N-of-1

Tracking Human Gut Microbiome Changes Resulting from a Colonoscopy

Larry Smarr[1,2,5]; Embriette R. Hyde[4]; Daniel McDonald[4]; William J. Sandborn[2,3]; Rob Knight[2,4,5]

[1]University of California San Diego, Calit2, La Jolla, CA, USA; [2]University of California San Diego, Center for Microbiome Innovation, La Jolla, CA, USA; [3]University of California San Diego, Gastroenterology, San Diego, CA, USA; [4]University of California San Diego, Pediatrics, La Jolla, CA, USA; [5]University of California San Diego, Computer Science and Engineering, La Jolla, CA USA

1. Introduction

The human gut microbiome is known to vary on daily time scales and have major ecological shifts in disease states. Pharmaceutical interventions have also been observed to cause sudden changes in microbial composition. Here, I (Larry Smarr) quantify the daily shifts in my gut microbiome composition after a bowel cleanse and colonoscopy by sequencing the V4 region of the bacterial 16S rRNA gene of frozen stool samples which I collected and stored. My team's results reproduce the known result that there is an immediate change in the microbiome ecology after the colonic intervention with recovery after several weeks, but we discover new dynamical insights on the daily time scale. In particular, we find that 1) alpha-diversity decreases dramatically post-colonoscopy, followed by a slow recovery; 2) there is a dramatic evolution in the microbial ecology post-cleanse until a return to the pre-colonoscopy state after two weeks; and 3) some microbial taxa change in relative abundance by up to 40-fold. This N=1 experiment indicates that the gut microbiome can have major changes on a daily basis when the host system has a major insult to its environment.

Correspondence to:
Larry Smarr
University of California San Diego
9500 Gilman Drive #0436
La Jolla, CA 92093–0436
USA
E-mail: lsmarr@ucsd.edu

Methods Inf Med 2017; 56: 442–447
https://doi.org/10.3414/ME17-01-0036
received: April 7, 2017
accepted: September 8, 2017

2. What Did I Do (and Why)?

I have been engaged in a Quantified Self project [1] for the last decade, carrying out intensive monitoring of my body's biomarkers in blood and stool, human and microbiome genomics, and 3D imaging. I am able to do this because I am a UC San Diego Professor with access to computational, sequencing, and visualization facilities that are not normally available to an individual. One result [2, 3] of this undertaking was a self-diagnosis of Inflammatory Bowel Disease (IBD), a surprise at the time given that I had no overt symptoms and felt fine. Due to the growing interest in the role that the gut microbiome may play in the pathogenesis and evolution of this disease state, I began collecting and storing my stool in late 2011. Since then I have collaborated with the J. Craig Venter Institute and the UC San Diego Center for Microbiome Innovation (CMI) to analyze my gut microbiome, both by comparing my gut microbiome to that of healthy individuals and individuals with IBD, in addition to developing a long-term time series analysis of my microbial ecology.

The gut microbiome is comprised of bacterial species and strains that taxonomically group into a number of different phyla. In my earlier work my research group carried out deep metagenomic analysis [4, 5] of 255 healthy subjects in the NIH Human Microbiome Project, as well as a year and half time series on my stool. In the top graph in ▶Figure 1, one can see that in healthy people, two phyla, the Bacteroides and Firmicutes dominate the gut, a result which was also published by the Human Microbiome Consortium in 2012 [6, 7].

I collected my first sample on December 28, 2011, when my blood serum complex reactive protein (CRP) was 27 times the upper limit for healthy. As can be seen in ▶Figure 1, the abundance of Phylum Bacteroidetes (red) in my sample was 15-fold reduced in abundance compared to the Healthy average (HMP) in the top graph. Additionally, a number of normally rare species in the Firmicutes dominated that phylum, while Proteobacteria were also elevated (*E. coli*, a species normally present at a relative abundance of a few tenths of a percent in healthy individuals was at a relative abundance of 10% in my sample). Phyla Euryarchaeota (in the Archaea Kingdom) and Fusobacteria were also greatly elevated (520x and 150x, respectively) compared to the healthy average. Together, these data clearly demonstrate that in my disease state, the microbial ecology in my gut was far from the healthy population.

It also became apparent, when analyzing samples collected at intervals of one to several months after that initial inflamed state, that my gut microbiome was very unstable over time (▶Figure 2), with rapid collapse and blooming of specific species across time. This is radically different from microbiome trajectory of healthy individuals, which is quite stable over time-though events such as travel, infection, or dietary changes can rapidly and drastically affect the composition of the gut microbiome [8, 9].

© Schattauer 2017

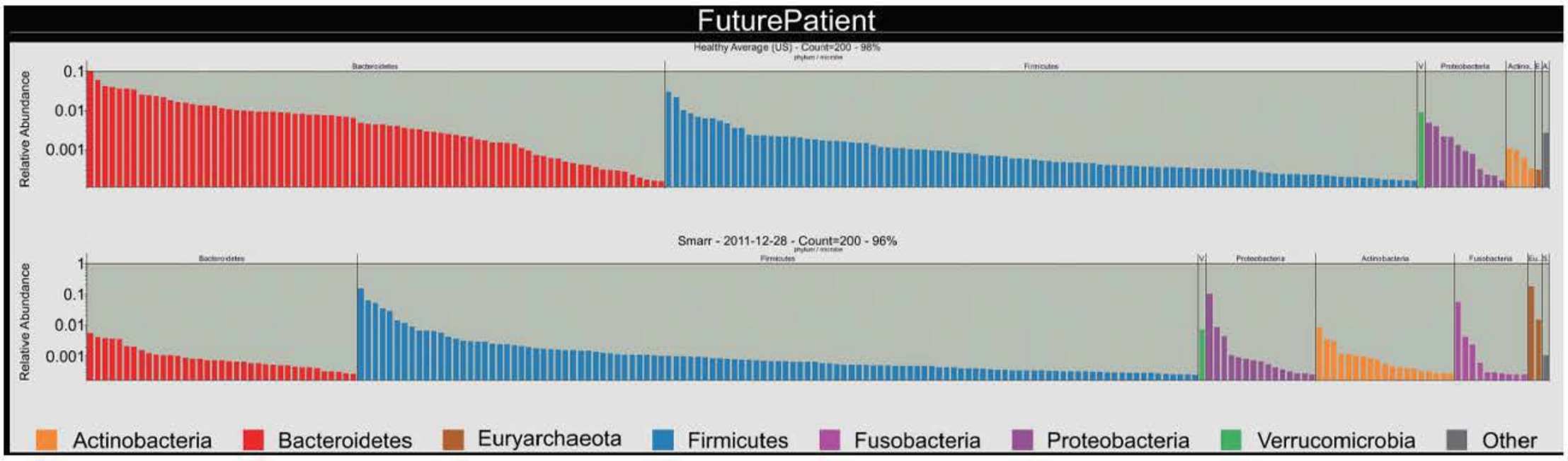


Figure 1 The quantitative gut microbiome ecology of individuals screened by the Human Microbiome Project as healthy (top graph) and my ecology at my point of highest inflammation (lower graph), as sequenced and computed in [4]. Each bar graph displays the relative abundance (on a vertical logarithmic scale) of the 200 most abundant species in descending order of relative abundance. The species are organized into phyla (each phyla colored differently). Here, an exponential falloff in relative abundance (straight line on semi-log scale) is apparent when moving to lower levels of relative abundance. This indicates that a few species dominate each phylum, with a long tail of increasingly rare species. Graphical software developed by Andrew Prudhomme, Calit2's Qualcomm Institute, UC San Diego.

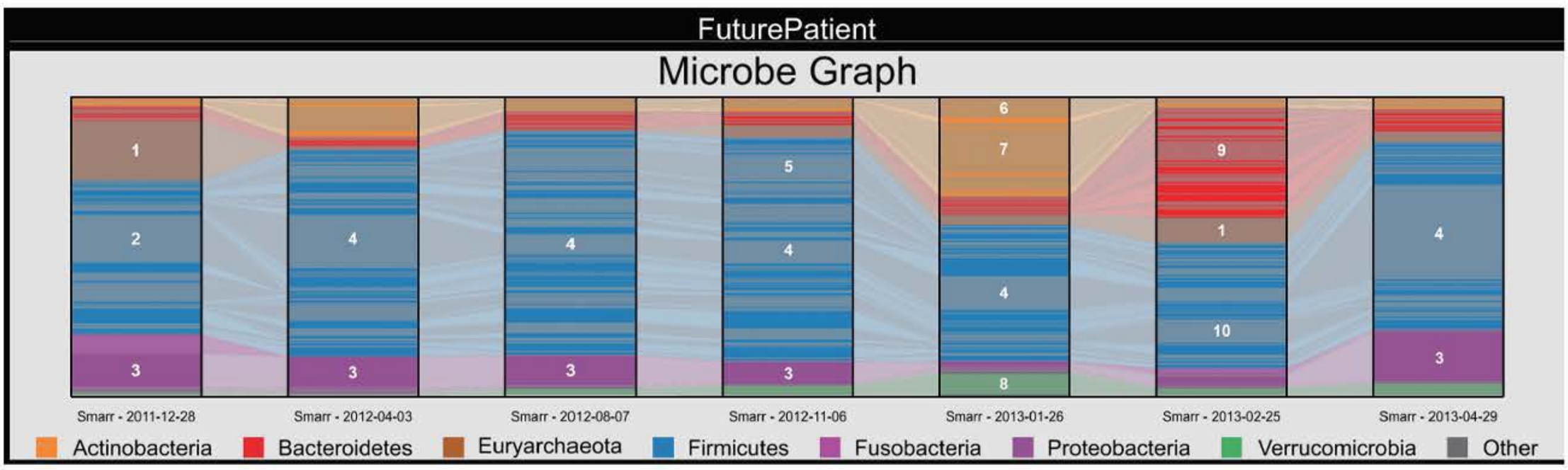


Figure 2 **My dynamic microbiome ecology over a year and half.** Each bar is a stacked bar graph of the relative abundance of species organized by phylum, with each bar representing an individual sample (ordered by collection date) The change in the relative abundance of species across time can be tracked through the shaded areas between bars. Species (phylum) at a relative abundance of 5% or more are numbered and include: 1. *Methanobrevibacter smithii* (Euryarchaeota), 2. *Parvimonas micra* (Firmicutes), 3. *Escherichia coli* (Proteobacteria), 4. *Faecalibacterium prausnitzii* (Firmicutes), 5. *Dorea longicatena* (Firmicutes), 6. *Actinomyces odontolyticus* (Actinobacteria), 7. *Bifidobacterium animalis* (Actinobacteria), 8. *Akkermansia muciniphila* (Verrucomicrobia), 9. *Bacteroides intestinalis* (Bacteriodetes), and 10. *Ruminococcus bromii* (Firmicutes).

Knowing that keeping my gut microbiome in a healthy state was essential to my health as the microbiome host, I became interested in the microbiome dynamics on short time scales and the impact of external influences. My chance to study this came during the summer of 2015, when I was 67 years old and I was scheduled for an annual colonoscopy performed by William Sandborn – a procedure requiring a colon cleanse with a polyethylene glycol solution (PEG). On the five days leading up to July 6, 2015, when I did the colon cleanse, I obtained and froze stool samples each day. Late in the day on July 6 I drank consecutively two single-liter doses of MoviPrep, a commonly used PEG solution. On July 7 I had the colonoscopy, during which a large number of biopsies (~50) were taken. Starting the next day (July 8) I took daily stool samples until July 31. This gave me a long daily time series of stool samples from which I would later be able to sequence the microbiome to analyze in detail the changes in my gut microbiome after a major insult to their host environment.

I hypothesized, based on the research literature [10, 11, 12] that there would be a major transitory impact on my gut microbiome (though some of these papers have reported inconsistent results). The current literature reports several approaches to assessing gut microbiome alterations post-colonoscopy; however, those using next-generation sequencing techniques do not have daily stool samples. One of the most recent studies (published in 2015) is also the largest, to my knowledge: stool was collected from 23 healthy subjects undergoing a colon cleanse immediately prior and immediately after the cleanse, and then again at 14 and 28 days post-cleanse [12]. This study used the 16S rRNA gene V1 and V6 hypervariable regions, whereas our study uses V4, leading to different methods for

© Schattauer 2017

assaying microbial membership. Differences in target variable region may especially alter taxonomic results due to bias associated with different 16S rRNA variable regions. One valuable contribution of [12] is that they performed qPCR and were able to show that the bacteria content of the stool was decreased 35x immediately after the cleanse, indicating that the impact of the cleanse on the microbiome state is quite substantial.

In contrast to this study, my N of 1 approach was characterized by daily sampling for six days prior to the colonoscopy and 24 days after the colonoscopy, providing a level of time-resolution that doesn't exist in any other colon-cleanse microbiome study. Additionally, our study makes use of the most recent techniques developed and benchmarked for high quality microbiome studies. In spite of these differences in detail, the overall result reported in [12] is reproduced here, namely that the microbiome ecology substantially recovers to the pre-colonoscopy state after several weeks.

Through this sampling effort, I hoped to understand in detail the impact the procedure can have on microbiome composition, and how long it would take to recover, as well as the daily trajectory of that recovery. By partnering with the CMI, my samples underwent state-of-the-art processing and analysis techniques, the results of which I present here.

3. How Did I Do It?

My colonoscopy showed narrowing of the sigmoid colon with the lumenal diameter reduced down to 9 mm, associated with an inflammatory mass, and diverticulosis. My pathology of the sigmoid colon biopsies showed focal mildly active colitis. I collected a small bit of fecal material daily from July 1, 2015 through July 31, 2015 using a sterile BD BBL CultureSwab (Becton, Dickinson, and Company, catalogue number 220135) broken off into a 1.5mL Eppendorf tube. These tubes were stored in an ice pack-insulated box inside my home freezer (▶Figure 3) to preserve the sample and prevent damage from the automatic freeze-thaw cycles that are common in home freezers. Samples were then delivered



Figure 3 Sample collection and home storage. My stool samples are collected, labeled, and immediately frozen in my home freezer. Note that the bar coded 1.5ml vials are surrounded by frozen ice packs to thermally insulate the vials from any freeze-thaw cycles that are common in home refrigerators.

frozen to the Knight Lab at the University of California San Diego (UC San Diego), where they were immediately stored at –80°C until sample processing.

DNA was extracted using the MOBIO PowerSoil DNA Extraction kit (MOBIO, Carlsbad, CA) following protocols benchmarked through the Earth Microbiome Project (www.earthmicrobiome.org/protocols-and-standards/dna-extraction-protocol/). The V4 of the 16S rRNA gene was then amplified with PCR again following protocols benchmarked through the Earth Microbiome Project ([13]; see also www.earthmicrobiome.org/protocols-and-standards/16s/). This approach facilitates the capture of most of the bacteria present in a sample due to the ubiquitous nature of the 16S rRNA gene and the "universal" V4 primer. The V4 region recapitulates the full length 16S rRNA gene at approximately genus level resolution. Barcoded amplicons were then sequenced on an Illumina HiSeq 2500 using the 2x 150bp protocol at the Institute for Genomic Medicine at UC San Diego.

The resulting sequences were analyzed by Embriette Hyde and Daniel McDonald in the Knight Lab. Raw sequences were demultiplexed using QIIME 1.9.1 [14] following the default parameters for split_libraries_fastq.py, with the exception of the minimum PHRED quality score, which was set to 19 instead of the default of 3. We changed this default as this is the quality score recommended for the downstream analysis using Deblur (http://msystems.asm.org/content/2/2/e00191–16).

While historically microbiome studies have involved clustering sequences into "operational taxonomic units (OTUs)", Deblur and other similar tools have recently emerged as an improved way to infer the most likely set of input sequences in the context of a non-trivial error rate during sequencing, and to obtain maximal resolution within the data (i.e., differentiating by a single nucleotide). This improves resolution over traditional OTU picking methods, which often use an arbitrary sequence similarity of 97% for clustering. After quality filtering, trimming, and deblurring the data, some samples had fewer than 1000 reads remaining and were eliminated from downstream analyses.

Because some downstream analysis steps, including the computation of the PD Whole Tree diversity metric (alpha diversity), requires a phylogeny tree, we then used SEPP ([15]; see also https://github.com/smirarab/sepp/wiki/SEPP-on-Greengenes) to insert the V4 fragment sequences resulting from Deblur into the Greengenes tree [16]. Taxonomy was assigned to the sequences using the taxonomy present on the internal nodes of the already annotated Greengenes tree, and this taxonomy information was added to the biom-format observation matrix containing unique sequences and their abundances in each sample (traditionally called an "OTU table" in the microbiome research field). Using QIIME 1.9.1, an alpha diversity metric (PD Whole Tree) and the unweighted UniFrac distance metric were computed on rarefied BIOM [17] tables, and taxa summaries were obtained from these tables in order to obtain a clear picture of community changes over time. For some of these analyses, replicate samples collected on the same day were merged so that each date had a single representative sample associated with it.

© Schattauer 2017

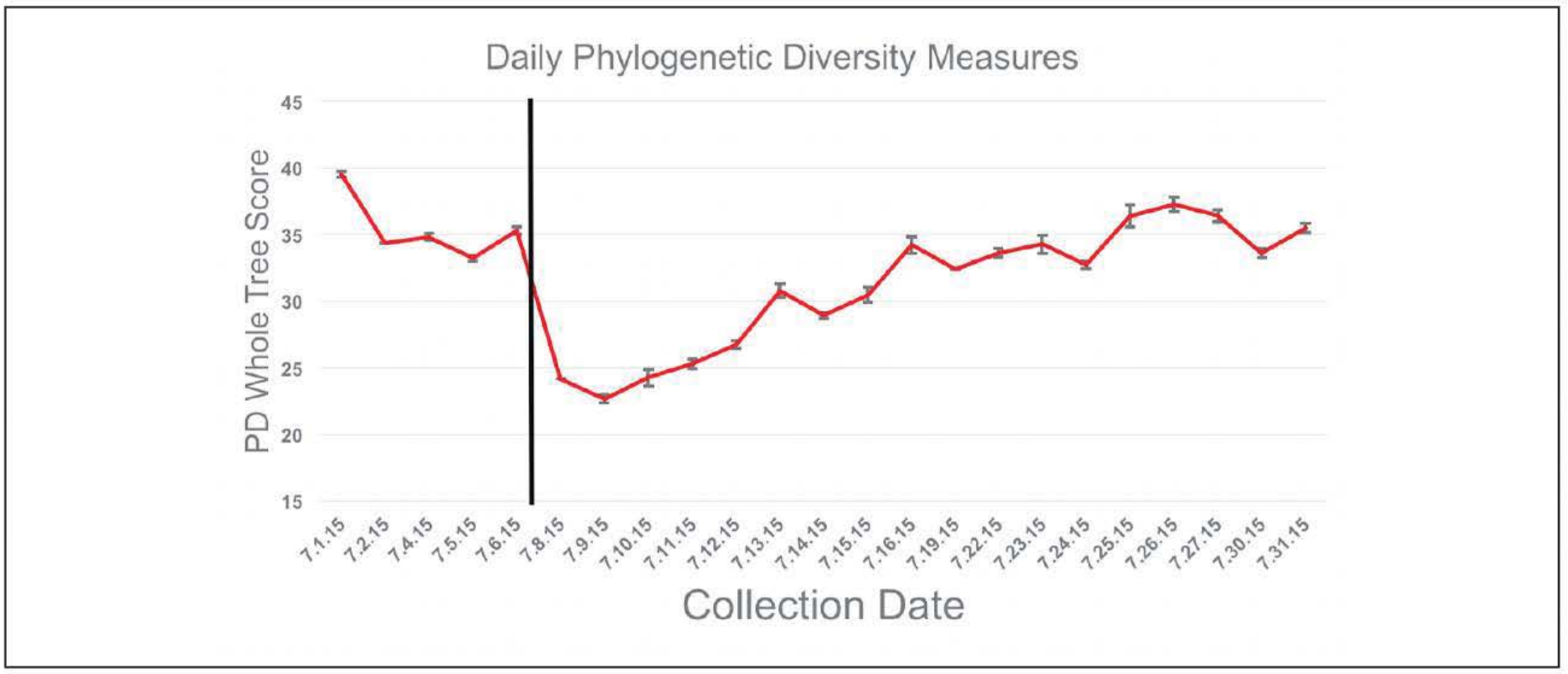


Figure 4 Alpha diversity is decreased after a colon cleanse and colonoscopy, but recovers in less than two weeks. The phylogenetic diversity (alpha diversity) score is plotted for each date on which a sample was collected. Phylogenetic diversity drops abruptly after the colonoscopy (indicated by black vertical line) and slowly increases again. Data are the mean +/– standard deviation of the PD Whole Tree metric calculation on ten random sub-samplings of the OTU table to a rarefaction depth of 25510 sequences per sample.

4. What Did I Learn?

The main question I wanted to answer was exactly how my microbiome ecology was disrupted by the colon cleanse and colonoscopy. Our team approached the answer at three levels, first by investigating the change in microbiome diversity over time, second by exploring the evolution of the microbiome ecology as a whole, and third by following the change in relative abundance of specific microbial taxa.

Using 16S rRNA sequencing revealed that the microbial community *diversity* was altered immediately after the colon cleanse and colonoscopy. As seen in ▶Figure 4, phylogenetic diversity (PD Whole Tree), a measure of the diversity within a sample (alpha diversity) that takes into account evolutionary relationships, dropped significantly (Monte Carlo permuted p-value = 0.009) from an average of 34–39 in the days prior to the colonoscopy to an average of 22–24 in the two days immediately following the colonoscopy (July 7, 2015, indicated by the vertical black line in ▶Figure 4). The lowest average diversity was noted two days post-colonoscopy and slowly increased over the following days. By July 16th, the average phylogenetic diversity was once again in the range of 34 and significantly (Monte Carlo permuted p-value = 0.003) higher than the average phylogenetic diversity of samples collected in the first 9 days post-colonoscopy, indicating a near-complete recovery of diversity. This is the most striking difference between my work and prior published work assessing microbiome changes associated with colon cleanses and colonoscopies. For instance, in [12] it is noted that in that and previous studies: "… we found no statistical differences in the diversity of the microbiota between the time points or treatment groups in agreement with prior studies addressing the impact of lavage. [12]".

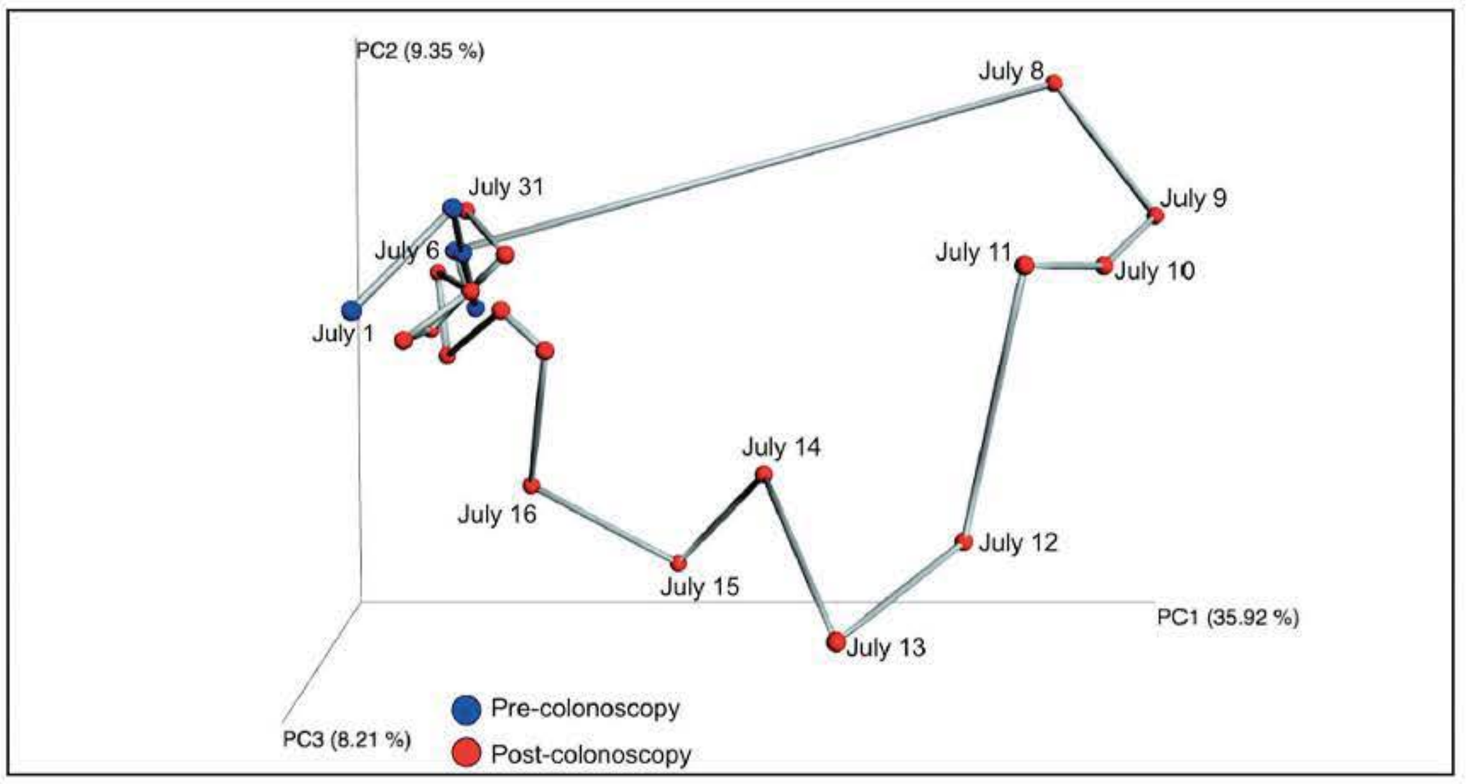


Figure 5 An unweighted UniFrac-based PCoA plot reveals an immediate and drastic microbial community change after the colonoscopy, followed by a period of recovery until after July 16, 2015, when post-colonoscopy samples clusters again with pre-colonoscopy samples. The gray line traces the change in microbial community structure from day-to-day by connecting samples in sequential order. Blue dots=pre-colonoscopy; red dots=post-colonoscopy. Note how close the beginning date (July 1) and the final date (July 31) are to each other, showing the microbial ecology was able to recover to its pre-colonoscopy state. Data shown here have been rarefied to 25951 sequences per sample.

© Schattauer 2017

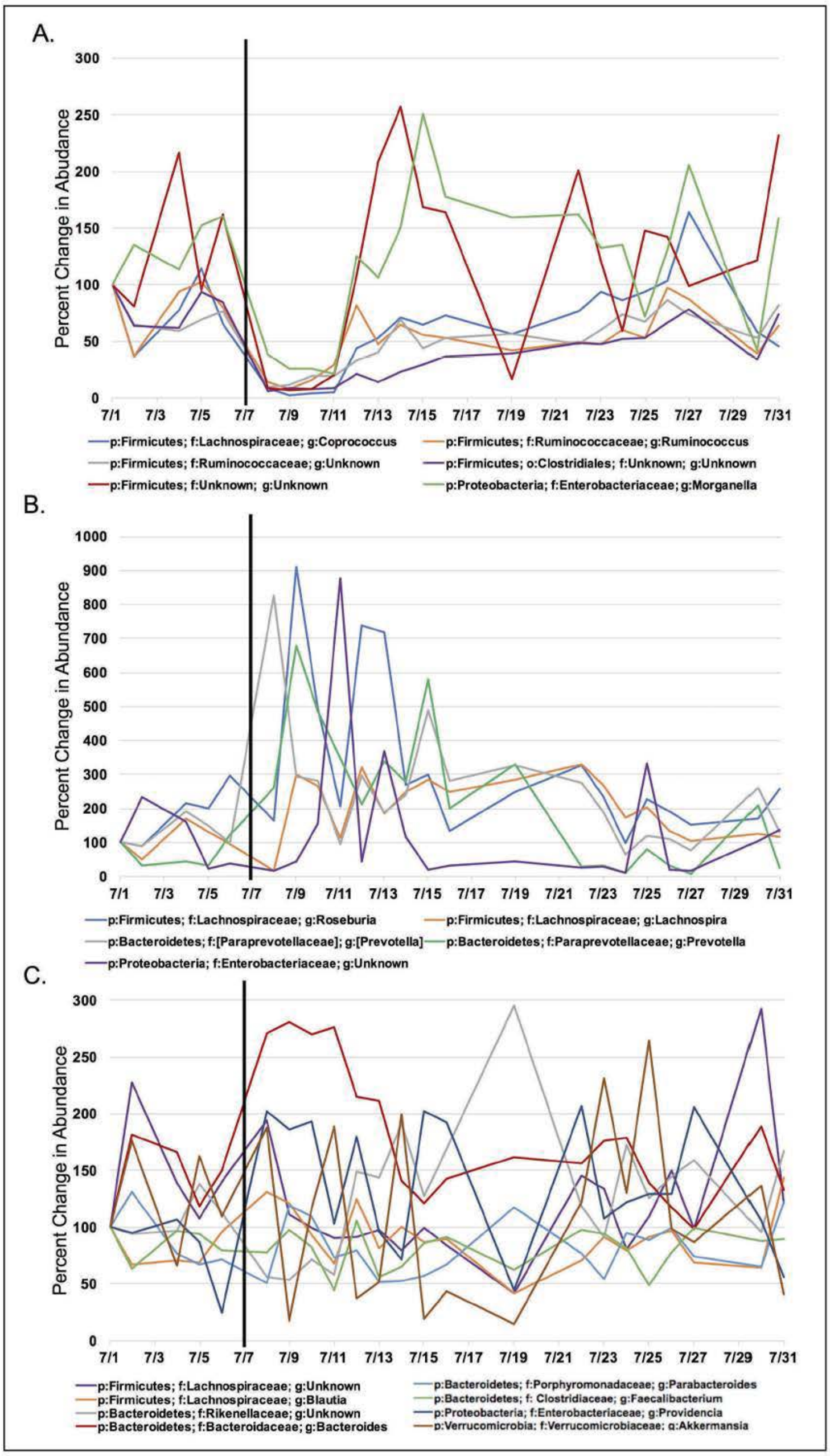


Figure 6 Percent change in abundance of select genera after colon cleanse and colonoscopy. Bacterial genera which increased more than 3x **(A)**, decreased more than 3x **(B)**, or remained roughly stable after the colon cleanse and colonoscopy (vertical black bar on July 7) are plotted as daily percent change compared to the first sampling date (July 1). The relative abundance of each genus at each time point was divided by its relative abundance on July 1 and then plotted as a percent change. Some genera were unable to be classified to the genus level; the phylum, family, and deepest level of classification is listed. All dates are in 2015.

Next, we examined how the microbial communities shifted through time by computing pairwise sample distances using UniFrac and then projecting those distances into principal coordinates analysis (PCoA) space. In the dramatic graph in ▶Figure 5, all pre-colonoscopy samples (blue dots) cluster together, followed by an immediate large shift to a different location in PCoA space the day after the colonoscopy (July 8). The ecology then enters a "wandering stage" with a gradual movement back to the pre-colonoscopy cluster by July 19, 2015 (all post-colonoscopy samples are colored red). This large shift in PCoA space indicates that the drop in diversity we observe comes from a major shift in the underlying taxa forming the ecology.

Notably, all samples collected before July 8 and from July 19 and later form a tight cluster, while those samples collected in the 9 days following the colonoscopy are characterized by large movements through PCoA space, indicating significant bacterial community changes during the "recovery" days followed by a relative stability of the community within nine days after the colonoscopy. Indeed, assessing differences in unweighted UniFrac distances revealed that there was no significant difference between the mean distances of pre- vs. post-colonoscopy and post-colonoscopy vs. recovered samples. This data further supports that the microbial communities in samples collected during the nine days immediately following the colonoscopy were significantly different from the microbial communities in samples collected prior to the colonoscopy and in samples categorized as "recovered", which did not significantly differ from each other.

Next, we examined changes in the relative abundance of genera across time. In this discussion, we focus on the most abundant genus, those with relative abundance of 1% or higher. In [12] it was noted that changes in taxon relative abundance up to a factor of 3x were observed post-colonoscopy. In contrast, in our work with much finer time resolution (▶Figures 6A and 6B) we found a number of genera that exhibited much larger changes in relative abundance. For instance, genus *Rosburia* jumped over 9-fold post-colonoscopy com-

© Schattauer 2017

pared to its July 1 value and genus Coprococcus dropped by 40-fold post-colonoscopy compared to its July 1 value. Additionally, instead of just decreasing or increasing and then coming back to the prior state, several taxa underwent large fluctuations above and below the pre-colonoscopy state. For example, *Roseburia* increased immediately post-colonoscopy, returned to pre-colonoscopy levels, increased again, and then finally remained slightly higher than pre-colonoscopy levels. While average relative abundance levels experienced nearly full recovery after a few weeks, day to day variations (some wide) indicated that, at least for some taxa, two weeks was not sufficient for a full recovery from the colon cleanse and colonoscopy.

Despite the large fluctuations illustrated in ▶ Figures 6A and 6B, some of the major members of the gut community prior to the colonoscopy remained major members after the colonoscopy. In ▶ Figure 6C we show a number of abundant genus that had excursions less than 3x. We find that the genus *Bacteroides, Morganella,* and *Faecalibacterium* were among the top five genera detected at all time points, both before and after the colonoscopy. While some of these genera (e.g., *Bacteroides*) had variations up to 3x, others (e.g. *Faecalibacterium and Parabacteroides)* were seen to be quite stable over the entire time series.

5. Discussion

This experiment demonstrates the clear value of using an N of 1 time series to analyze in detail the time evolution of the human gut microbiome driven by external interventions. While these trends need to be confirmed through studies with higher Ns, the heterogeneity in the subject population in a typical averaged population study will often wash out the finer detail that was resolved here. This N of 1 study assessing gut microbiome changes in response to a colon cleanse and colonoscopy is contrasted to previous studies, which average over dozens of patients, such as [12], and have coarse time resolution. During my experiment, I sampled daily before and for three weeks after the colon cleanse. My results show that longitudinal analysis within individuals allows for detection of details that would be occluded by studies which average across multiple subjects, thereby providing guidelines for future studies.

We see now in high resolution, both in time and in taxa, that the colon cleanse and colonoscopy is quite disruptive to the gut microbiome, although that damage mostly goes away, as seen in earlier studies, within two weeks. The results presented here confirm just how dynamic an individual's gut microbiome can be. The fact that while many taxa keep their relative abundance, while others change up to 40x, confirms that we are witnessing a transient redistribution of the population. Given there is such major disruption in the microbial ecology from bowel preparation for colonoscopy, it is likely that there are corresponding transient changes in the microbial metabolic pathways available to the gut. More work will be needed to derive the metabolic pathways from the taxonomy, but the data presented here support an imperative for such investigations.

Acknowledgment

We would like to thank members of the Knight Lab and Calit2's Qualcomm Institute for helpful discussions on this project.

References

1. Smarr L. Quantifying your body: a how-to guide from a systems biology perspective. Biotechnol J 2012; 7(8): 980–991.
2. Smarr L. Quantified Health: A 10-year Detective Story of Digitally Enabled Genomic Medicine. The Strategic News Service Newsletter, Special Letter 2011; 14(36). Available from: https://www.stratnews.com/recent/mode/show/issue/2011-09-29/.
3. Smarr L. Can You Coordinate the Dance of Your Body's 100 Trillion Microorganisms? TEDMED 2013. Available from: http://www.tedmed.com/talks/show?id=18018.
4. Wu S, et al. Large memory high performance computing enables comparison across human gut microbiome of patients with autoimmune diseases and healthy subjects. Proceedings of the Conference on Extreme Science and Engineering Discovery Environment: Gateway to Discovery (XSEDE '13). Article 25, 2013.
5. Yazdani M, et al. Using Machine Learning to Identify Major Shifts in Human Gut Microbiome Protein Family Abundance in Disease. IEEE International Conference on Big Data, 2016. p. 1272–1280.
6. Huttenhower C, et al. Structure, function and diversity of the healthy human microbiome. Nature 2012; 486(7402): 207–214.
7. Human Microbiome Project Consortium. A framework for human microbiome research. Nature 2012; 486(7402): 215–221.
8. Caporaso JG, et al. Moving pictures of the human microbiome. Genome Biol 2011; 12(5): R50.
9. David LA, et al. Host lifestyle affects human microbiota on daily timescales. Genome Biol 2014; 15: R89.
10. Drago L, Toscano M, De Grandi R, Casini V, Pace F. Persisting changes of intestinal microbiota after bowel lavage and colonoscopy. Eur J Gastroenterol Hepatol 2016; 28(5): 532–537.
11. O'Brien CL, Allison GE, Grimpen F, Pavli P. Impact of colonoscopy bowel preparation on intestinal microbiota. PLoS One 2013; 8(5): e62815.
12. Jalanka J, et al. Effects of bowel cleansing on the intestinal microbiota. Gut 2015; 64(10): 1562–1568.
13. Caporaso JG, et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J 2012; 6(8): 1621–1624.
14. Caporaso JG, et al. QIIME allows analysis of high-throughput community sequencing data. Nat Methods 2010; 7(5): 335–336.
15. Mirarab S, Nguyen N, Warnow T. SEPP: SATé-enabled phylogenetic placement. Pac Symp Biocomput 2012: 247–258.
16. McDonald D, et al. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. ISME J 2012; 6(3): 610–618.
17. McDonald D, et al. The Biological Observation Matrix (BIOM) format or: how I learned to stop worrying and love the ome-ome. Gigascience 2012; 1(1): 7.

© Schattauer 2017