

Review

Understanding the Extent and Sources of Variation in Gut Microbiota Studies; a Prerequisite for Establishing Associations with Disease

Volker Mai^{1,*}, Maria Ukhanova¹ and David J. Baer²

¹ Microbiology and Cell Science, Emerging Pathogens Institute, University of Florida, Gainesville, FL 32611, USA; E-Mail: mukhanova@ufl.edu

² US Department of Agriculture, Agricultural Research Service, Beltsville Human Nutrition Research Center, Beltsville, MD 20705, USA; E-Mail: David.Baer@ars.usda.gov

* Author to whom correspondence should be addressed; E-Mail: vmai@ufl.edu; Tel.: +1-352-273-9398; Fax: +1-352-392-5922.

Received: 16 April 2010; in revised form: 10 August 2010 / Accepted: 12 August 2010 /

Published: 27 August 2010

Abstract: Humans harbor distinct commensal microbiota at various anatomic sites. There has been renewed interest in the contributions of microbiota activities to human health and disease. The microbiota of the gut is the most complex of all anatomic sites in terms of total numbers of bacteria that interact closely with the mucosal immune system and contribute various functions to host physiology. Especially in the proximal large intestine a diverse microbiota ferments complex substrates such as dietary fiber and host mucins, but also metabolizes bile acids and phytoestrogens that reach the large intestine. It is now well established that microbiota composition differs between but over time also within individuals. However, a thorough understanding of the sources of variations in microbiota composition, which is an important requirement for large population based microbiota studies is lacking. Microbiota composition varies depending on what kind of sample is collected, most commonly stool samples, stool swabs or superficial rectal or intestinal biopsies, and the time of collection. Microbiota dynamics are affected by life style factors including diet and exercise that determine what nutrients reach the proximal colon and how fast these nutrients pass through (transit time). Here we review sample collection issues in gut microbiota studies and recent findings about dynamics in microbiota composition. We recommend standardizing human microbiota analysis methods to facilitate comparison and

pooling between studies. Finally, we outline a need for prospective microbiota studies in large human cohorts.

Keywords: dynamics; gut microbiota; metagenome; community structure

1. Introduction

There has been a recurring interest in studying associations between the human commensal gut microbiota and health, ever since Metchnikoff suggested in the early twentieth century that microbiota composition can be positively modified by consuming ‘beneficial’ microbes, such as milk fermenting lactic acid bacteria [1]. Although microbiota have been studied ever since, using mainly conventional culture based methods until the development of molecular tools towards the end of the last century, it is the availability of high throughput sequencing technologies that now provides the tools needed for in depth studies. Through its immense metabolic capabilities the gut microbiota contributes to human physiology by transforming complex nutrients such as dietary fiber or intestinal mucins that otherwise would be lost to the human host into simple sugars, short chain fatty acids and other nutrients that can be absorbed [2]. Furthermore, the microbiota produces some essential vitamins including vitamin K, vitamin B12 and folic acid, contributes to intestinal bile acid metabolism and recirculation, transforms potential carcinogens such as *N*-nitroso compounds [NOCs] and heterocyclic amines [HCAs] and activates bioactive compounds including phytoestrogens [3,4]. Differences in environmental factors including diet as well as hosts genetics are thought to contribute to microbiota diversity [5].

1.1. Gut Microbiota and GI Diseases

Inflammatory bowel diseases [(IBD), ulcerative colitis and Crohn’s disease (CD)] are distinct chronic inflammatory disorders of the gastrointestinal tract. The etiology of IBD is not known, but it is believed that in at least a subset of genetically susceptible individuals an abnormal immune response directed against gut bacteria contributes [6]. CD is heterogeneous with varying locations of disease activity and a broad spectrum of disease severity. Immunosuppressive medications are often used but the remission rates for immune modulators and anti-TNF therapy are only about 40%. IBD is a disease of westernized, industrialized countries but is now being increasingly recognized in developing countries including India and China, is increasingly occurring in children under the age of 10 years old, and in African Americans. Genome wide association studies have identified numerous CD risk loci including genes in the innate immune response that directly interact with intracellular bacteria such as NOD2 (CARD15) and the autophagy proteins, ATG16L1 and IRGM [7]. The increasing incidence of IBD cannot be explained by genetics and the environmental influences remain speculative but gut microbiota remains a focus of attention [6]. To date IBD has been linked to microbiota composition in a variety of studies [8,9] and successful interventions using antibiotics, prebiotics and probiotics have been reported. In addition to reports of differences in fecal microbiota composition mucosa adherent bacteria also differ between IBD cases and healthy controls and sample fixation is crucial for detecting this difference [10].

There is emerging evidence for the potential contribution of gut microbiota to colorectal carcinogenesis (CRC) [11-13]. Recent, data from our group suggests that specific groups of bacteria can be more frequently detected in subjects harboring large intestinal adenomas when compared to normal controls [14]. Colorectal adenomas are not necessarily good surrogate markers for CRC risk, only 10% of polyps are thought to have the potential to proceed to CRC. CRC is a relatively rare disease that can develop after decades of exposure to risk factors that might include distorted microbiota diversity or as of yet unknown pathogens that cause chronic infections. Determination of an association between gut microbiota and CRC will require long term prospective studies in large human cohorts. In order to be sufficiently powered such prospective cohort studies would have to follow thousands of individuals of many years to observe a sufficient number of cases to allow for an appropriate analysis.

Necrotizing enterocolitis (NEC) is a disease of preterm infants that recently has received interest for potential contributions of microbiota [15,16]. Preterm infants are often delivered by C-section, they frequently receive antibiotics and their feeding habits differ from those of full term infants [17]. Under these circumstances, the normal development of gut microbiota [18] likely is distorted and possibly can initiate a disproportional immune response that results in NEC. Currently, prospective cohort studies are underway that are aimed at elucidating this association.

A crucial requirement for the design of future studies aimed at correlating microbiota with various GI and other diseases is a thorough quantitative understanding of the variation in microbiota composition between individuals but also within them over time is. Below we discuss recent advances as well as remaining shortcomings in our knowledge of microbiota variations.

1.2. Gut microbiota Variation and Dynamics

Currently neither the extent of gut microbiota diversity and activities nor its dynamics within and between individuals are fully understood [2]. Bacteria belonging to a few phyla, particularly Firmicutes and Bacteroidetes, dominate in most healthy individuals [19-21]. Estimates for the total number of bacterial species comprising the collective gut microbiome have recently been extended up to 40,000 [22], but largely due to the large amount of emerging sequence data the bacterial species concept is prone to undergo revision. A recent study has explored variations in microbiota composition in nine subjects at different anatomic sites at four time points one day and three months apart [23]. This study showed different microbiota dynamics for each anatomic site. Unifrac, a bioinformatics tool for 16S rRNA sequence data analysis that uses shared branch length in a phylogenetic tree generated with the combined sequence set of individual samples [24,25] was used to determine microbiota composition. Based on the Unifrac distance metric Costello *et al.* suggested that gut microbiota community exhibited minimal intra-individual variation over 90 days [23]. However, that claim appears somewhat subjective as their data shows variation both within (Unifrac distance = 0.58) as well as between unrelated individuals (Unifrac distance = 0.72). Another extensive gut microbiota study in twin pairs and their mothers (154 individuals) reported an intra-individual Unifrac distance of 0.69 over an average of 57 days, which compared to a Unifrac distance of 0.8 between unrelated individuals [26]. As the Unifrac distance within individuals in one study is very close to the Unifrac distance between unrelated individuals in the other study it appears arbitrary to assign labels such as

minimal or extensive to the variation observed in these studies. What is clear from this body of work that has tremendously expanded our knowledge of human microbiota diversity is that each individual retains a distinct microbiota over time. This finding is encouraging for future efforts aimed at correlating microbiota diversity with health status. Overall microbiota diversity, which is expressed well by the Unifrac metric, is only one potential correlate with disease. The presence or absence of potentially beneficial/pathogenic microbes is another measure that can be linked with health status. As sequencing technologies have improved to allow for the generation of hundreds of thousands (454) or even millions (Illumina, Solid) reads/run the correct binning of sequences into operational taxonomic units is a daunting task that will require the development of new algorithms [27] and their implementation into user friendly software packages.

1.3. The Core Microbiome

The combined microbial gene pool, studied by metagenomic approaches, exceeds the complexity of the human genome, extending the metabolic abilities of the human/microbiota “supra- or superorganism” [8,28]. Extensive metagenomic sequencing has now given us deeper insight into prevalent intestinal microbial genes, identifying a core microbiome [26,29]. Although the bacterial species composition varies among individuals and over time (see above), the core activities encoded by the microbiome appear more consistent. This is not surprising as most microbes share a minimal gene set and those microbes best adapted to the gut environment either share a set of genes allowing them to outcompete other microbes or develop into synergistic partners. While it might not be relevant for studies aimed at associating microbiota metabolites with disease, such as recent work using microbial metabolite profiling in urine [30,31], interactions with the mucosal immune system likely depend heavily on structure. Thus, studying both the diversity and species composition of the microbiota as well as its metabolic capabilities has promise for investigating potential associations with health and disease.

2. Barriers in Population Based Microbiota Studies

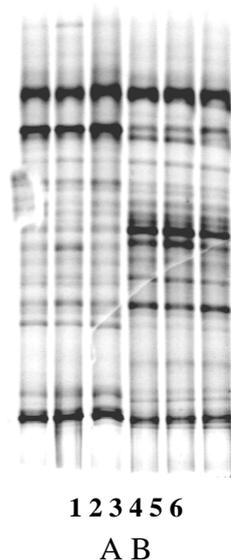
2.1. Gut Microbiota Sampling

There has been an ongoing discussion in the field regarding appropriate sample collection. Specifically, mucosa attached microbiota, as determined in biopsy samples, and fecal microbiota, collected as fecal swabs or whole stools, have been utilized [32,33]. Although feces or fecal swabs are the most convenient samples they do not accurately reflect the microbiota composition or activities in the proximal large intestine. Due to changes in substrates, pH and water content *etc.* the composition of the microbiota changes while intestinal contents are slowly moved through the colon, collected in the rectum and excreted. Similarly, rectal biopsies are not representative of the proximal gut microbiota. Colon biopsies also do not represent microbiota in its physiologic state as the extensive colon preparation to clean intestinal contents removes some of the outer mucus layer and thus mucosa attached microbes as well as their normal attachment sites. The bowel prep has been shown to affect markers of proliferation in intestinal epithelium [34], suggesting that it affects colon physiology and

thus likely the mucosa and associated microbiota. Indeed, a PEG bowel preparation results in moderate to severe loss of superficial mucus in 96% of patients [35].

There has been some work aimed at determining the biostructure of microbiota in a stool sample. Using pinched fecal cylinders to investigate changes in microbiota structure from the outside of formed feces (closest to the mucosa while inside the gut) towards the inside (luminal bacteria) Swidsinski *et al.* reported a clear structure that was distorted in subjects with idiopathic diarrhea [36]. The observation of structure would argue against the validity of using fecal swabs to evaluate overall microbiota composition. Feces form while the luminal contents move through the large intestine where water is reabsorbed. Luminal contents are then mixed in the rectum before they are evacuated as formed feces. The portion of the feces first evacuated might harbor microbiota different from that at the end of the pellet. To address this concern, we compared microbiota profiles in the front and the end of two fecal pellets from two individuals with the homogenized samples by DGGE (Figure 1). Pearson and Dice coefficient based similarity matrices were generated and Shannon Diversity indexes calculated to compare microbiota composition and diversity in the different parts of the stool samples. For both stools, microbiota was similar in all three parts. Thus, a pea sized piece from any part of the pellet, which is sufficient for most microbiota analyses, can be collected by individuals and transported safely in small vials containing DNA preserving solutions. This circumvents the need for subjects to deliver fresh samples within a few hours of defecation to a laboratory.

Figure 1. Variation within stool sample. DGGE profiles (V6-V8) region of two stool samples (A and B) in the front of the fecal pellet (Lanes 1 and 4), the back (lanes 2 and 5) and a homogenized sample (lanes 3 and 6). DGGE was performed on an 8% [wt/vol] acrylamide gel with a gradient from 40% at the top to 50% at the bottom at a temperature of 60 °C. 100% denaturing conditions were defined as 7M urea and 40% formamide. Gels were run for 16 h at 65 V and stained with Cyber Green. Gels were scanned in using Quantity One (Biorad) and analyzed using Diversity Database software (Biorad). Microbiota diversity in profiles was compared (i) by generating dendograms based on similarity matrixes calculated using Pearson and Dice coefficients and (ii) by calculating the Shannon Diversity index for each profile based on the number and intensity of bands.



Gut evacuates, such as those induced by a colon prep, likely contain both luminal as well as mucosa attached microbiota. Although it might not be realistic to attempt collecting such samples in large population based studies, there clearly is a need to improve and validate gut microbiota sample collection procedures.

2.2. Understanding Microbiota Dynamics in Studies with Controlled Diets

Dietary intake is the origin for most of the substrates that reach the large intestine where they are available for microbial fermentation. Other substrates available for bacterial fermentation include mucins and bile acids that are released into the lumen. Thus, variations in diet are likely one of the most significant sources of microbiota dynamics. Exercise affects host physiology including hormone levels and transit time and thus is another factor that can shape microbiota composition. Genetic differences clearly can affect gut microbial ecology [37] but recent studies in twins and their mothers do not suggest that genetic differences are a major source of microbiota variation in humans [26].

We have previously reported microbiota variation in a mouse study with strictly controlled diets and an exercise regimen [11]. Mice for this study were bred in dedicated breeding colonies to reduce variation in genetics and age. Mice were single housed to avoid competition for food as a source of variation. In one group that was calorie restricted mice even received the same amounts of the same diet. As expected the different diets and exercise affected microbiota composition. Somewhat surprisingly, mice within each group also showed some differences in microbiota composition. This suggests that there are source other than host genetics, age, diet and exercise that contribute to microbiota dynamics. We hypothesize that initial gut colonization is a chance event that sets up a cascade of microbiota establishment that contributes to differences between individuals.

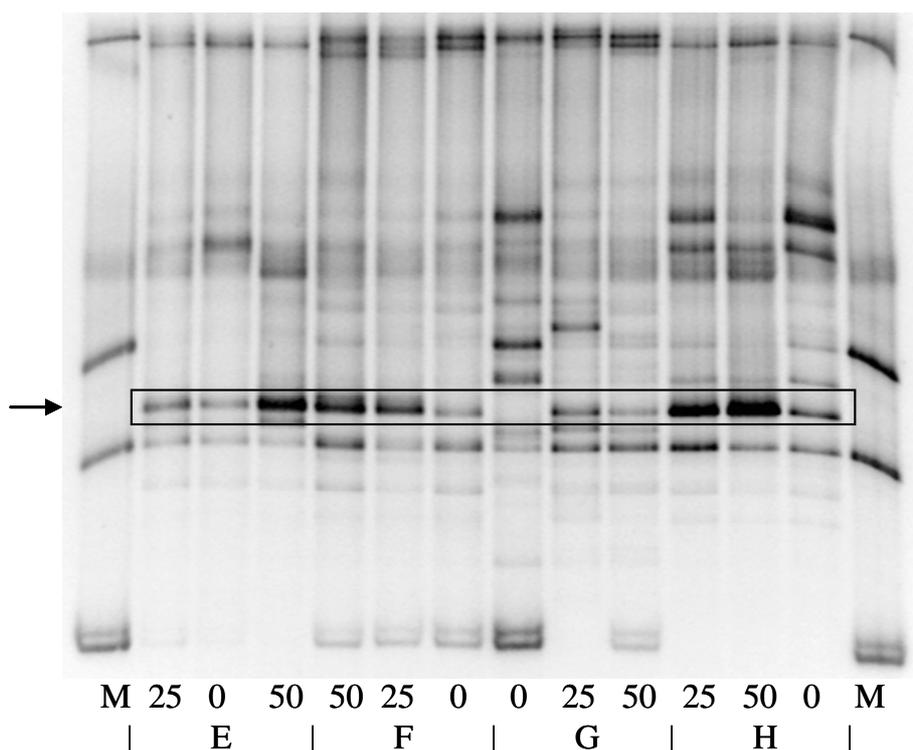
The effects of drastic dietary changes have recently been investigated in a mouse model inoculated with human fecal microbiota. Changing diets from a low-fat, plant rich diet to one high in fat and refined sugars shifted microbiota composition within one day [38]. Changes were fairly consistent in individual mice, but even on the class level differences between mice were clearly detectable.

From the variation observed in animals it is not surprising that microbiota also varies in individuals that consume controlled diets. Human feeding studies are frequently performed to study the effects of a dietary intervention on various aspects of human physiology. Subjects usually consume some diets under supervision, but frequently take meals to work or prepared food packages home for the weekend. A crossover design, in which individuals serve as their own controls, reduces inter-individual variation. Using such a study design we have previously reported effects of black tea drinking on gut microbiota composition [39]. Effects of tea drinking on microbiota, likely due to the tea polyphenols that have antimicrobial effects, were seen within two weeks using quantitative 16S rRNA based fluorescent *in situ* hybridization (FISH). We also observed in this study that microbiota variations within less than two weeks in individuals within the same intervention period in which they consumed the same amounts of the same foods in a weekly rotation.

Dietary fibers represent some of the main substrates that reach the large intestine. Resistant maltodextrin (RM) is a low caloric food ingredient that in many ways behaves similar to dietary fiber. RM contains a mixture of oligosaccharides and polysaccharides. In a double-blinded randomized crossover feeding study we have recently determined that RM supplementation has a prebiotic effect

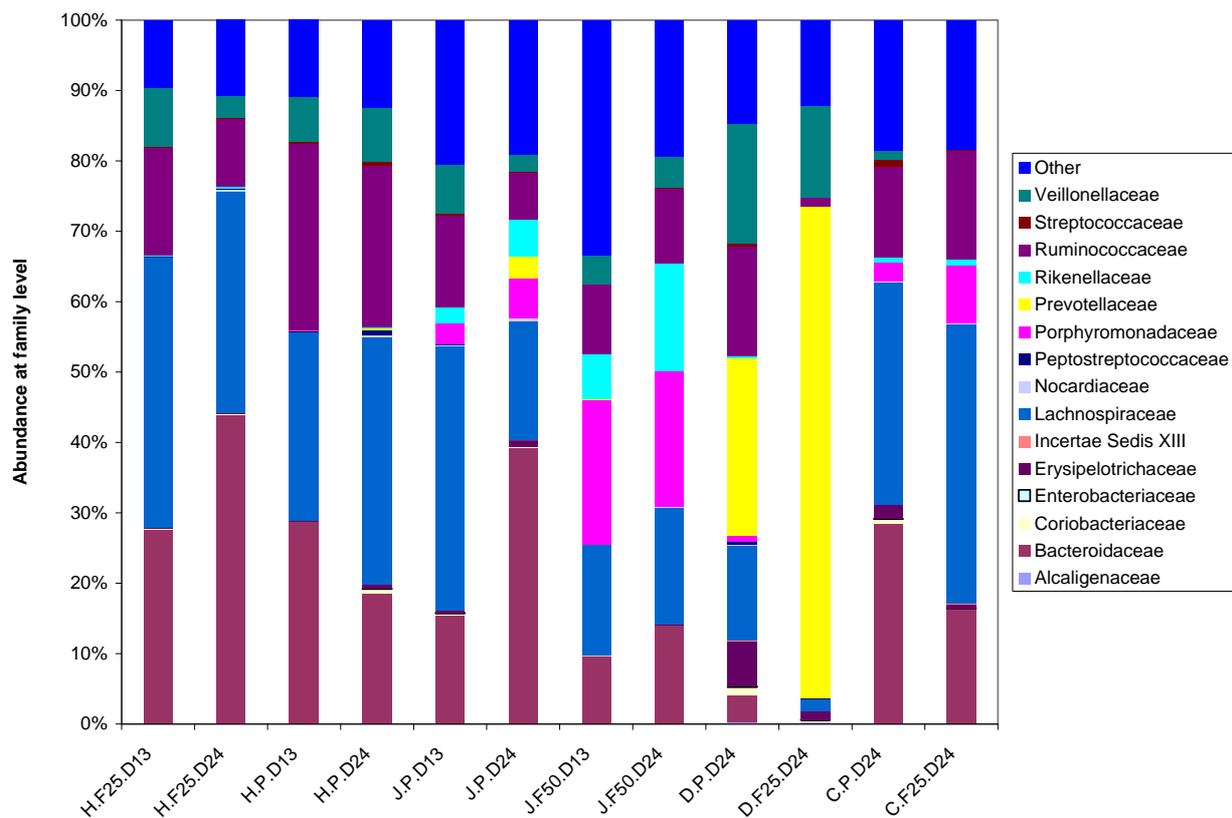
and enriches for Bifidobacteria (unpublished). Effects of RM could be detected by the appearance of a strong band generated by simple DGGE profiling (Figure 2). To determine variation within subjects we also performed an in depth 16S rRNA analysis using a bar-coded 454 pyrosequencing approach. For four individuals, two with four time points each and two others with two time points, we obtained a total of 283,672 sequence reads with an average read number of 23,639/sample and an average length of 225 nucleotides. The varying proportions of bacterial families between subjects showed that each of the subjects harbored a unique microbiota (Figure 3). Although subjects consumed a controlled diet, the proportions of dominant bacterial families changed to some extent during the same intervention period.

Figure 2. Effects of RM on microbiota. DGGE profiles for subjects E, F, G, H on Day 24 during placebo (0), RM 25 (25) and RM 50 (50) intervention periods. M = molecular marker, black arrow indicates position of band increased during RM supplementation (box). DNA was eluted from the marked band and submitted to sequencing. DGGE gels were run and diversity in each profile was calculated as described in Figure 1. Profiles for each individual are presented in the order that they received each intervention.



Although we obtained an average of 23,639 reads/sample, for two of the individuals approximately 100,000 total sequences reads, we still did not reach saturation. This observation indicates that even with the current high throughput sequencing methods detecting rare bacterial species with high confidence will be difficult to achieve in large population studies with many samples. Although identifying rare species might not be crucial for describing overall diversity in microbiota composition it might be necessary to identify rare potentially opportunistic pathogens that cause low level but persistent localized infections.

Figure 3. 454 sequencing for microbiota variation. 16s rRNA based abundance of dominant OTU's on the family level. OTU's grouping to each of the 15 most dominant OTU's and all other OTU's combined are color coded. 16S rRNA sequences were analyzed using the RDP pyrosequencing pipeline using standard settings for removal of low quality reads. Samples from subjects H, J, D, and C were analyzed during placebo(P), RM25 (F25) or RM50 (F50) on days 13 (D13) and Day 24(D24). Each column shows the bacterial composition of one fecal sample based on Megablast searches against the Greengenes database. Except for samples H.F25.D24 (H 25 in Figure 2) and H.P. D24 (H 0 in Figure 2) samples differ from those shown in Figure 2.



As the data from microbiota studies with controlled diets suggests, factors other than diet can affect microbiota composition. In some recent reports gut microbiota dynamics within individuals is described as minimal. We need to be careful with attaching subjective labels to microbiota variation and instead thrive to develop objective measures that allow us to compare variation across studies.

2.3. The Need for Standardization of Analysis Methodology

Although NIH is proceeding with an ambiguous Human Microbiota Project (HMP) (<http://nihroadmap.nih.gov/hmp/>) few attempts have to date been made to standardize methodology. There have been some efforts to standardize sample collection and validate DNA extraction protocols [40], but different groups use different primer/probe sets that target different regions of the 16S rRNA gene. Furthermore, large 16S rRNA sequence datasets are analyzed using different bioinformatics approaches and sequences are compared to different databases. In a time in which technology is rapidly developing it is natural that different groups explore various options to identify

the methods and study designs best suited to pursue their specific interests. However, this results in the generation of extensive data sets that cannot easily be pooled to increase the statistical power needed to test the significance of potential associations with human health. Another group that explores vast microbial diversities, the International Census on Marine Microbes (<http://icomm.mbl.edu/>), has taken early steps to facilitate smoother exchange of results. We should standardize protocols and develop an efficient structure that better facilitates coordination of studies and data exchange.

3. Conclusions

We have entered an exciting period of discovering the multiple contributions of host associated microbiota to human health. New sequencing technologies are now facilitating studies at a depth that was previously not possible requiring the development of novel analytical tools for optimal data mining. Recent studies have crucially improved our understanding of the large variation in microbiota composition at various anatomic sites, between individuals and over time. Multiple studies have focused on comparing microbiota composition in diseased individuals with those in healthy controls. To study potential contributions of microbiota to the development of various diseases, and not simply continue to describe how a disease state is associated with changed microbiota, we need to design and perform prospective studies in which we evaluate microbiota in at risk individuals before the disease has developed. In the foreseeable future 16S rRNA based studies will continue to be of high utility for studying microbial diversity. However, an integrated function oriented approach that includes metagenomic, transcriptomic and proteomic studies will be instrumental in developing a better understanding of how complex microbial communities might contribute to host physiology. Such studies will require strong multidisciplinary teams and large cohorts that follow individual over long periods of time until a sufficient number of individual develops the disease(s) of interest. The recent explosion of detailed knowledge of ways in which microbiota can affect health seems to have justified the large additional investment that will be required to perform such studies.

Acknowledgements

A grant from the American Cancer Society (MRS GT CCE-107301) has funded part of the work performed in Mai's lab.

References and Notes

1. Metchnikoff, E. *The Prolongation of Life*; Heinemann: London, UK, 1907; pp.73-84.
2. Dethlefsen, L.; Eckburg, P.B.; Bik, E.M.; Relman, D.A. Assembly of the human intestinal microbiota. *Trends Ecol. Evol.* **2006**, *21*, 517-523.
3. Jones, B.V.; Begley, M.; Hill, C.; Gahan, C.G.; Marchesi, J.R. Functional and comparative metagenomic analysis of bile salt hydrolase activity in the human gut microbiome. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 13580-13585.

4. Lunn, J.C.; Kuhnle, G.; Mai, V.; Frankenfeld, C.; Shuker, D.E.; Glen, R.C.; Goodman, J.M.; Pollock, J.R.; Bingham, S.A. The effect of haem in red and processed meat on the endogenous formation of N-nitroso compounds in the upper gastrointestinal tract. *Carcinogenesis* **2007**, *28*, 685-690.
5. Khachatryan, Z.A.; Ktsoyan, Z.A.; Manukyan, G.P.; Kelly, D.; Ghazaryan, K.A.; Aminov, R.I. Predominant role of host genetics in controlling the composition of gut microbiota. *PLoS ONE* **2008**, *3*, e3064.
6. Baumgart, D.C.; Carding, S.R. Inflammatory bowel disease: cause and immunobiology. *Lancet* **2007**, *369*, 1627-1640.
7. Barrett, J.C.; Hansoul, S.; Nicolae, D.L.; Cho, J.H.; Duerr, R.H.; Rioux, J.D.; Brant, S.R.; Silverberg, M.S.; Jewell, D.; Satsangi, J.; Mathew, C.G.; Parkes, M.; Georges, M.; Daly, M.J. Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat. Genet.* **2008**, *40*, 955-962.
8. Peterson, D.A.; Frank, D.N.; Pace, N.R.; Gordon, J.I. Metagenomic approaches for defining the pathogenesis of inflammatory bowel diseases. *Cell Host. Microbe* **2008**, *3*, 417-427.
9. Sokol, H.; Seksik, P.; Rigottier-Gois, L.; Lay, C.; Lepage, P.; Podglajen, I.; Marteau, P.; Dore, J. Specificities of the fecal microbiota in inflammatory bowel disease. *Inflamm. Bowel. Dis.* **2006**, *12*, 106-111.
10. Swidsinski, A.; Loening-Baucke, V.; Vanechoutte, M.; Doerffel, Y. Active Crohn's disease and ulcerative colitis can be specifically diagnosed and monitored based on the biostructure of the fecal flora. *Inflamm. Bowel. Dis.* **2008**, *14*, 147-161.
11. Mai, V.; Colbert, L.H.; Perkins, S.N.; Schatzkin, A.; Hursting, S.D. Intestinal microbiota: a potential diet-responsive prevention target in ApcMin mice. *Mol. Carcinog.* **2007**, *46*, 42-48.
12. Scanlan, P.D.; Shanahan, F.; Clune, Y.; Collins, J.K.; O'Sullivan, G.C.; O'Riordan, M.; Holmes, E.; Wang, Y.; Marchesi, J.R. Culture-independent analysis of the gut microbiota in colorectal cancer and polyposis. *Environ. Microbiol.* **2008**, *10*, 789-798.
13. Mai, V.; Draganov, P.V. Recent advances and remaining gaps in our knowledge of associations between gut microbiota and human health. *World J. Gastroenterol.* **2009**, *15*, 81-85.
14. Culpepper, T.; Vedam-Mai, V.; Valentine, J.; Polyak, S.; Mai, V. Gut Microbiota composition correlates with Colorectal Polyp Prevalence. In *Proceedings of the 110th General Meeting American Society for Microbiology*; ASM: San Diego, CA, USA, 23-27 Mai 2010.
15. Young, C.; Sharma, R.; Handfield, M.; Mai, V.; Neu, J. Biomarkers for infants at risk for necrotizing enterocolitis: clues to prevention? *Pediatr. Res.* **2009**, *65*, 91R-97R.
16. Mshvildadze, M.; Neu, J.; Shuster, J.; Theriaque, D.; Li, N.; Mai, V. Intestinal microbial ecology in premature infants assessed with non-culture-based techniques. *J. Pediatr.* **2009**, *156*, 20-25.
17. Neu, J.; Mshvildadze, M.; Mai, V. A roadmap for understanding and preventing necrotizing enterocolitis. *Curr. Gastroenterol. Rep.* **2008**, *10*, 450-457.
18. Favier, C.F.; De Vos, W.M.; Akkermans, A.D.L. Development of bacterial and bifidobacterial communities in feces of newborn babies. *Anaerobe* **2003**, *9*, 219-229.
19. Tannock, G.W. Molecular methods for exploring the intestinal ecosystem. *Br. J. Nutr.* **2002**, *87*, S199-S201.

20. Eckburg, P.B.; Bik, E.M.; Bernstein, C.N.; Purdom, E.; Dethlefsen, L.; Sargent, M.; Gill, S.R.; Nelson, K.E.; Relman, D.A. Diversity of the human intestinal microbial flora. *Science* **2005**, *308*, 1635-1638.
21. Zoetendal, E.G.; Vaughan, E.E.; De Vos, W.M. A microbial world within us. *Mol. Microbiol.* **2006**, *59*, 1639-1650.
22. Frank, D.N.; Pace, N.R. Gastrointestinal microbiology enters the metagenomics era. *Curr. Opin. Gastroenterol.* **2008**, *24*, 4-10.
23. Costello, E.K.; Lauber, C.L.; Hamady, M.; Fierer, N.; Gordon, J.I.; Knight, R. Bacterial community variation in human body habitats across space and time. *Science* **2009**, *326*, 1694-1697.
24. Lozupone, C.; Hamady, M.; Knight, R. UniFrac—an online tool for comparing microbial community diversity in a phylogenetic context. *BMC Bioinformatics* **2006**, *7*, 371.
25. Hamady, M.; Lozupone, C.; Knight, R. Fast UniFrac: facilitating high-throughput phylogenetic analyses of microbial communities including analysis of pyrosequencing and PhyloChip data. *ISME J.* **2010**, *4*, 17-27.
26. Turnbaugh, P.J.; Hamady, M.; Yatsunencko, T.; Cantarel, B.L.; Duncan, A.; Ley, R.E.; Sogin, M.L.; Jones, W.J.; Roe, B.A.; Affourtit, J.P.; Egholm, M.; Henrissat, B.; Heath, A.C.; Knight, R.; Gordon, J.I. A core gut microbiome in obese and lean twins. *Nature* **2009**, *457*, 480-484.
27. Sun, Y.; Cai, Y.; Liu, L.; Yu, F.; Farrell, M.L.; McKendree, W.; Farmerie, W. ESPRIT: estimating species richness using large collections of 16S rRNA pyrosequences. *Nucleic Acids Res.* **2009**, *37*, e76.
28. Gill, S.R.; Pop, M.; Deboy, R.T.; Eckburg, P.B.; Turnbaugh, P.J.; Samuel, B.S.; Gordon, J.I.; Relman, D.A.; Fraser-Liggett, C.M.; Nelson, K.E. Metagenomic analysis of the human distal gut microbiome. *Science* **2006**, *312*, 1355-1359.
29. Qin, J.; Li, R.; Raes, J.; Arumugam, M.; Burgdorf, K.S.; Manichanh, C.; Parkhill, J.; Weissenbach, J.; Bork, P.; Ehrlich, S.D.; Wang, J. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **2010**, *464*, 59-65.
30. Martin, F.P.; Wang, Y.; Sprenger, N.; Yap, I.K.; Lundstedt, T.; Lek, P.; Rezzi, S.; Ramadan, Z.; van Bladeren, P.; Fay, L.B.; Kochhar, S.; Lindon, J.C.; Holmes, E.; Nicholson, J.K. Probiotic modulation of symbiotic gut microbial-host metabolic interactions in a humanized microbiome mouse model. *Mol. Syst. Biol.* **2008**, *4*, 157.
31. Martin, F.P.; Sprenger, N.; Yap, I.K.; Wang, Y.; Bibiloni, R.; Rochat, F.; Rezzi, S.; Cherbut, C.; Kochhar, S.; Lindon, J.C.; Holmes, E.; Nicholson, J.K. Panorganismal gut microbiome-host metabolic crosstalk. *J. Proteome. Res.* **2009**, *8*, 2090-2105.
32. Blaut, M.; Collins, M.D.; Welling, G.W.; Dore, J.; van Loo, J.; de Vos, W. Molecular biological methods for studying the gut microbiota: the EU human gut flora project. *Br. J. Nutr.* **2002**, *87*, S203-S211.
33. Zoetendal, E.G.; von Wright, A.; Vilpponen-Salmela, T.; Ben Amor, K.; Akkermans, A.D.; de Vos, W.M. Mucosa-associated bacteria in the human gastrointestinal tract are uniformly distributed along the colon and differ from the community recovered from feces. *Appl Environ. Microbiol.* **2002**, *68*, 3401-3407.

34. Croucher, L.J.; Bury, J.P.; Williams, E.A.; Riley, S.A.; Corfe, B.M. Commonly used bowel preparations have significant and different effects upon cell proliferation in the colon: a pilot study. *BMC Gastroenterol.* **2008**, *8*, 54.
35. Bucher, P.; Gervaz, P.; Egger, J.F.; Soravia, C.; Morel, P. Morphologic alterations associated with mechanical bowel preparation before elective colorectal surgery: a randomized trial. *Dis. Colon. Rectum* **2006**, *49*, 109-112.
36. Swidsinski, A.; Loening-Baucke, V.; Verstraelen, H.; Osowska, S.; Doerffel, Y. Biostructure of fecal microbiota in healthy subjects and patients with chronic idiopathic diarrhea. *Gastroenterology* **2008**, *135*, 568-579.
37. Salzman, N.H.; Hung, K.; Haribhai, D.; Chu, H.; Karlsson-Sjoberg, J.; Amir, E.; Tegatz, P.; Barman, M.; Hayward, M.; Eastwood, D.; Stoel, M.; Zhou, Y.; Sodergren, E.; Weinstock, G.M.; Bevins, C.L.; Williams, C.B.; Bos, N.A. Enteric defensins are essential regulators of intestinal microbial ecology. *Nat. Immunol.* **2009**, *11*, 76-82.
38. Turnbaugh, P.J.; Ridaura, V.K.; Faith, J.J.; Rey, F.E.; Knight, R.; Gordon, J.I. The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic mice. *Sci. Transl. Med* **2009**, *1*, 6ra14.
39. Mai, V.; Katki, H.; Clevidence, B.; Hursting, S.; Harmsen, H.; Schatzkin, A. Changes in the fecal flora composition of human volunteers in a double-blind randomized black tea feeding study. *J. Nutr.* **2002**, *132*, 3551S.
40. Salonen, A.; Nikkila, J.; Jalanka-Tuovinen, J.; Immonen, O.; Rajilic-Stojanovic, M.; Kekkonen, R.A.; Palva, A.; De Vos, W.M. Comparative analysis of fecal DNA extraction methods with phylogenetic microarray: Effective recovery of bacterial and archaeal DNA using mechanical cell lysis. *J. Microbiol. Methods* **2010**, *81*, 127-134.

© 2010 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).