26.6%
Subdoligranulum	**Clostridia**	2.5%	0.8%	33.6%
Lachnospira	**Clostridia**	4.2%	1.2%	29.5%
Anaerostipes	**Clostridia**	2.1%	0.4%	18.6%
Coprococcus	**Clostridia**	1.8%	1.2%	68.0%
Roseburia	**Clostridia**	6.1%	1.5%	24.2%
Ruminococcus	**Clostridia**	21.1%	2.8%	13.3%
Eubacterium	**Clostridia**	4.2%	0.9%	22.6%
Papillibacter	**Clostridia**	6.8%	1.8%	26.5%
Holdemania	**Mollicutes**	1.8%	2.3%	130.8%
Streptococcus	**Bacilli**	2.3%	1.4%	62.2%
Collinsella	**Actinobacteria**	1.1%	0.7%	65.8%
Bifidobacterium	**Actinobacteria**	5.5%	4.8%	86.9%
Bacteroides	**Bacteroidetes**	5.5%	1.6%	29.7%
Alistipes	**Bacteroidetes**	1.4%	0.9%	62.5%
Prevotella	**Bacteroidetes**	1.1%	2.0%	184.1%
Verrucomicrobium	**Verrucomicrobiae**	1.0%	0.8%	84.6%

Table 3.7 shows genera present at 1.0% or more. Standard Deviation and Coefficient of Variation were calculated to illustrate commonality of each genus among samples, as well as highlight potential differences which may give weight to the presence of specific genera.

The order of presence is of note for another reason, being that while there has been work done with *Ruminococcus*, *Faecalibacterium*, *Bifidobacterium*, and *Roseburia* with respect to nutrient requirements, there still remains some gaps in the knowledge surrounding the intricacies of *Papillibacter*. While work has been performed on lower-level taxonomical groups, much of

these studies have been focused on higher groups, possibly due to the stability seen with higher taxonomies. The fact that *Papillibacter* is among the more abundant members in the intestinal environment, suggests that its role in aiding and coordinating metabolic function in the intestine warrants future study.

Figure 3.4 Variation of Intestinal Genera Among Healthy Adults

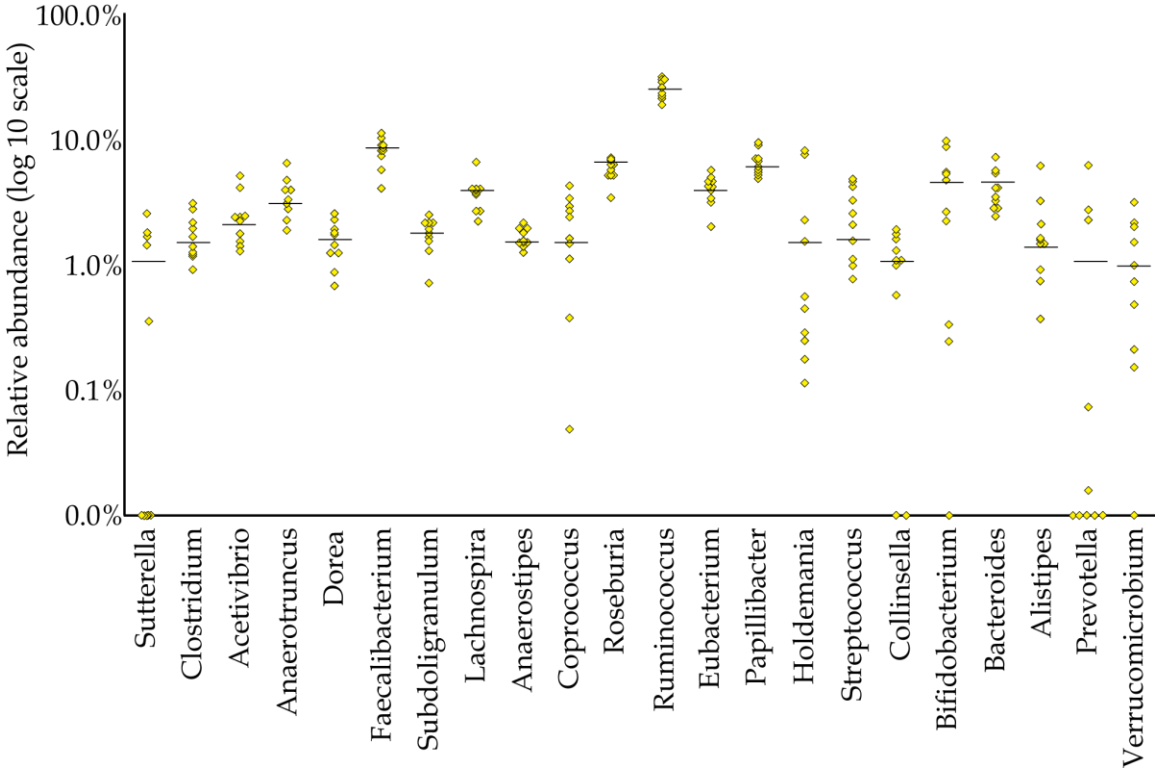


Figure 3.4 shows relative abundances and of genera for each sample along with average relative abundance of each genus.

Because microorganisms are essential for proper intestinal health, it becomes a curious issue to address individual uniqueness at these lower taxonomical levels. *Faecalibacterium, Papillibacter,* and *Roseburia* all had similar variation as calculated through coefficient of variation (26%, 26%, and 24% respectively), and abundances between samples show that there are few outliers (defined by samples with abundance values two or more standard deviations from the mean) (Table 3.8). This illustrates that these genera belong to a functionally important group with respect to intestinal homeostasis. Figure 3.4 shows how extensive the variation was for each

of the top 22 genera. From this image it can be inferred that some genera, carrying less variation across samples, possess necessary (and potentially unknown) functions. The variation observed in *Faecalibacterium* may be an indicator of patient diet as this organism is known to be associated with dietary fiber and particle-phase fecal content [38].

Table 3.8 Relative Abundance of Intestinal Genera Across Adult Samples

Genus	aHLT01	aHLT02	aHLT03	aHLT04	aHLT05	aHLT06	aHLT07	aHLT08	aHLT09	aHLT10
Sutterella	0.0%	2.7%	0.0%	0.3%	2.4%	2.4%	3.4%	0.0%	0.0%	0.0%
Clostridium	1.1%	3.0%	3.4%	1.3%	0.9%	2.3%	1.2%	1.9%	2.4%	1.2%
Acetivibrio	1.7%	1.6%	2.5%	5.2%	1.6%	2.6%	4.3%	2.6%	2.6%	1.9%
Anaerotruncus	4.8%	2.5%	2.7%	6.7%	3.5%	2.3%	3.6%	1.9%	2.8%	1.6%
Dorea	2.0%	2.4%	2.6%	0.6%	1.8%	1.6%	1.8%	2.7%	3.3%	3.0%
Faecalibacterium	9.0%	8.9%	6.2%	3.8%	8.1%	9.6%	9.0%	9.9%	12.0%	10.6%
Subdoligranulum	1.7%	0.7%	2.3%	3.4%	2.4%	2.8%	2.1%	3.1%	3.1%	3.1%
Lachnospira	4.3%	2.8%	7.4%	3.3%	4.2%	4.4%	4.0%	4.0%	3.3%	4.3%
Anaerostipes	1.6%	1.9%	2.8%	2.5%	1.9%	2.2%	1.8%	2.5%	1.9%	1.8%
Coprococcus	1.2%	0.0%	2.3%	0.4%	1.4%	2.4%	4.4%	1.5%	2.6%	2.0%
Roseburia	7.7%	3.6%	7.4%	4.7%	8.1%	5.0%	6.6%	6.7%	5.8%	5.1%
Ruminococcus	22.8%	18.8%	23.2%	22.1%	20.4%	15.8%	17.9%	22.1%	24.0%	24.2%
Eubacterium	4.1%	4.6%	4.6%	4.4%	3.6%	2.2%	3.4%	4.9%	5.6%	4.3%
Papillibacter	7.2%	5.3%	4.9%	10.2%	5.7%	9.3%	6.7%	7.3%	5.0%	6.1%
Holdemania	1.5%	0.5%	5.7%	6.4%	2.1%	0.3%	0.4%	0.2%	0.2%	0.6%
Streptococcus	3.9%	1.7%	1.0%	0.7%	4.2%	0.9%	1.3%	2.1%	2.6%	4.4%
Collinsella	0.0%	2.1%	0.7%	0.0%	1.1%	1.7%	1.1%	1.3%	1.9%	1.0%
Bifidobacterium	3.4%	12.3%	0.2%	0.0%	0.3%	13.0%	3.9%	7.9%	7.6%	6.7%
Bacteroides	5.7%	4.0%	5.6%	6.8%	7.1%	4.0%	4.9%	8.6%	3.5%	4.5%
Alistipes	2.2%	0.9%	0.8%	3.7%	1.1%	1.5%	1.2%	0.5%	1.1%	1.3%
Prevotella	0.0%	5.9%	0.0%	0.0%	2.0%	0.0%	2.6%	0.1%	0.0%	0.0%
Verrucomicrobium	2.8%	1.5%	0.7%	1.2%	1.0%	1.6%	0.0%	0.5%	0.3%	0.2%

Table 3.8 indicates individual abundances for genera across 10 samples. Of note is the abundance range for *Bifidobacterium* (0%-13%). None of the samples that were lacking particular groups appeared to have this made up with higher numbers of another, potentially related, group.

The variation seen among the ten samples suggested that some bacterial genera may be affected by age. Table 3.9 highlights calculated Spearman Rank correlation values. We observed a negative relationship between *Bifidobacterium* and subject age; and while there are no elderly individuals in this group and there have not been reports indicating any sort of change in bacterial

group abundances among young and middle aged adults, studies in elderly individuals indicate that bacterial groups do differ between younger children and elderly individuals [84]. Interestingly enough, aHLT06 had the highest abundance of *Bifidobacterium* at 13% and age of 44; aHLT02 being a close partner with 12% abundance however 22 years of age.

Table 3.9 Correlations of Bacterial Genera Abundances and Volunteer Age

Genus	Spearman Rank Correlation
Ruminococcus	0.02
Faecalibacterium	-0.31
Papillibacter	0.48
Roseburia	0.02
Bifidobacterium	-0.40
Bacteroides	0.48
Lachnospira	0.29
Eubacterium	-0.24
Anaerotruncus	0.14
Acetivibrio	0.50
Subdoligranulum	0.83
Streptococcus	-0.48
Dorea	-0.62
Anaerostipes	0.52
Clostridium	0.02
Coprococcus	0.07
Holdemania	0.43
Alistipes	0.43
Sutterella	-0.33
Collinsella	-0.31
Prevotella	-0.45
Verrucomicrobium	0.14

Table 3.9 shows Spearman Rank Correlation between age and abundance for each genus (aHLT09 and aHLT10 were omitted from correlation calculation as no age could be obtained for these two samples). There were no significant values among the correlation ($p<0.05$).

While it cannot be said that the correlations of these genera are statistically significant, this may be explained by the fact that volunteers within this group already harbor a well developed microbiota, some of the older volunteers could be supplementing their diets with probiotics, and the large gaps between ages of such a small population mask what might otherwise be a significant observation. Further characterization of these organisms (along with larger sample size) will aid in an understanding, not only of their role in digestion, but whether

the trend between abundance and age is real. This trend presents a notable possibility that certain components of any of these individuals' diets could contribute to higher numbers of organisms, for example, the case *Bifidobacterium*, as there exist many consumer products that are marketed as probiotic and contain multiple bacterial organisms, *Bifidobacterium* among them. To more accurately determine the extent that a decreasing/increasing pattern of bacterial organisms within a young-middle aged adult group would require a larger sample group with decreased jumps in age.

Detection of Known Phylo-species Inhabiting the Intestinal Tract

An advantage of using microarrays is the ability to accurately probe at the phylo-species level. The Microbiota Array interrogates for 775 possible phylo-species, 66 of which have been characterized and have actual binary names. Out of the 66 characterized phylo-species, 41 were detected to some degree. All 8 known *Bacteroides spp.* were found to be present (average range 0.01%-1.02%), along with all of *Bifidobacterium* (3 species, range 1.02%-1.73%), and *Faecalibacterium* (1 species, 1.53%). Known *Bifidobacterium* species *B. catenulatum,* and *B. longum* displayed the highest variation with respect to abundance, with at least 2-3 samples containing these genera in amounts of 2.3%-4.3% (data not shown), however as stated above this is potentially related to consumption of this genus in commercial probiotic products.

It was observed that each individual harbored a unique set of phylo-species, as evidenced by the occurrence of probesets in each lone sample. This observation led to the idea that there exists a "core" microbiota, or those phylo-species that are shared among all samples; a "disposable" microbiota in which phylo-species occur in more than one sample but not all; and finally a "unique" microbiota that exists in each particular sample. With this sample group it was seen that out of 775 possible phylo-species, 589 were detected in at least one sample. Of the 589 detected phylo-species, 384 were considered disposable and shared in more than one but not all

samples; there were 113 phylo-species present in all samples and labeled as the core microbiota, leaving 89 total phylo-species unique to one of the samples (Figure 3.5). The core microbiota was made up of 26 genera, mainly *Ruminococcus* species (37 species, 37% of core), with *Roseburia* (19 species, 17% of core), *Papillibacter* (10 species, 9% of core), and *Faecalibacterium* (9 species, 8% of core) making up the rest of the majority. Beyond this there were 9 genera with 2-4 species among the core (including *Bacteroides* with 2) and 13 with only 1 species in the core microbiota.

To test how well we sampled the population and whether the core we observed was a good representation of what might occur *in vivo*, sample presence and absence data were loaded into Matlab, and a rarefaction script was applied to determine average core phylo-species across all ten samples. Rarefaction compares the species richness of multiple samples, especially of differing sizes. The output of the rarefaction analysis indicates the number of species with respect to the number of individuals sampled. The curves that are created as a result indicate total species richness at the point of plateau. The Matlab script accomplishes this by using converted presence and absence values (P/A into 1/0). The binary data is imported into Matlab in matrix format (775x10), each row representing a phylo-species and column representing samples. The script then compares each sample combination out of the total sample population and averages the number of phylo-species that are present in the total population; then the sample population is increased by 1 and the script repeats the comparisons. In other words, the first loop of the script counts how many phylo-species are in each particular sample, then records the average of the sum. The next loop looks at all pair wise comparisons and again records the average of the sum. This process repeats, until the population size reaches the actual sample population, in this case 10.

Results showed that as sample size decreases, the average number of phylo-species that are shared among samples increases as would be expected (Figure 3.6). While the fit line in Figure 6 does not truly plateau at 113 phylo-species, this lack of leveling out indicates that there may be more phylo-species that make up the true human distal gut core.

Figure 3.5 Core, Shared, and Unique Microbiomes of the Adult Distal Colon

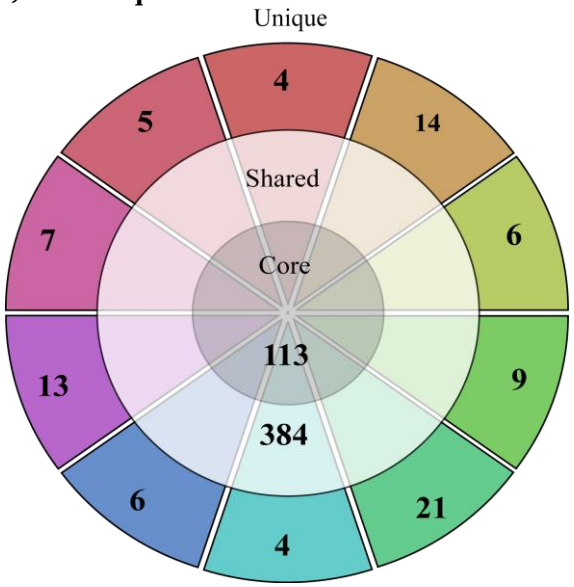


Figure 3.5 depicts phylo-species that are shared among all samples (Core), among more than one but not all (Shared), and those that are specific to each sample (Unique). Each individual triangle represents a singular sample, the middle donut is indicative of a "meshing" of phylo-species that appear to be present in more than one occurrence, and the inner circle contains those phylo-species that are significant enough to populate all samples**.**

Figure 3.6 Rarefaction of Healthy Adult Volunteers

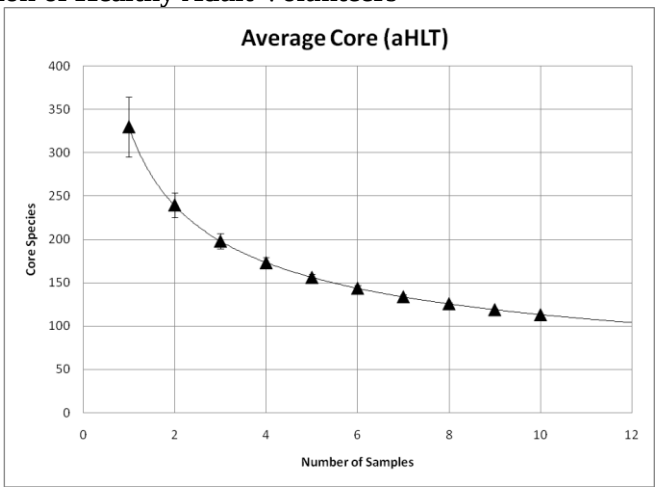


Figure 3.6 illustrates the rarefaction curve of healthy adult samples. As sample group size increases, Core Species decrease and begin to show signs of plateau. Values are average number of core species, error bars indicate Standard Error. The line represents a polynomial equation fit.

The distribution of unique, shared, and core phylo-species among this sample group may suggest that subtle differences between people are enough to separate individuals with respect to the microbiota; however, this cannot be determined nor concretely stated given the size of the sample group.

4. DISCUSSION

The development of the microarray and other such culture-independent techniques have allowed researchers to gain invaluable knowledge into the human intestinal microbiota. Here we used a custom designed Microbiota Array to assess the intestinal microbiota of 10 healthy adult volunteers. The benefits of using this technology include the ability to assess quantitatively the intestinal microbiota, whereas high-throughput sequencing and DGGE lack such ability. Another advantage of the microarray based approach is the relative ease with which this technology can be applied to larger populations. The construction and development of the Microbiota Array allowed us to quantify accurately the intestinal microbiota of healthy adult volunteers.

The initial analysis showed that Firmicutes was the dominant phylum in all samples, followed by Actinobacteria, and Bacteroidetes. This observation is in agreement with previous observations [26] [4], and shows that the higher taxonomical levels are stable and relatively free from change, further evidenced by looking at the degree of variation among the interrogated samples. The ratio of Firmicutes and Bacteroidetes shows that the former is approximately 8-9 times more abundant than the latter. This ratio agrees with results seen by Mariat et al., who observed that healthy adults had 11 times more Firmicutes than Bacteroidetes [34]. The results seen here are in good concordance with others. Future works would do well to include dietary intake information to assess further the extent to which diet contributes to the Firmicutes/Bacteroidetes ratio. Phylum level analysis is an initial point for examining the intestinal environment; however, it does not illustrate potential differences among populations.

At the class level we observed that Clostridia was dominant and represented the major constituent of the Firmicutes phylum. The second most abundant class was *Bacteroides* and was the only representative of the Bacteroidetes phylum. Since none of the volunteers had been on any antibiotics up to 3 months prior to the study, and none had reported any intestinal maladies, it

can be inferred that the presence of Clostridia is a product of nonpathogenic enteric flora, and not due to establishment of pathogenic *Clostridium* species, such as *Clostridium difficile* which is a known intestinal pathogen. Members of the class Clostridia aid the human host in fermentation and degradation of fiber sources in the diet. Bacteroidetes was the second most abundant class in all ten samples, which is consistent with referenced works and is justified, as this class of bacteria is beneficial in degrading polysaccharides [11]. It was noted that Actinobacteria was quite variable across the obtained samples, an interesting result as this class was represented mainly by the genus *Bifidobacterium* which is a component of many over-the-counter probiotic products, as is discussed below. Interestingly, there were cases where the number of detected phylo-species for a particular class was low even though the overall abundance of the class was high. The results from the class level are in good concordance with previous studies using fluorescent *in-situ* hybridization [85-87], and indicated to us that the class level gives enough information to allow insight into the intestinal microbiota, and represents an accurate account of major groups inhabiting the intestinal tract. The Microbiota Array is able to quantify large communities, which allowed us to determine the microbiota structure at the genus level as well as enabled us to assess the presence of known phylo-species. Halting analysis at the class level also hides trends and differences that potentially hold vital information for proper understanding of intestinal microbiota relationships.

At the genus level *Ruminococcus*, a member of Clostridia, was most abundant. In fact, the top four genera, in accordance with abundance, belonged to Clostridia. Of these four, *Ruminococcus* is known to have a substantial number of cellulolytic complexes which are utilized by species of this genus to attach to plant-derived dietary components and degrade cellulose, a key component of the plant cell wall structure [44]. This genus of bacteria is also known to be associated with solid-phase fecal matter. The dominant presence, along with the ability of this

genus to degrade plant polysaccharides and attach to solid-state fecal content, illustrates that this genus is a primary degrader in bacterial catabolism of ingested nutrients, supplying byproducts to other bacterial genera to be used as primary nutrient sources. *Faecalibacterium*, also a member of Clostridia, was next abundant. This genus' ability to utilize acetate and produce butyrate, suggests that species of this group coordinate the catabolism and utilization of byproducts from primary degraders to bacteria and colonocytes that use short chain fatty acids (mainly butyrate) as primary nutrient sources. A species of this genus has also been shown to increase production of interleukin-10, a potent anti-inflammatory cytokine, further solidifying the necessity of *Faecalibacterium* [35].

Bifidobacterium was the major representative of Actinobacteria, however; showed a tremendous amount of variation between samples. Potential reasons for such variation include a pattern of decreasing presence with age and consumption of probiotic products such as yogurt or supplements. A negative correlation was seen between volunteer age and abundance of *Bifidobacterium* and therefore it can be inferred that the degree of variation observed may indicate a decreasing abundance of this genus with age. A much more detailed analysis into the roles these genera play in intestinal homeostasis related to metabolism would shed much more light onto this area of research; unfortunately because of the harsh conditions imposed on these organisms in trying to cultivate them, many of them remain uncharacterized and thus our focus remains on those we are able to study. Lastly, it has been shown here that analysis of intestinal microbiota can be taken down to the genus and species levels. While it is acknowledged here that analysis at lower taxonomies shows more variations and less consistencies, however; it is important to include these observations. Furthermore, stopping analysis at the Phylum and Class levels results in incomplete findings, essentially camouflaging patterns and relationships that are potentially important.

The ability of the Microbiota Array to assess accurately large populations validates it as a valuable tool for studying intestinal microbiota. We observed that species contribution among all ten samples fit into three microbiome groups; a "core" microbiome in which species are present in all samples, a "shared" or interchangeable microbiome which can be defined as species that are present in more than 1 sample yet are not shared among all, and finally each sample was seen to harbor a "unique" set of phylo-species; a set of phylo-species that were only seen in each particular sample. To test the accuracy of the core microbiome observed, a rarefaction script was applied to the samples. The results of the rarefaction agreed with the observation and also showed that as the sample size increased, the difference between the calculated core at each population size decreased. From this result it can be inferred that our observed core microbiome is very close to what might be present *in vivo*. The core microbiome was comprised mainly of *Ruminococcus* (37 species) species, followed by *Roseburia* (19 species) and *Papillibacter* (10 species). This indicates that these genera provide important function to proper health of the intestinal tract, given that *Roseburia* and *Ruminococcus* both contain genes for catabolizing components of the human diet that would otherwise be indigestible [88] [89] [90]. What is interesting here is the presence of *Papillibacter* because there is such a lack of evidence illustrating this genus' purpose and function with respect to intestinal health.

The work described here illustrates the ability of the Microbiota Array to quantify accurately the intestinal microbiota of healthy adult volunteers. We analyzed and compared different taxonomical levels to better understand how variability and community structure differ at these lower levels. Previous studies have focused on higher phylogenetic groups, possibly because of such variability at the genus and phylo-species levels, and while it is not always advantageous to report such variability, it is important to accurately show true community

makeup in individuals. The advantage of the Microbiota Array over phylogenetic microarrays was the use of gDNA as starting material. Previous approaches utilized cDNA synthesized from RNA, however RNA is more unstable than DNA; therefore it stands to reason that using DNA as starting material yields results more consistent with actual microbial numbers. Admittedly there are limitations in this study, mainly the small groups size; and while it would be beneficial to have much larger sample groups there remains difficulties in acquiring samples from healthy volunteers. Our ability to assess such low taxonomies is superior to other molecular approaches because we can combine the results of the Microbiota Aray with other techniques such as NMR and 16S rRNA arrays. The abilities of the Microbiota Array, along with combination of differing groups and techniques, will allow us to accurately detect minute changes in intestinal homeostasis and open venues to better understanding of intestinal relationships. Once we understand the role of intestinal microbiota in human health and disease, we will be able to detect changes in health much sooner, resulting in shorter and more productive treatment strategies.

Appendix A

Wash Buffer A: Non-Stringent Wash Buffer
For 1L:

Reagent	Amount (mL)
SSPE (20X)	300
Tween-20 (10%)	1
Water	699

Filter through 0.2µm filter.

Wash Buffer B: Stringent Wash Buffer
For 1L:

Reagent	Amount (mL)
MES Buffer (12X)	83.3
NaCl (5M)	5.2
Tween-20 (10%)	1
Water	910.5

Filter through 0.2µm filter.

References

1. Whitman, W.B., D.C. Coleman, and W.J. Wiebe, *Prokaryotes: the unseen majority.* Proc Natl Acad Sci U S A, 1998. **95**(12): p. 6578-83.
2. Turroni, F., et al., *Human gut microbiota and bifidobacteria: from composition to functionality.* Antonie Van Leeuwenhoek, 2008. **94**(1): p. 35-50.
3. Sghir, A., *Molecular diversity and phylogeny of human colonic bacteria.* Molecular Ecology in Gastrointestinal Systems, 1999.
4. Eckburg, P.B., et al., *Diversity of the human intestinal microbial flora.* Science, 2005. **308**(5728): p. 1635-8.
5. Suau, A., et al., *Fusobacterium prausnitzii and related species represent a dominant group within the human fecal flora.* Syst Appl Microbiol, 2001. **24**(1): p. 139-45.
6. Bik, E.M., et al., *Molecular analysis of the bacterial microbiota in the human stomach.* Proc Natl Acad Sci U S A, 2006. **103**(3): p. 732-7.
7. Aas, J.A., et al., *Defining the normal bacterial flora of the oral cavity.* J Clin Microbiol, 2005. **43**(11): p. 5721-32.
8. Dethlefsen, L., et al., *Assembly of the human intestinal microbiota.* Trends Ecol Evol, 2006. **21**(9): p. 517-23.
9. Leitch, E.C., et al., *Selective colonization of insoluble substrates by human faecal bacteria,* in *Environ Microbiol*. 2007. p. 667-79.
10. Hayashi, H., M. Sakamoto, and Y. Benno, *Fecal microbial diversity in a strict vegetarian as determined by molecular analysis and cultivation.* Microbiol Immunol, 2002. **46**(12): p. 819-31.
11. Liszt, K., et al., *Characterization of bacteria, clostridia and Bacteroides in faeces of vegetarians using qPCR and PCR-DGGE fingerprinting.* Ann Nutr Metab, 2009. **54**(4): p. 253-7.
12. Walker, A.W., et al., *pH and peptide supply can radically alter bacterial populations and short-chain fatty acid ratios within microbial communities from the human colon.* Appl Environ Microbiol, 2005. **71**(7): p. 3692-700.
13. Vrieze, A., et al., *The environment within: how gut microbiota may influence metabolism and body composition.* Diabetologia. **53**(4): p. 606-13.
14. Turnbaugh, P.J., et al., *An obesity-associated gut microbiome with increased capacity for energy harvest.* Nature, 2006. **444**(7122): p. 1027-31.
15. Ley, R.E., et al., *Microbial ecology: human gut microbes associated with obesity.* Nature, 2006. **444**(7122): p. 1022-3.
16. Backhed, F., et al., *The gut microbiota as an environmental factor that regulates fat storage.* Proc Natl Acad Sci U S A, 2004. **101**(44): p. 15718-23.
17. Mueller, S., et al., *Differences in fecal microbiota in different European study populations in relation to age, gender, and country: a cross-sectional study.* Appl Environ Microbiol, 2006. **72**(2): p. 1027-33.
18. Flint, H.J., et al., *Interactions and competition within the microbial community of the human colon: links between diet and health.* Environ Microbiol, 2007. **9**(5): p. 1101-11.
19. Benus, R.F., et al., *Association between Faecalibacterium prausnitzii and dietary fibre in colonic fermentation in healthy human subjects.* Br J Nutr, 2010: p. 1-8.
20. Xu, J., et al., *Evolution of symbiotic bacteria in the distal human intestine.* PLoS Biol, 2007. **5**(7): p. e156.
21. Li, M., et al., *Symbiotic gut microbes modulate human metabolic phenotypes.* Proc Natl Acad Sci U S A, 2008. **105**(6): p. 2117-22.

22. Comstock, L.E., *Importance of glycans to the host-bacteroides mutualism in the mammalian intestine.* Cell Host Microbe, 2009. **5**(6): p. 522-6.
23. Possemiers, S., et al., *The intestinal environment in health and disease - recent insights on the potential of intestinal bacteria to influence human health.* Curr Pharm Des, 2009. **15**(18): p. 2051-65.
24. Rautava, S. and W.A. Walker, *Commensal bacteria and epithelial cross talk in the developing intestine.* Curr Gastroenterol Rep, 2007. **9**(5): p. 385-92.
25. Obata, T., et al., *Indigenous opportunistic bacteria inhabit mammalian gut-associated lymphoid tissues and share a mucosal antibody-mediated symbiosis.* Proc Natl Acad Sci U S A, 2010. **107**(16): p. 7419-24.
26. Rajilic-Stojanovic, M., H. Smidt, and W.M. de Vos, *Diversity of the human gastrointestinal tract microbiota revisited.* Environ Microbiol, 2007. **9**(9): p. 2125-36.
27. Hooper, L.V. and A.J. Macpherson, *Immune adaptations that maintain homeostasis with the intestinal microbiota.* Nat Rev Immunol, 2010. **10**(3): p. 159-69.
28. Zasloff, M., *Antimicrobial peptides of multicellular organisms.* Nature, 2002. **415**(6870): p. 389-95.
29. Tanabe, H., et al., *Denatured human alpha-defensin attenuates the bactericidal activity and the stability against enzymatic digestion.* Biochem Biophys Res Commun, 2007. **358**(1): p. 349-55.
30. Kraehenbuhl, J.P. and M. Corbett, *Immunology. Keeping the gut microflora at bay.* Science, 2004. **303**(5664): p. 1624-5.
31. Clarke, M.B. and V. Sperandio, *Events at the host-microbial interface of the gastrointestinal tract III. Cell-to-cell signaling among microbial flora, host, and pathogens: there is a whole lot of talking going on.* Am J Physiol Gastrointest Liver Physiol, 2005. **288**(6): p. G1105-9.
32. Kaper, J.B. and V. Sperandio, *Bacterial cell-to-cell signaling in the gastrointestinal tract.* Infect Immun, 2005. **73**(6): p. 3197-209.
33. Pai, R. and G. Kang, *Microbes in the gut: a digestable account of host-symbiont interactions.* Indian J Med Res, 2008. **128**(5): p. 587-94.
34. Mariat, D., et al., *The Firmicutes/Bacteroidetes ratio of the human microbiota changes with age.* BMC Microbiol, 2009. **9**: p. 123.
35. Sokol, H., et al., *Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients.* Proc Natl Acad Sci U S A, 2008. **105**(43): p. 16731-6.
36. Segain, J.P., et al., *Butyrate inhibits inflammatory responses through NFkappaB inhibition: implications for Crohn's disease.* Gut, 2000. **47**(3): p. 397-403.
37. Pryde, S.E., et al., *The microbiology of butyrate formation in the human colon.* FEMS Microbiol Lett, 2002. **217**(2): p. 133-9.
38. Walker, A.W., et al., *The species composition of the human intestinal microbiota differs between particle-associated and liquid phase communities.* Environ Microbiol, 2008. **10**(12): p. 3275-83.
39. Hooper, L.V., T. Midtvedt, and J.I. Gordon, *How host-microbial interactions shape the nutrient environment of the mammalian intestine.* Annu Rev Nutr, 2002. **22**: p. 283-307.
40. Xu, J. and J.I. Gordon, *Inaugural Article: Honor thy symbionts.* Proc Natl Acad Sci U S A, 2003. **100**(18): p. 10452-9.
41. Backhed, F., et al., *Host-bacterial mutualism in the human intestine.* Science, 2005. **307**(5717): p. 1915-20.
42. Stappenbeck, T.S., L.V. Hooper, and J.I. Gordon, *Developmental regulation of intestinal angiogenesis by indigenous microbes via Paneth cells.* Proc Natl Acad Sci U S A, 2002. **99**(24): p. 15451-5.

43. Ley, R.E., et al., *Obesity alters gut microbial ecology*, in *Proc Natl Acad Sci U S A*. 2005. p. 11070-5.
44. Flint, H.J., et al., *Polysaccharide utilization by gut bacteria: potential for new insights from genomic analysis*, in *Nat Rev Microbiol*. 2008. p. 121-31.
45. Robert, C. and A. Bernalier-Donadille, *The cellulolytic microflora of the human colon: evidence of microcrystalline cellulose-degrading bacteria in methane-excreting subjects.* FEMS Microbiol Ecol, 2003. **46**(1): p. 81-9.
46. Abell, G.C., et al., *Phylotypes related to Ruminococcus bromii are abundant in the large bowel of humans and increase in response to a diet high in resistant starch.* FEMS Microbiol Ecol, 2008. **66**(3): p. 505-15.
47. Tlaskalova-Hogenova, H., et al., *Commensal bacteria (normal microflora), mucosal immunity and chronic inflammatory and autoimmune diseases.* Immunol Lett, 2004. **93**(2-3): p. 97-108.
48. Palmer, C., et al., *Development of the human infant intestinal microbiota.* PLoS Biol, 2007. **5**(7): p. e177.
49. Dominguez-Bello, M.G., et al., *Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns.* Proc Natl Acad Sci U S A, 2010. **107**(26): p. 11971-5.
50. Biasucci, G., et al., *Cesarean delivery may affect the early biodiversity of intestinal bacteria.* J Nutr, 2008. **138**(9): p. 1796S-1800S.
51. Biasucci, G., et al., *Mode of delivery affects the bacterial community in the newborn gut.* Early Hum Dev, 2010.
52. Penders, J., et al., *Factors influencing the composition of the intestinal microbiota in early infancy.* Pediatrics, 2006. **118**(2): p. 511-21.
53. Makivuokko, H., et al., *The effect of age and non-steroidal anti-inflammatory drugs on human intestinal microbiota composition.* Br J Nutr, 2010. **103**(2): p. 227-34.
54. Woodmansey, E.J., et al., *Comparison of compositions and metabolic activities of fecal microbiotas in young adults and in antibiotic-treated and non-antibiotic-treated elderly subjects.* Appl Environ Microbiol, 2004. **70**(10): p. 6113-22.
55. Woodmansey, E.J., *Intestinal bacteria and ageing.* J Appl Microbiol, 2007. **102**(5): p. 1178-86.
56. Enck, P., et al., *The effects of ageing on the colonic bacterial microflora in adults.* Z Gastroenterol, 2009. **47**(7): p. 653-8.
57. Claesson, M.J., et al., *Microbes and Health Sackler Colloquium: Composition, variability, and temporal stability of the intestinal microbiota of the elderly.* Proc Natl Acad Sci U S A, 2010.
58. Costello, E.K., et al., *Bacterial community variation in human body habitats across space and time.* Science, 2009. **326**(5960): p. 1694-7.
59. Dicksved, J., et al., *Molecular analysis of the gut microbiota of identical twins with Crohn's disease.* ISME J, 2008. **2**(7): p. 716-27.
60. Turnbaugh, P.J., et al., *A core gut microbiome in obese and lean twins.* Nature, 2009. **457**(7228): p. 480-4.
61. Turnbaugh, P.J., et al., *Organismal, genetic, and transcriptional variation in the deeply sequenced gut microbiomes of identical twins.* Proc Natl Acad Sci U S A, 2010. **107**(16): p. 7503-8.
62. Willing, B., et al., *Twin studies reveal specific imbalances in the mucosa-associated microbiota of patients with ileal Crohn's disease.* Inflamm Bowel Dis, 2009. **15**(5): p. 653-60.

63. Zoetendal, E.G., et al., *DNA isolation protocols affect the detection limit of PCR approaches of bacteria in samples from the human gastrointestinal tract.* Syst Appl Microbiol, 2001. **24**(3): p. 405-10.
64. Ley, R.E., D.A. Peterson, and J.I. Gordon, *Ecological and evolutionary forces shaping microbial diversity in the human intestine.* Cell, 2006. **124**(4): p. 837-48.
65. Hehemann, J.H., et al., *Transfer of carbohydrate-active enzymes from marine bacteria to Japanese gut microbiota.* Nature, 2010. **464**(7290): p. 908-12.
66. Moore, W.E. and L.V. Holdeman, *Human fecal flora: the normal flora of 20 Japanese-Hawaiians.* Appl Microbiol, 1974. **27**(5): p. 961-79.
67. Moore, W.E. and L.H. Moore, *Intestinal floras of populations that have a high risk of colon cancer.* Appl Environ Microbiol, 1995. **61**(9): p. 3202-7.
68. Wang, R.F., et al., *Design and evaluation of oligonucleotide-microarray method for the detection of human intestinal bacteria in fecal samples.* FEMS Microbiol Lett, 2002. **213**(2): p. 175-82.
69. Wehkamp, J., et al., *Paneth cell antimicrobial peptides: topographical distribution and quantification in human gastrointestinal tissues.* FEBS Lett, 2006. **580**(22): p. 5344-50.
70. Kerckhoffs, A.P., et al., *Lower Bifidobacteria counts in both duodenal mucosa-associated and fecal microbiota in irritable bowel syndrome patients.* World J Gastroenterol, 2009. **15**(23): p. 2887-92.
71. Codling, C., et al., *A molecular analysis of fecal and mucosal bacterial communities in irritable bowel syndrome.* Dig Dis Sci, 2009. **55**(2): p. 392-7.
72. Conte, M.P., et al., *Gut-associated bacterial microbiota in paediatric patients with inflammatory bowel disease.* Gut, 2006. **55**(12): p. 1760-7.
73. Guarner, F. and J.R. Malagelada, *Gut flora in health and disease.* Lancet, 2003. **361**(9356): p. 512-9.
74. Palmer, C., et al., *Rapid quantitative profiling of complex microbial populations.* Nucleic Acids Res, 2006. **34**(1): p. e5.
75. Paliy, O., et al., *High-throughput quantitative analysis of the human intestinal microbiota with a phylogenetic microarray.* Appl Environ Microbiol, 2009. **75**(11): p. 3572-9.
76. Harmsen, H.J., et al., *Extensive set of 16S rRNA-based probes for detection of bacteria in human feces.* Appl Environ Microbiol, 2002. **68**(6): p. 2982-90.
77. Hattori, M. and T.D. Taylor, *The human intestinal microbiome: a new frontier of human biology.* DNA Res, 2009. **16**(1): p. 1-12.
78. Gill, S.R., et al., *Metagenomic analysis of the human distal gut microbiome.* Science, 2006. **312**(5778): p. 1355-9.
79. Rajilic-Stojanovic, M., et al., *Development and application of the human intestinal tract chip, a phylogenetic microarray: analysis of universally conserved phylotypes in the abundant microbiota of young and elderly adults.* Environ Microbiol, 2009. **11**(7): p. 1736-51.
80. Rigsbee, L., et al., *Optimizing the analysis of human intestinal microbiota with phylogenetic microarray.* FEMS Microbiol Ecol, 2010. **75**(2): p. 332-42.
81. Rainer, J., et al., *CARMAweb: comprehensive R- and bioconductor-based web service for microarray data analysis.* Nucleic Acids Res, 2006. **34**(Web Server issue): p. W498-503.
82. Ramirez-Flandes, S. and O. Ulloa, *Bosque: integrated phylogenetic analysis software.* Bioinformatics, 2008. **24**(21): p. 2539-41.
83. Durban, A., et al., *Assessing gut microbial diversity from feces and rectal mucosa.* Microb Ecol, 2010. **61**(1): p. 123-33.
84. He, T., et al., *Composition of Faecal Microbiota of Elderly People.* Microbial Ecology in Health and Disease, 2003. **15**(4): p. 153-159.

85. Rigottier-Gois, L., et al., *Fluorescent hybridisation combined with flow cytometry and hybridisation of total RNA to analyse the composition of microbial communities in human faeces using 16S rRNA probes*, in *FEMS Microbiol Ecol*. 2003. p. 237-45.
86. Sokol, H., et al., *Specificities of the fecal microbiota in inflammatory bowel disease*, in *Inflamm Bowel Dis*. 2006. p. 106-11.
87. Lay, C., et al., *Colonic microbiota signatures across five northern European countries*, in *Appl Environ Microbiol*. 2005. p. 4153-5.
88. Hippe, B., et al., *Quantification of butyryl CoA:acetate CoA-transferase genes reveals different butyrate production capacity in individuals according to diet and age.* FEMS Microbiol Lett, 2010.
89. Scott, K.P., et al., *Microbes and Health Sackler Colloquium: Substrate-driven gene expression in Roseburia inulinivorans: Importance of inducible enzymes in the utilization of inulin and starch.* Proc Natl Acad Sci U S A, 2010.
90. Tap, J., et al., *Towards the human intestinal microbiota phylogenetic core.* Environ Microbiol, 2009. **11**(10): p. 2574-84.