Title: Temporal Variability is a Personalized Feature of the Human Microbiome

Authors: Gilberto E. Flores[a, *, #], J. Gregory Caporaso[b, c, d, #], Jessica B. Henley[e], Jai Ram Rideout[d, f], Daniel Domogala[b], John Chase[b], Jonathan W. Leff[e, g], Yoshiki Vázquez-Baeza[h], Antonio Gonzalez[i], Rob Knight[i, j, k], Robert R. Dunn[l], Noah Fierer[e, g, *].

Affiliations:

[a]Department of Biology, California State University, Northridge, Northridge, CA 91330-8303.

[b]Department of Biological Sciences, Northern Arizona University, Flagstaff, AZ, 86011.

[c]Institute for Genomics and Systems Biology, Argonne National Laboratory, Argonne, IL, 60439.

[d]Center for Microbial Genetics and Genomics, Northern Arizona University, Flagstaff, AZ 86011.

[e]Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO 80309.

[f]Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY 10029.

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.509v1 | CC-BY 4.0 Open Access | rec: 22 Sep 2014, publ: 22 Sep 2014

[g]Department of Ecology and Evolutionary Biology, University of Colorado, Boulder, CO 80309.

[h]Department of Computer Science, University of Colorado, Boulder, CO 80309.

[i]BioFrontiers Institute, University of Colorado, Boulder, CO 80309.

[j]Department of Chemistry and Biochemistry University of Colorado, Boulder, CO 80309.

[k]Howard Hughes Medical Institute, University of Colorado, Boulder, CO 80309.

[l]Department of Biology and Keck Center for Behavioral Biology, North Carolina State University, Raleigh, NC, 27607

Email addresses: GEF – gilberto.flores@csun.edu; JGC – gregcaporaso@gmail.com; JBH – jessicabhenley@gmail.com; JRR – jai.rideout@gmail.com; DD – ddd38@nau.edu; JC – chase.john@gmail.com; JWL – leff.jonathan@gmail.com; YVB – yoshiki89@gmail.com; AG – antgonza@gmail.com; RK – rob.knight@colorado.edu; RRD – rroberdeaudunn@gmail.com; NF – noahfierer@gmail.com.

***Corresponding authors:**

Gilberto E. Flores

18111 Nordhoff Street

Northridge, CA, 91330-8303

818-677-4276

gilberto.flores@csun.edu

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.509v1 | CC-BY 4.0 Open Access | rec: 22 Sep 2014, publ: 22 Sep 2014

Noah Fierer

University of Colorado, Boulder

216 UCB

Boulder, CO, 80309-0216

303-492-5615

noahfierer@gmail.com

[#]**These authors contributed equally to this work.**

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.509v1 | CC-BY 4.0 Open Access | rec: 22 Sep 2014, publ: 22 Sep 2014

Abstract:

Background: It is now apparent that the complex microbial communities found on and in the human body (the human microbiome) vary across individuals. What has largely been missing from previous studies is an understanding of how these communities vary over time within individuals. To the extent to which it has been considered, it is often assumed that temporal variability is negligible for healthy adults. Here we address this gap in understanding by profiling the forehead, gut (fecal), palm, and tongue microbial communities in 85 adults, weekly over three months.

Results: We found that skin (forehead and palm) varied most in the number of taxa present, whereas gut and tongue communities varied more in the relative abundances of taxa. Within each body habitat, there was a wide range of temporal variability across the study population, with some individuals consistently harboring more variable communities than others. The best predictor of these differences in variability across individuals was microbial diversity; individuals with more diverse gut or tongue communities were less variable than individuals with less diverse communities.

Conclusions: This expanded sampling allowed us to observe consistently high levels of temporal variability in both diversity and community structure in all body habitats studied. These findings suggest that temporal dynamics may need to be considered when attempting to link changes in microbiome structure to changes in health status. Furthermore, our findings show that, not only is the composition of an individual's

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.509v1 | CC-BY 4.0 Open Access | rec: 22 Sep 2014, publ: 22 Sep 2014

microbiome highly personalized, but their degree of temporal variability is also a personalized feature.

Keywords:
Human microbiome, microbial community ecology, bacteria, 16S rRNA, DNA sequencing, Illumina, temporal dynamics

Background:

The increasing recognition that commensal and mutualistic microorganisms are necessary for many aspects of normal human physiology has altered the traditional pathogen-dominated view of human-bacterial interactions [1, 2]. However, before we can begin to manage, restore, and/or exploit our microbial partners in ways that promote human health, we first must have a comprehensive understanding of how and why these communities vary through time. Previous studies that have characterized human associated microbial communities over time have been based on relatively few individuals [3, 4], intermittent sampling intervals [2, 5, 6], single body habitats [4, 7-10] or focused on disease states [11], leaving us with an incomplete picture of the range of normal variability in the human microbiome.

Here, we investigated the temporal dynamics of forehead, gut (feces), palm, and tongue microbial communities of 85 college-age adults (median age = 21 years) from three U.S. universities. Samples were self-collected weekly over a three-month period beginning in January 2012. Bacterial and archaeal communities were characterized using high-

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.509v1 | CC-BY 4.0 Open Access | rec: 22 Sep 2014, publ: 22 Sep 2014

throughput sequencing of the variable region 4 (V4) of the 16S rRNA gene [12]. In total, we generated 170,563,932 quality-filtered sequences from 3,655 samples, with all analyses conducted on samples rarified to exactly 10,000 sequences per sample. To identify potential drivers of variability, we collected demographic, lifestyle, and hygiene data at the initiation of the sampling period using a standardized 49-question survey (Additional file 1). Weekly questionnaires were used to track changes in health status, medication use, menstrual cycle for women, and other dramatic changes in routine behavior (Additional file 2) De-identified responses to all questions are provided in Additional file 3.

Results and Discussion:

To quantify the amount of temporal variability in diversity of each body habitat, we calculated the coefficient of variation (CV = standard deviation/mean) for three alpha diversity metrics (phylogenetic diversity, phylotype richness, and Shannon index) for each individual [13]. Low CV values indicate that an individual had relatively stable alpha diversity levels, whereas high CV values indicate than an individual had variable levels of alpha diversity over the three-month study period. As evident in Figure 1a, there was a wide range of variability within each body habitat indicating that some individuals varied more than others. When we compare values across body habitats, we see that skin surfaces, particularly the palm, exhibited higher levels of temporal variability in diversity than gut or tongue (Figure 1a). These patterns were generally consistent regardless of the alpha diversity metric used. Skin surfaces also hosted the most diverse communities we surveyed (Additional file 4), as theory would predict given that uncovered regions like

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.509v1 | CC-BY 4.0 Open Access | rec: 22 Sep 2014, publ: 22 Sep 2014

the forehead and palm are more regularly exposed to a larger number of taxa able to disperse onto the skin surface [14] (i.e. a large species pool). Ecological theory also predicts that habitats with large species pools should vary more through time [14], which is what we observed here.

This high degree of temporal variability in alpha diversity levels was matched by high variability (and hence instability) in community membership (Figure 1b). Comparing the proportion of phylotypes shared among time intervals within an individual shows that fewer phylotypes were shared through time in skin communities than in the gut or tongue communities. For example, on average only 15% of the phylotypes observed on the palm skin surface (excluding singletons on a per individual basis) were observed at any other point in time, whether samples were collected one or six weeks apart. A similar pattern was observed when we used median unweighted UniFrac values [15] (a phylogenetic metric of community membership) for each body habitat, where turnover was found to be greater for the skin than for the tongue and gut communities (Figure 2a and Additional file 5a). In contrast, variability in community structure, which accounts for phylotype abundance (median weighted UniFrac), was lower on the forehead than the other body habitats (Figure 2b and Additional file 5b), suggesting that the nature of variability differs depending on the body habitat in question. On the tongue and in the gut, changes in the relative abundance of persistent taxa (i.e., those taxa that are consistently present over time) drive the temporal dynamics, whereas temporal variability in forehead communities appears to be driven more by the presence or absence of transient taxa on the skin surface. For the palm, both membership and structure appear highly dynamic, likely due to

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.509v1 | CC-BY 4.0 Open Access | rec: 22 Sep 2014, publ: 22 Sep 2014

frequent hand washing and exchange of microbes with the numerous objects we touch on a daily basis, including our other body parts.

Using median UniFrac values for each individual as our metric of temporal variability in community membership (unweighted) and structure (weighted), we found that individuals differed dramatically not only in the composition of their microbial communities (Additional file 6), as has been observed previously [2, 5, 16, 17], but also in the degree of temporal variability observed in their microbial communities (Figure 2 and Additional file 7). This has been previously shown in vaginal communities [8], but we show here that this is a general characteristic of microbial communities across human body habitats. The variability of microbial communities in one body habitat, in general, did not predict the variability of microbial communities of other body habitats. The exception was the two skin habitats, where individuals that had more variable forehead communities also had more variable palm communities (Additional files 8 and 9). This finding suggests that the factors that contribute to intra-personal temporal variability in microbiome composition are shared across skin habitats, but not necessarily across other body habitats. Furthermore, relatively few individuals exhibited a significant time-decay relationship [18]; in general, samples collected closer together in time did not harbor more similar communities than those collected further apart in time (Additional file 10). These results highlight that attempts to predict what type of communities to expect in a given body habitat based on data collected during the previous week (or weeks) may be difficult for most individuals. However, it is important to note that if we had sampled more frequently (e.g., on a daily basis [3]) or for a longer period of time, we may have

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.509v1 | CC-BY 4.0 Open Access | rec: 22 Sep 2014, publ: 22 Sep 2014

been able to identify a stronger relationship between elapsed time and the composition of the communities within body habitats.

Having established that the degree and nature of variability was specific to each body site and was in itself an important parameter that distinguished individuals from one another, we next sought to identify factors associated with this variation across individuals. Based on previous work [4, 19], we expected that antibiotic usage would lead to profound shifts in the structure of an individual's microbiome. Indeed, within a given individual, the largest shifts observed in community membership coincided with the time points that the individual reported having taken oral antibiotics ($p < 0.001$ for both unweighted and weighted UniFrac, Monte Carlo t-test with 1000 iterations). Across the study population, however, with the exception of palm communities, we did not find a significant effect of antibiotics on variability in community membership and structure; individuals who took antibiotics did not, on average, have more variable communities than those that did not take antibiotics over the time period of this experiment (Figure 2). Our observation that antibiotic use was not associated with increased temporal variability in microbial communities across the study population could be due to the fact that we did not control for the timing of sampling relative to antibiotic use, dosage, or type of antibiotics used by the individuals sampled here, or it may be because microbial community responses to antibiotics are highly individualized, as suggested by recent work [4, 20].

We next used generalized linear models (GLMs) to identify which other factors or combination of factors best predicted why some individuals harbored more variable

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.509v1 | CC-BY 4.0 Open Access | rec: 22 Sep 2014, publ: 22 Sep 2014

185 microbial communities than others. For these models, we again used median weighted or
186 unweighted UniFrac values of each individual as our response variables for each body
187 habitat. Potential predictive factors were compiled from the initial survey responses
188 (Additional file 2) and we only included factors for which we had sufficient replication in
189 survey responses (Additional file 3). Presented models included factors with a
190 significance value < 0.05. As shown in Table 1, our models were often able to explain
191 much of the variability in the temporal stability of microbial communities across
192 individuals, but the strength of the models was dependent on the body habitat in question
193 or the distance metric used. Common predictive factors observed in multiple body
194 habitats included median alpha diversity values (Shannon Index), university affiliation
195 and antibiotic use (Table 1). However, the strongest predictive variable for most body
196 habitats was median diversity, measured using the Shannon index, suggesting an overall
197 relationship between diversity and variability. Other factors appeared to have a body site-
198 specific affect. For example, the number of roommates helped explain a significant
199 amount of variability in the structure (weighted) of forehead microbial communities, a
200 pattern that may driven by the exchange of skin bacteria between individuals sharing a
201 common living area.
202
203 To explore the relationship between diversity and temporal variability in greater detail,
204 we generated single-factor linear models using median Shannon index values as our
205 metric of diversity and either median weighted or unweighted UniFrac values as our
206 metric of stability (Figure 3). With these models, we observed statistically significant
207 negative correlations between diversity and compositional variability for the gut and

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.509v1 | CC-BY 4.0 Open Access | rec: 22 Sep 2014, publ: 22 Sep 2014

tongue communities; individuals with more diverse communities were less variable (more stable) than individuals with less diverse communities. In contrast, a positive relationship was observed between forehead community diversity and structural variability while no relationship was evident for palm communities. Similar directional patterns were observed with the other diversity metrics (Additional file 11). Our finding that microbial communities which experience lower rates of immigration (the gut and tongue) exhibit a positive diversity-stability relationship parallels patterns observed in many plant and animal communities where increases in species diversity often result in more stable communities and communities that are more resistant to invasions (i.e. the portfolio effect) [21]. Although the health implications of the diversity-stability relationships observed here remain undetermined, recent work has shown that gut communities of lower diversity are often associated with disease phenotypes in humans [22].

Individuals that had more stable communities harbored taxonomically distinct communities compared with those found in more variable individuals (Figure 4). For example, individuals with stable forehead communities had a greater relative abundance of *Staphylococcaceae* and *Corynebacteriaceae*, whereas individuals with highly variable forehead communities were enriched in *Streptococcaceae* and *Lactobacillaceae* (Figure 4a). The trade-off between *Staphylococcaceae* and *Lactobacillaceae* is intriguing because several *Lactobacillaceae* species inhibit attachment of *Staphylococcaceae* to epithelial cells [23, 24]. In the gut, two of the dominant groups of Firmicutes, *Clostridiaceae* and *Lactobacillaceae*, were more abundant in variable individuals, whereas the

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.509v1 | CC-BY 4.0 Open Access | rec: 22 Sep 2014, publ: 22 Sep 2014

231 *Bacteroidaceae* (the dominant family within the Bacteroidetes phylum) were most
232 abundant in stable individuals (Figure 4b). A higher Firmicutes:Bacteriodetes ratio has
233 been observed in guts of obese individuals [25, 26], but we did not have enough diversity
234 in body mass index (BMI) to formally test if temporal variability may also be associated
235 with obesity. Although the mechanisms underlying these patterns remain unclear, these
236 observations highlight the likely importance of bacterial interactions in determining the
237 stability of human-associated microbial communities.
238
239 **Conclusions:**
240 Our findings suggest that the high degree of temporal variability in alpha diversity levels,
241 community membership, and community structure observed across the sampled body
242 habitats and across study participants is important to consider when designing studies to
243 assess linkages between the human microbiome and health. Although the variability in
244 community composition among individuals typically exceeds the temporal variability
245 within individuals over time, a pattern we also observed here (Additional file 6), the
246 intra-individual temporal variability is considerable and the degree of variability that an
247 individual experiences over time may be a (largely unexplored) factor in determining
248 disease state or differential treatment success. Further, because variability through time
249 can be high, samples collected at one point in time may not adequately characterize an
250 individual's microbiome, even if focusing on only the more abundant phylotypes (Figure
251 1b, dark shades). If the effect size of a change in disease state on the human microbiome
252 is sufficiently large (e.g., the loss of a major lineage), this intra-individual temporal
253 variability may be irrelevant. However, if changes in disease state are associated with

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.509v1 | CC-BY 4.0 Open Access | rec: 22 Sep 2014, publ: 22 Sep 2014

more subtle shifts in microbial community composition, it would be important to control for this temporal variability before one could establish causal linkages between changes in the microbiome and changes in health status. It is now well established that there is considerable inter-individual variability in the composition of the human microbiome [5, 17], leading to the concept of a "personal microbiome", and we are beginning to establish causal relationships between composition of the microbiome and disease [27]. Here we show that there is also a high-degree of inter-individual variability in the stability of the human gut, tongue, forehead, and palm microbiome. As a result, we suggest that the "personal microbiome" concept should be extended to incorporate the rate of change of an individual's microbiome, in addition to its composition (a feature which distinguishes the "personal microbiome" from the "personal genome") and that future investigations into associations between features of the microbiome and host phenotype may want to consider temporal variability as a potential explanatory factor.

Methods

Subject recruitment and sample collection.

Volunteers were recruited from three Universities (University of Colorado, Boulder (UCB), Northern Arizona University (NAU), and North Carolina State University (NCSU)) in January/February of 2012 and asked to donate weekly self-collected samples for a minimum of ten weeks using sterile, pre-labeled, double-tipped swabs (Becton, Dickinson and Company, Sparks, MD, USA.). Participants were instructed to sample two skin habitats (foreheads and palms) and the surface of their tongue by swabbing for 10-15

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.509v1 | CC-BY 4.0 Open Access | rec: 22 Sep 2014, publ: 22 Sep 2014

277 seconds. Gut (fecal) samples were collected by touching cotton swabs to used toilet paper
278 so that a small amount of fecal material was transferred to each pair of swabs. Volunteers
279 were asked to collect samples before showering and as close to drop-off times as possible
280 without placing samples in freezers to avoid freeze-thaw cycles. One representative at
281 each University collected samples from students and placed them in a -20 °C freezer until
282 shipping on dry ice to the UCB where all sample processing occurred. Volunteers were
283 also asked to provide a variety of demographic and behavioral metadata at the initiation
284 of the project using a scantron-based survey (Additional file 1). Weekly questionnaires
285 (Additional file 2) were also provided with sampling kits to collect information on
286 changes in health status, medication use, and menstruation for women. At the conclusion
287 of the study prior to publication, study participants were provided their personalized
288 results via a password-protected website (http://my-microbes.qiime.org). All volunteers
289 were made aware of the nature of this project and gave written consent in accordance
290 with protocols approved by each University's Institutional Review Board (IRB) (UCB
291 409.13; NAU 12.0169; NCSU 2443). Per IRB regulations, volunteers were able to drop
292 out of the study at any time and were not required to answer any or all survey questions.
293
294 *Sample processing.*
295 Samples from NAU and NCSU were shipped on dry ice to UCB at the conclusion of
296 sampling. Upon arrival, individual swabs were linked with Personal IDs using digital
297 barcodes and logged into an Excel worksheet. Swabs were then sorted by body habitat
298 and the tip of one duplicate swab was aseptically cut into single wells in 2 ml 96-well
299 deep-well plates (Axygen Inc., Union City, CA, USA). Plates were sealed with silicone

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.509v1 | CC-BY 4.0 Open Access | rec: 22 Sep 2014, publ: 22 Sep 2014

Axymat sealing mats (Axygen Inc., Union City, CA, USA). Each plate contained negative control samples that included swab blanks (sterile swabs), extraction blanks (reagents), and a PCR control. Forehead, gut, and tongue plates also included positive controls that were collected from one individual at the initiation of the project and stored/shipped with samples at each university. No differences were observed in community membership or structure in positive control samples.

DNA extraction, PCR amplification, and sequencing.

DNA extraction and PCR amplification of the variable region 4 (V4) of the 16S rRNA gene using Illumina adapted universal primers 515F/806R [12, 28] was conducted using the direct PCR protocol as previously described [29].

Aliquots (4 µl) from the fecal and tongue extracts were transferred into 384-well plates for triplicate PCR reactions, while skin aliquots (forehead and palm, 4 µl) were transferred into 96-well plates. PCRs were conducted in triplicate 20 µl reactions and thermal cycling conditions for the 384-well plates were: initial denaturation for 3 min at 94 °C; 35 cycles (94 °C, 60 sec; 50 °C, 60 sec; 72 °C, 105 sec) followed by a final elongation for 10 min at 72 °C. Conditions for the 96-well plates were identical except for shorter denaturation (94 °C, 45 sec) and elongation (72 °C, 90 sec) steps. PCR products from triplicate reactions of each sample were pooled, visualized on an agarose gel, and quantified using the PicoGreen dsDNA assay (Invitrogen, Carlsbad, CA, USA). Positive amplicons from each body habitat (forehead, gut, palm, and tongue) were then pooled in equimolar concentrations into composite samples that were cleaned using a

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.509v1 | CC-BY 4.0 Open Access | rec: 22 Sep 2014, publ: 22 Sep 2014

single-tube MoBio Ultraclean PCR Clean-up Kit (MoBio Laboratories, Carlsbad, CA USA). Each body habitat was sequenced on an individual lane (4 lanes total) of an Illumina HiSeq2000 instrument at the University of Colorado BioFrontiers Institute Advanced Genomics Facility.

Data processing.

All data processing was performed using QIIME 1.6.0-dev unless otherwise noted. The specific processing steps were as follows. Raw fastq data was demultiplexed and quality filtered as described previously [30]. Sequences that passed quality filtering were clustered into phylotypes (Operational Taxonomic Units, OTUs) at 97% sequence identity using a uclust-based [31] closed-reference protocol against the 12_10 revision of the Greengenes database [32], where reads that did not match a sequence in the reference set at least 97% identity were excluded from subsequent analyses. The taxonomy of each phylotype was assigned as the taxonomy associated with the Greengenes sequence defining that OTU. The Greengenes phylogenetic tree was used for phylogenetic diversity calculations. A median of 49242.0 sequences was collected per sample. After removing phylotypes appearing in negative controls at high abundance ($\geq$ 0.5% across all controls) [29], all samples were rarefied to 10,000 sequences for all downstream analyses unless otherwise noted.

Potentially mislabeled samples were detected using the random forest classification approach described previously [33]. Briefly, the full sample-by-phylotype abundance matrix (i.e., OTU table) was filtered to exclude phylotypes that were observed in fewer

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.509v1 | CC-BY 4.0 Open Access | rec: 22 Sep 2014, publ: 22 Sep 2014

than 10 samples. The OTU table was then randomly subsampled to exactly 1000 sequences per sample. Three samples achieved a probability of being mislabeled greater than 90%, and were excluded from all downstream analyses.

Alpha diversity metrics (phylogenetic diversity (PD), phylotype richness, and Shannon Index) were computed as implemented in QIIME. Comparisons of alpha diversity presented in this study are computed at exactly 10,000 sequences per sample. Beta diversity was computed using the weighted and unweighted UniFrac metrics [15] at exactly 10,000 sequences per sample.

The *time series samples* were defined as the set of samples that came from an individual's body site where at least seven samples were collected and successfully sequenced from that individual's body site over the ten-week collection period. For example, if six fecal samples and seven forehead samples were sequenced from an individual, their fecal samples would not be included in any time series analyses, but their forehead samples would be. This resulted in 75 gut time series sample collections, 80 tongue time series sample collections, 80 forehead time series sample collections, and 61 palm time series sample collections.

All QIIME commands for performing these processing steps can be found in Additional file 12.

Statistical analysis.

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.509v1 | CC-BY 4.0 Open Access | rec: 22 Sep 2014, publ: 22 Sep 2014

To assess the temporal variability of within sample diversity (alpha diversity), we calculated the coefficient of variation (CV) for three diversity metrics (phylogenetic diversity – PD, OTU richness, and Shannon index) for each body habitat of each individual through time. Individual values were used to determine the per body site median across the study population, with higher values indicative of more variable communities.

Variability in community composition (beta diversity) was determined per body habitat by calculating the median weighted and unweighted UniFrac distances for each individual over time. With this metric, communities with a higher median value are more variable whereas a lower value indicates more stable communities. (Note that because we summarize temporal data in a single measurement, we do not need to account for lack of independence of temporal samples from a single individual in evaluations based on this metric.) Differences across body sites for both alpha- and beta-diversity were assessed using the non-parametric Kruskal-Wallis one-way analysis of variance with pairwise comparisons made using the Mann-Whitney U-test, as implemented in R.

To determine the number of phylotypes shared by an individual over different windows of time, we converted the OTU tables of each body habitat to a presence/absence matrix, split it by individual, filtered out singletons, and determined the number of OTUs found in exactly two samples, three samples, four samples, and so on up to seven samples using a custom R script. Samples did not have to be from consecutive weeks. We repeated this analysis on only the top 10% most abundant OTUs per individual. The numbers of

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.509v1 | CC-BY 4.0 Open Access | rec: 22 Sep 2014, publ: 22 Sep 2014

phylotypes shared per individual were then averaged across individuals for each window of time and each body habitat.

For each body habitat, the study population was divided into quartiles based on median intra-individual UniFrac values where the 1st quartile was defined as 'stable', the 2nd and 3rd quartiles as 'average', and the 4th quartile as 'variable.' To determine if certain taxa were more or less abundant in the different quartiles (i.e. stability classes), we rank transformed the most abundant bacterial families (> 1% across individuals) for each body habitat and tested for differences between the groups using the nonparametric Kruskal-Wallis analysis of variance.

adonis [34], ANOSIM [35], and PERMDISP [34] (using 999 permutations) were used to test for differences in community composition between individuals at each body site. The statistical methods were used to analyze both weighted and unweighted UniFrac distance matrices, with only the *time series samples* being included in the analyses.

To determine the affect of antibiotic use on community variability, we grouped individuals based on their usage (yes or no) and used the non-parametric Mann-Whitney U-test to test for differences between the two groups. Spearman rank correlations were used to determine if community variability as measured using median UniFrac distances was correlated across pairs of body habitats. To assess if patterns in community composition could be related to time between sampling events, Mantel tests (Spearman-rank correlations on 999 permutations) were conducted for each individual using both

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.509v1 | CC-BY 4.0 Open Access | rec: 22 Sep 2014, publ: 22 Sep 2014

weighted and unweighted UniFrac values and Manhattan time-distance matrices calculated in R using the VEGAN package (33). Using the mean of the different alpha diversity metrics (PD, phylotype richness, and Shannon index) as our metrics of diversity and median UniFrac distances (both weighted and unweighted) as our metric of community variability on a per individual basis, we constructed linear models for each body habitat across individuals to examine the relationship between diversity and stability.

We identified key predictors of the variability in composition of bacterial assemblages using generalized linear models. We used a model simplification procedure, removing non-significant terms ($\alpha = 0.05$) in a stepwise fashion [36], to explore the relative contributions of the various terms included in the start model. Model simplification approaches have been criticized [37], but in the absence of strong *a priori* information on the drivers of variability of bacterial assemblages, this approach is a useful first step [38]. The final models we present are those that exclusively include variables that explain significant variation in our dependent variables. We also used model simplification in which final models were those in which Bayesian information criterion (BIC) was minimized. However, these "best" models ended up including all variables we tested and so here we focus on those variables with significant explanatory power.

To determine if the weeks where individuals reported taking antibiotics were the weeks where they experience the largest changes in their gut community compositions, we ran per-body-site one-tailed, rank-based Monte Carlo t-tests. The adjacent-week UniFrac

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.509v1 | CC-BY 4.0 Open Access | rec: 22 Sep 2014, publ: 22 Sep 2014

distances were compiled for each individual on a per-body-site basis (i.e., the distance between their gut samples on week 1 and week 2, week 2 and week 3, and so on). Each individual's UniFrac distances were ranked from smallest to largest, and assigned their rank value. Across individual ranks were grouped into distributions based on whether they occurred in a week where the individual reported taking antibiotics or not. Those distributions were then compared with a one-tailed Monte Carlo t-test with 1000 iterations.

Data availability: Sequence data and accompanying de-identified metadata has been deposited in the EMBI under accession number (**ERP005150-ERP005153**) and can also be found in the QIIME Database under the study ID **(will add ID prior to publication)** (http://www.microbio.me/qiime/).

Abbreviations:

BIC: Bayesian information criterion; BMI: body mass index; CA: California; CV: coefficient of variation; dsDNA: double-stranded deoxyribonucleic acid; GLM: generalized linear model; IRB: Institutional Review Board; MD: Maryland; ml: milliliter; NAU: Northern Arizona University; NCSU: North Carolina State University; OTU: operational taxonomic unit; PCR: polymerase chain reaction; PD: phylogenetic diversity; QIIME: Quantitative Insights Into Microbial Ecology; rRNA: ribosomal ribonucleic acid; sec: seconds; UCB: University of Colorado, Boulder; µl: microliter; U.S.: United States; USA: Unites States of America; V4: variable region 4.

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.509v1 | CC-BY 4.0 Open Access | rec: 22 Sep 2014, publ: 22 Sep 2014

Acknowledgments:

We thank MJ Gebert, G Humphrey and CL Lauber for logistical support throughout the project; N Rountree for sample collection; A Shade for providing custom R scripts; G Ackermann for IRB assistance; and M Watwood, E Schwartz, and J Wilder for allowing us to recruit volunteers in their classes. Computational support was provided by an Amazon Web Services research grant to JGC and RK Support for YVB, AG, and RK came from the National Institutes of Health, the Crohns and Colitis Foundation of America, and the Howard Hughes Medical Institute. RRD. was supported by grant #52006933 to NC State University from the Howard Hughes Medical Institute through the Precollege and Undergraduate Science Education Program (http://www.hhmi.org/grants/office/undergrad/).

Author Contributions:

GEF conceived of and designed the study, collected samples, analyzed and interpreted the data, and wrote the paper. JGC conceived of and designed the study, analyzed and interpreted the data, and wrote the paper. JBH conducted all wet lab work. JRR analyzed and interpreted the data. DD collected samples and analyzed and interpreted the data. JC analyzed and interpreted the data. JWL analyzed and interpreted the data. YVB analyzed and interpreted the data. AG analyzed and interpreted the data. RK conceived of and designed the study, analyzed and interpreted the data, and wrote the paper. RRD conceived of and designed the study, analyzed and interpreted the data, and wrote the paper. NF conceived of and designed the study, analyzed and interpreted the data, and wrote the paper.

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.509v1 | CC-BY 4.0 Open Access | rec: 22 Sep 2014, publ: 22 Sep 2014

Additional files:

The following additional data are available with the online version of this paper. Additional file 1 is the pre-study questionnaire used to collect demographic, lifestyle, and hygiene data on study participants. Additional file 2 is the weekly questionnaire used to collect information about changes in health status, medication use, stage of menstrual cycle for women, and any other dramatic changes in the routine of study participants. Additional file 3 is a list of all samples collected in this study with corresponding de-identified personal IDs of study subjects and answers to survey questions. Additional file 4 is a figure showing the amount of microbial diversity observed in each sample. Additional file 5 is a figure depicting the temporal variability observed in microbial community membership and structure for each body habitat of each individual. Additional file 6 is a table showing that the composition of each individual's microbiome is personalized through time. Additional file 7 is a figure showing how the microbial communities of selected individuals vary through time. Additional file 8 is a table showing the results of Spearman rank correlation of community membership across different body habitats. Additional file 9 is a table showing the results of Spearman rank correlation of community structure across different body habitats. Additional file 10 is a table of Mantel test results correlating microbial community membership and structure with time between samples (time distance-decay). Additional file 11 is a table of results correlating microbial diversity with temporal variability in community membership and structure for each body habitat. Additional file 12 is a list of all QIIME commands used in data processing.

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.509v1 | CC-BY 4.0 Open Access | rec: 22 Sep 2014, publ: 22 Sep 2014

References

1. Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, Gordon JI: The human microbiome project. *Nature* 2007, 449:804-810.
2. Structure, function and diversity of the healthy human microbiome. *Nature* 2012, 486:207-214.
3. Caporaso JG, Lauber CL, Costello EK, Berg-Lyons D, Gonzalez A, Stombaugh J, Knights D, Gajer P, Ravel J, Fierer N, et al: Moving pictures of the human microbiome. *Genome biology* 2011, 12:R50.
4. Dethlefsen L, Relman DA: Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation. *Proceedings of the National Academy of Sciences of the United States of America* 2011, 108 Suppl 1:4554-4561.
5. Costello EK, Lauber CL, Hamady M, Fierer N, Gordon JI, Knight R: Bacterial community variation in human body habitats across space and time. *Science* 2009, 326:1694-1697.
6. Grice EA, Kong HH, Conlan S, Deming CB, Davis J, Young AC, Bouffard GG, Blakesley RW, Murray PR, Green ED, et al: Topographical and temporal diversity of the human skin microbiome. *Science* 2009, 324:1190-1192.
7. Camarinha-Silva A, Jauregui R, Pieper DH, Wos-Oxley ML: The temporal dynamics of bacterial communities across human anterior nares. *Environmental microbiology reports* 2012, 4:126-132.
8. Gajer P, Brotman RM, Bai G, Sakamoto J, Schutte UM, Zhong X, Koenig SS, Fu L, Ma ZS, Zhou X, et al: Temporal dynamics of the human vaginal microbiota. *Science translational medicine* 2012, 4:132ra152.
9. Wu GD, Chen J, Hoffmann C, Bittinger K, Chen YY, Keilbaugh SA, Bewtra M, Knights D, Walters WA, Knight R, et al: Linking long-term dietary patterns with gut microbial enterotypes. *Science* 2011, 334:105-108.
10. David LA, Maurice CF, Carmody RN, Gootenberg DB, Button JE, Wolfe BE, Ling AV, Devlin AS, Varma Y, Fischbach MA, et al: Diet rapidly and reproducibly alters the human gut microbiome. *Nature* 2013.
11. Kong HH, Oh J, Deming C, Conlan S, Grice EA, Beatson MA, Nomicos E, Polley EC, Komarow HD, Murray PR, et al: Temporal shifts in the skin microbiome associated with disease flares and treatment in children with atopic dermatitis. *Genome research* 2012, 22:850-859.
12. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, Owens SM, Betley J, Fraser L, Bauer M, et al: Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *The ISME journal* 2012.
13. Shade A, Caporaso JG, Handelsman J, Knight R, Fierer N: A meta-analysis of changes in bacterial and archaeal communities with time. *The ISME journal* 2013, 7:1493-1506.
14. Fukami T: Community assembly along a species pool gradient: implications for multiple-scale patterns of species diversity. *Population Ecology* 2004, 46:137-147.

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.509v1 | CC-BY 4.0 Open Access | rec: 22 Sep 2014, publ: 22 Sep 2014

15. Lozupone C, Knight R: UniFrac: a new phylogenetic method for comparing microbial communities. *Applied and environmental microbiology* 2005, 71:8228-8235.
16. Ravel J, Gajer P, Abdo Z, Schneider GM, Koenig SS, McCulle SL, Karlebach S, Gorle R, Russell J, Tacket CO, et al: Vaginal microbiome of reproductive-age women. *Proceedings of the National Academy of Sciences of the United States of America* 2011, 108 Suppl 1:4680-4687.
17. Fierer N, Lauber CL, Zhou N, McDonald D, Costello EK, Knight R: Forensic identification using skin bacterial communities. *Proc Natl Acad Sci U S A* 2010, 107:6477-6481.
18. P. White E, B. Adler P, K. Lauenroth W, A. Gill R, Greenberg D, M. Kaufman D, Rassweiler A, A. Rusak J, D. Smith M, R. Steinbeck J, et al: A comparison of the species–time relationship across ecosystems and taxonomic groups. *Oikos* 2006, 112:185-195.
19. Dethlefsen L, Huse S, Sogin ML, Relman DA: The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing. *PLoS biology* 2008, 6:e280.
20. Maurice CF, Haiser HJ, Turnbaugh PJ: Xenobiotics shape the physiology and gene expression of the active human gut microbiome. *Cell* 2013, 152:39-50.
21. Tilman D: The ecological consequences of changes in biodiversity: A search for general principles. *Ecology* 1999, 80:1455-1474.
22. Le Chatelier E, Nielsen T, Qin J, Prifti E, Hildebrand F, Falony G, Almeida M, Arumugam M, Batto JM, Kennedy S, et al: Richness of human gut microbiome correlates with metabolic markers. *Nature* 2013, 500:541-546.
23. Prince T, McBain AJ, O'Neill CA: Lactobacillus reuteri protects epidermal keratinocytes from Staphylococcus aureus-induced cell death by competitive exclusion. *Applied and environmental microbiology* 2012, 78:5119-5126.
24. Zarate G, Nader-Macias ME: Influence of probiotic vaginal lactobacilli on in vitro adhesion of urogenital pathogens to vaginal epithelial cells. *Letters in applied microbiology* 2006, 43:174-180.
25. Ley RE, Turnbaugh PJ, Klein S, Gordon JI: Microbial ecology: human gut microbes associated with obesity. *Nature* 2006, 444:1022-1023.
26. Bervoets L, Van Hoorenbeeck K, Kortleven I, Van Noten C, Hens N, Vael C, Goossens H, Desager KN, Vankerckhoven V: Differences in gut microbiota composition between obese and lean children: a cross-sectional study. *Gut pathogens* 2013, 5:10.
27. Smith MI, Yatsunenko T, Manary MJ, Trehan I, Mkakosya R, Cheng J, Kau AL, Rich SS, Concannon P, Mychaleckyj JC, et al: Gut microbiomes of Malawian twin pairs discordant for kwashiorkor. *Science* 2013, 339:548-554.
28. Bates ST, Berg-Lyons D, Caporaso JG, Walters WA, Knight R, Fierer N: Examining the global distribution of dominant archaeal populations in soil. *ISME J* 2011, 5:908-917.

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.509v1 | CC-BY 4.0 Open Access | rec: 22 Sep 2014, publ: 22 Sep 2014

29. **Flores GE, Henley JB, Fierer N: A direct PCR approach to accelerate analyses of human-associated microbial communities. *PLoS One* 2012, 7:e44563.**
30. **Bokulich NA, Subramanian S, Faith JJ, Gevers D, Gordon JI, Knight R, Mills DA, Caporaso JG: Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. *Nature methods* 2013, 10:57-59.**
31. **Edgar RC: Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 2010, 26:2460-2461.**
32. **McDonald D, Price MN, Goodrich J, Nawrocki EP, DeSantis TZ, Probst A, Andersen GL, Knight R, Hugenholtz P: An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *The ISME journal* 2012, 6:610-618.**
33. **Knights D, Kuczynski J, Koren O, Ley RE, Field D, Knight R, DeSantis TZ, Kelley ST: Supervised classification of microbiota mitigates mislabeling errors. *The ISME journal* 2011, 5:570-573.**
34. **Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Wagner H: vegan: Community Ecology Package. In *R package version 20-5*, vol. http://CRAN.R-project.org/package=vegan; 2012.**
35. **Clarke KR: Non-parametric multivariate analyses of changes in community structure. *Australian Journal of Ecology* 1993, 18:117-143.**
36. **Crawley MJ: *Statistical Computing: An Introduction to Data Analysis Using S-Plus.* Chichester, UK: John Wiley & Sons, Ltd.; 2002.**
37. **Whittingham MJ, Stephens PA, Bradbury RB, Freckleton RP: Why do we still use stepwise modelling in ecology and behaviour? *The Journal of animal ecology* 2006, 75:1182-1189.**
38. **Anderson DR: *Model Based Inference in the Life Sciences: A Primer on Evidence.* New York, NY, USA: Springer Science+Business Media, LLC; 2008.**

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.509v1 | CC-BY 4.0 Open Access | rec: 22 Sep 2014, publ: 22 Sep 2014

Figure Legends

Figure 1. Body habitats exhibited different levels of temporal variability both in diversity (a) and membership (b). In **A**, each point represents the temporal variability of a single individual colored by gender (red=female, blue=male) with black bars representing the median for a given body habitat and metric. Statistical differences were observed for each metric across body habitats (Kruskal-Wallis, $p \leq 0.01$) and comparisons based on pairwise Mann-Whitney U-test are denoted by asterisks (* = corrected $p \leq 0.05$, ** = corrected $p \leq 0.01$). In **B**, the smaller, lighter shaded bars in each plot are for all phylotypes except singletons and the larger, darker bars are only for the 100 most abundant phylotypes for each individual. Error bars in **B** are ± 1 SEM.

Figure 2. Boxplots of unweighted (a) and weighted (b) intra-individual UniFrac distances for each body habitat. A broad range of temporal variability in microbial community membership (**a**) and structure (**b**) was observed across body habitats and within body habitats across individuals. Individuals are sorted by median in each plot. Green bars depict individuals who did not report antibiotic use during the study period while blue bars indicate individuals who took antibiotics. The median values for each body habitat are shown with vertical red lines. Dotted horizontal lines in each plot divide the study population into 1st and 4th quartiles and depict 'stable' and 'variable' individuals, respectively. Non-parametric Mann-Whitney U-tests were used to determine the affect of antibiotic use on temporal variability within each body habitat. P-values are shown in each panel. Note that statistical differences were observed for each metric across body habitats (Kruskal-Wallis, $p \leq 0.01$).

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.509v1 | CC-BY 4.0 Open Access | rec: 22 Sep 2014, publ: 22 Sep 2014

Figure 3. Relationship between diversity and variability of microbial communities associated with each body habitat. Diversity was measured as the median Shannon Diversity Index for each individual over the three-month sampling period. Variability was measured as intra-individual median weighted (white boxes) and unweighted (grey circles) UniFrac distance. Each point represents values of the time-series data for one individual. Spearman rank correlation coefficients are presented for statistically significant relationships ($p \leq 0.01$). Note that similar patterns were observed with other alpha diversity metrics (Additional file 11).

Figure 4. Average taxonomic composition was different among stability classes across individuals. Individuals were assigned to stability classes based on quartiles (1^{st} = stable (blue), 2^{nd} & 3^{rd} = average (red), 4^{th} = variable (green)) of median weighted UniFrac distances for each body habitat. Significant differences were observed across forehead (**a**) and gut (**b**) communities but not in palm (**c**) or tongue (**d**) communities as determined by rank transforming the most abundant bacterial families (> 1% in any group) for each body habitat and testing for differences between stability classes using the nonparametric Kruskal-Wallis analysis of variance. Significance is denoted with asterisks (* = corrected $p \leq 0.05$, ** = corrected $p \leq 0.01$).

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.509v1 | CC-BY 4.0 Open Access | rec: 22 Sep 2014, publ: 22 Sep 2014

674 **Table 1. Measured factors that influenced the temporal variability of the human**
675 **microbiome.** Generalized linear models (GLMs) were used to determine which of the
676 measured factors or combination of factors best predicted variability in microbiome
677 membership (unweighted UniFrac) and structure (weighted UniFrac). Unweighted
678 UniFrac distances are a metric of the phylogenetic dissimilarity of samples through time.
679 Weighted UniFrac distances weight dissimilarity both as a function of the phylogenetic
680 dissimilarity and the relative abundance of taxa (such that two samples with the same
681 phylogenetic dissimilarity are considered more different if one is dominated by a
682 particular taxon).

	Est	SS	F	P	BIC	R^2
Forehead – unweighted						
Antibiotic use	-0.015	0.010	8.76	0.004	-262.21	0.175
University	-0.119	0.006	5.41	0.023	-263.38	
Forehead – weighted						
Median Shannon	0.038	0.090	32.2	3.61 e -7	-190.67	0.580
Gender	-0.023	0.027	9.54	0.003	-193.54	
Number of roommates	-0.039	0.016	5.70	0.02	-196.16	
Gut – unweighted						
Median Shannon	-0.063	0.081	73.24	4.3 e -12	-240.90	0.570
Over the counter acne product	0.014	0.013	11.18	0.001	-249.97	
University	-0.014	0.007	6.64	0.012	-254.34	
Gut – weighted						
Median Shannon	-0.107	0.238	20.64	2.61 e -5	-85.83	0.319
Over the counter acne product	-0.034	0.065	5.65	0.021	-90.53	
University	-0.028	0.047	4.08	0.047	-90.55	
Palm – unweighted						
Exercise frequency	-0.033	0.022	15.74	2.00 e -4	-188.6	0.310
Lives with dogs	-0.014	0.010	7.18	0.009	-189.9	
Number of roommates	-0.016	0.008	5.77	0.019	-191.8	
Palm – weighted						
Antibiotic use	-0.026	0.024	4.97	0.029	-129.8	0.080
Tongue – unweighted						
Antibiotic use	-0.018	0.015	7.75	0.007	-217.82	0.215
Median Shannon	-0.038	0.010	5.5	0.022	-220.12	
Tongue – weighted						
No good model						

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.509v1 | CC-BY 4.0 Open Access | rec: 22 Sep 2014, publ: 22 Sep 2014

683

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.509v1 | CC-BY 4.0 Open Access | rec: 22 Sep 2014, publ: 22 Sep 2014

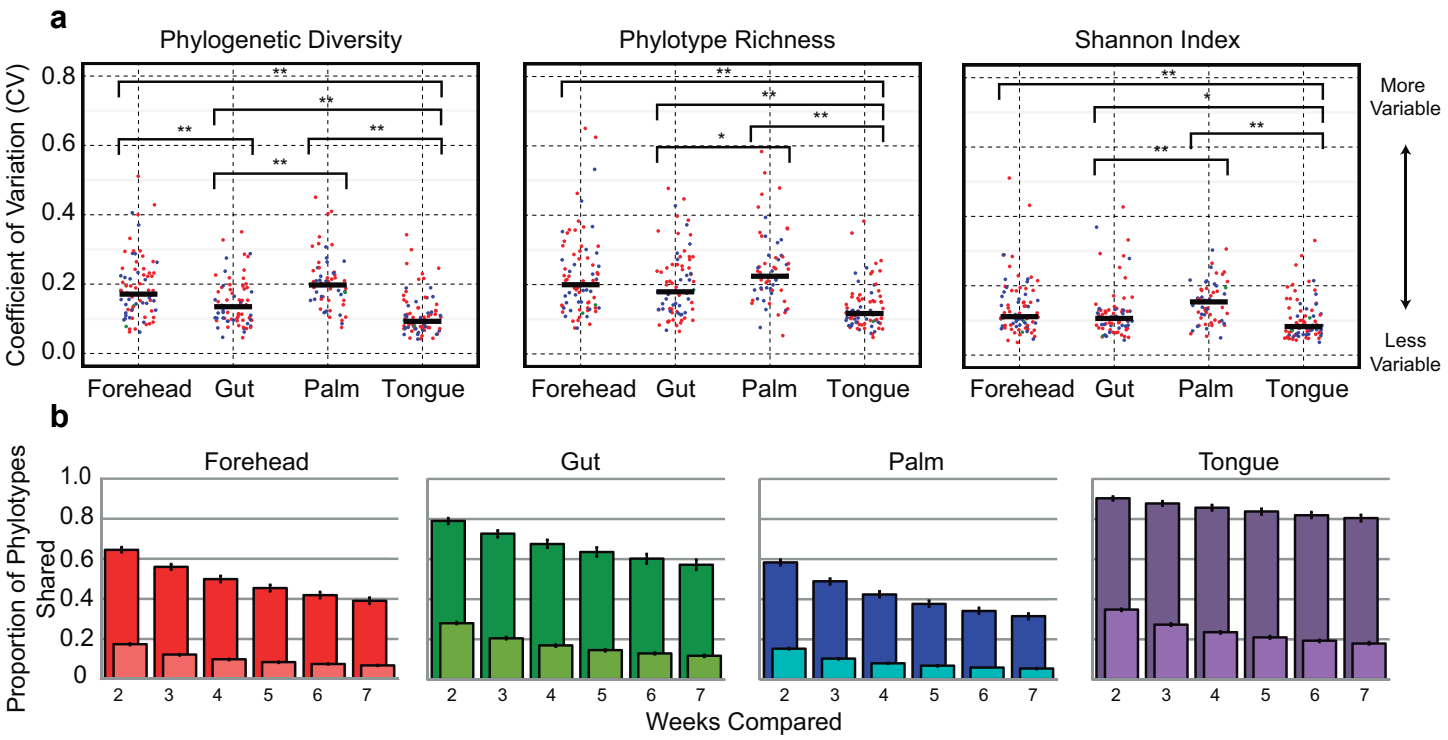

Figure 1

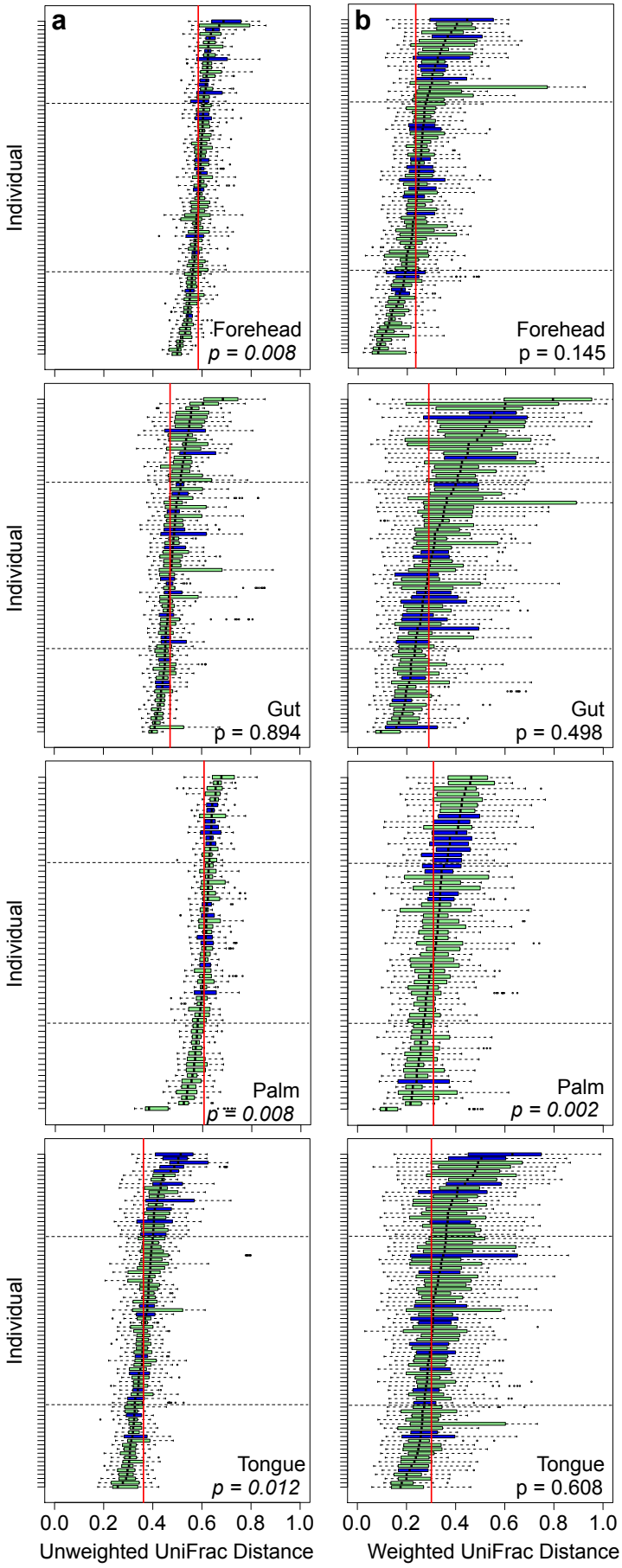

Figure 2

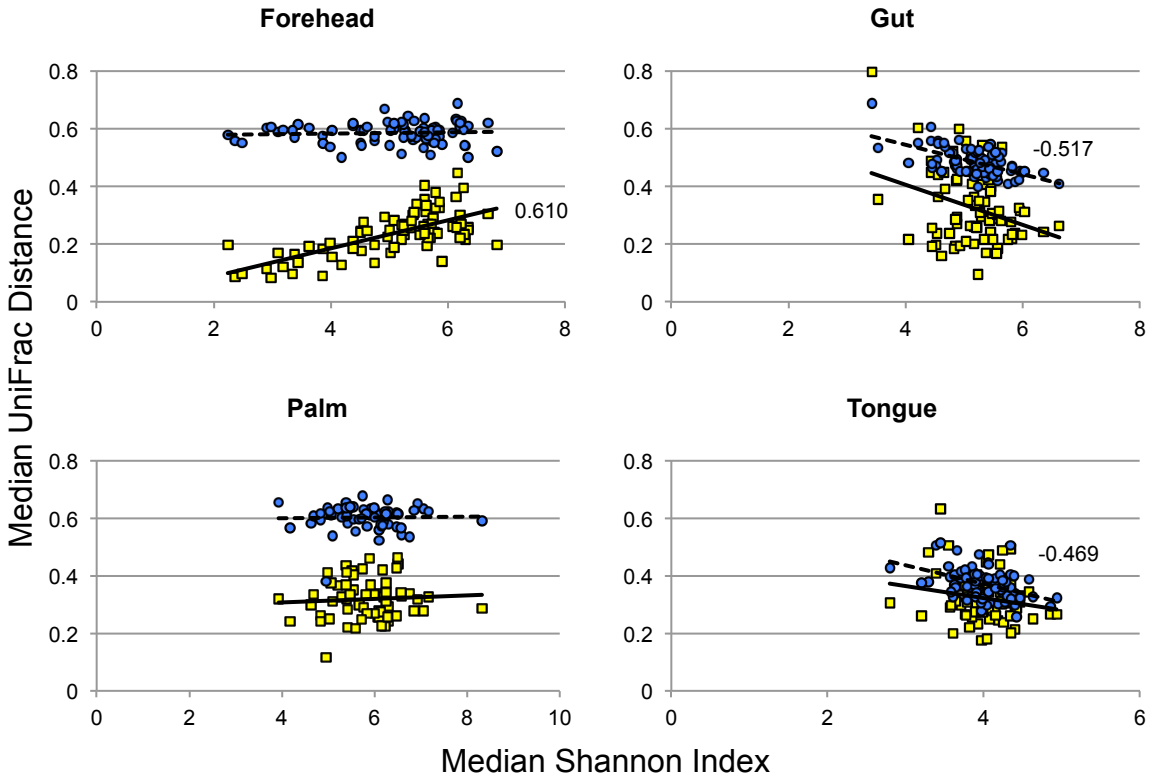

Figure 3

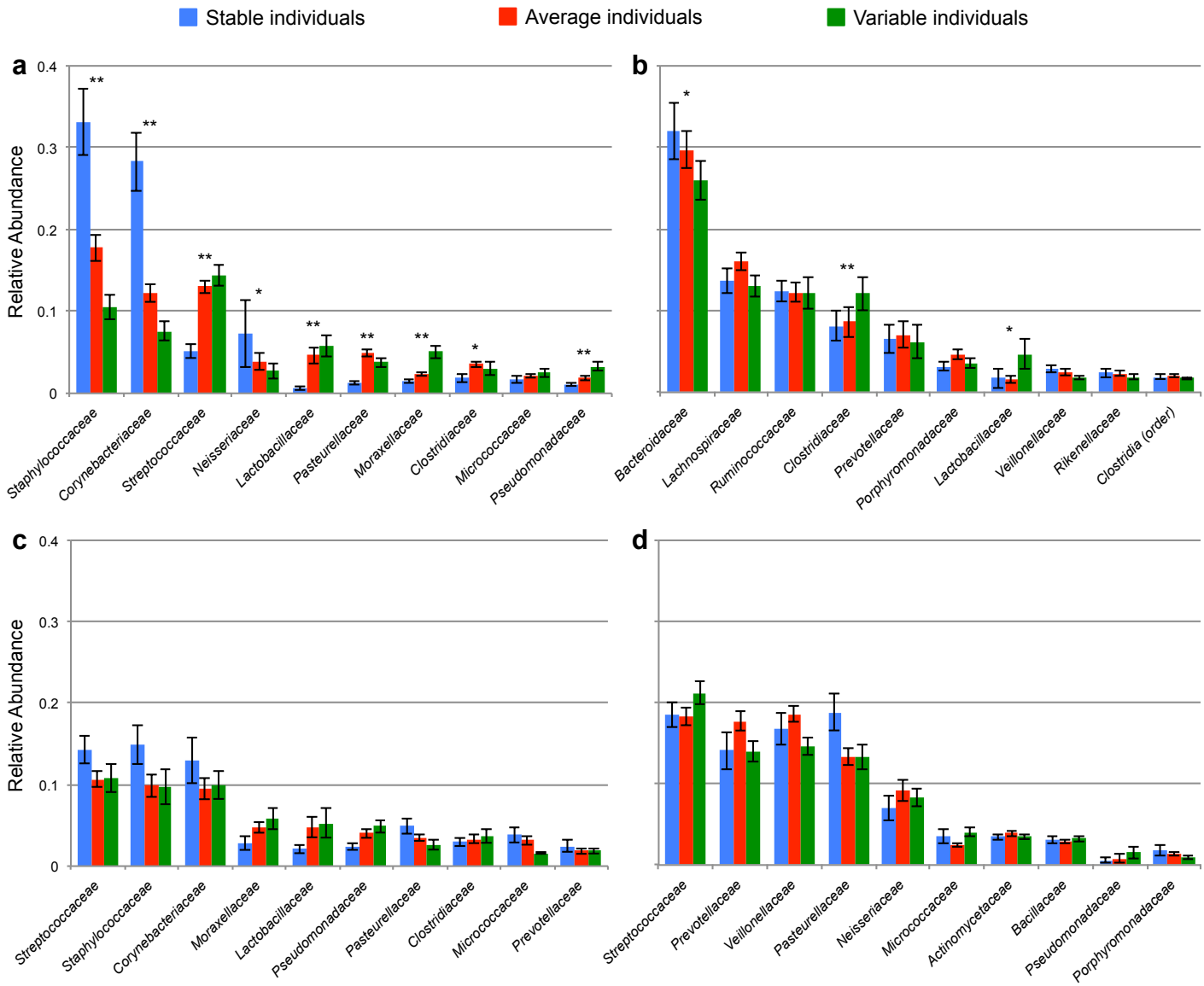

Figure 4