**Title:**

A phylogenetic model for the arrival of species into microbial communities

**Running Title:**

Phylogenetic community assembly of microbes

**Authors:**

John L. Darcy1, Alex D. Washburne2, Michael S. Robeson3, Tiffany Prest1, Steven K. Schmidt1

**Affiliations:**

1Department of Ecology and Evolutionary Biology, University of Colorado. Boulder, Colorado, 80309, USA.

2Department of Microbiology & Immunology, Montana State University. Bozeman, Montana, 59717, USA.

3Department of Biomedical Informatics, University of Arkansas for Medical Sciences. Little Rock, Arkansas, 72205, USA

**Corresponding author:**

J. L. Darcy

[Darcyj@Colorado.EDU](mailto:Darcyj@Colorado.EDU)

Department of Ecology and Evolutionary Biology

Campus Box 334

Boulder, CO, 80309, USA.

**Conflict of interest statement:**

The authors declare that no conflict of interest exists.

**Support:**

Funding was provided by an NSF grant for studying microbial community assembly following disturbance (DEB-1258160). The funding body had no role in study design, analysis, interpretation, or in the preparation of this manuscript.

Abstract

The rules and mechanisms underlying community assembly within the human microbiome are not known, and understanding when and why species are able to join communities is a formidable problem. To address this, we developed a mathematical model that describes the probabilities of different species joining a community over time, within a phylogenetic framework. We use our model to test three hypothetical assembly modes: underdispersion (species are more likely to join a community if a close relative is present), overdispersion (likelihood of joining is higher if a close relative is not present), and the neutral model (likelihood of joining is not related to phylogenetic relationships among species). We applied our model to time-series, high-throughput sequencing data from the human microbiome, and found that communities of the palms and feces had more strongly underdispersed patterns of community assembly than tongue communities. We also found that community assembly in an infant gut microbiome became less underdispersed (but still different than the neutral model) after the subject began consuming baby formula. Surprisingly, all data sets analyzed showed phylogenetic underdispersion, and the magnitude of this underdispersion varied by sampling location on the human body.

Introduction

Every non-sterile surface in the world is in some stage of community assembly, from a forest of tropical trees to the microbes in a mammalian gut. The communities of organisms inhabiting these environments are dynamic through time, and studying patterns of assembly may shine light on general rules that govern their change. Understanding these community assembly rules may aid habitat restoration (Palmer *et al.*, 1997; Temperton, 2004), the management of nature reserves with disturbances (Richards *et al.*, 1999; Bengtsson *et al.*, 2000), and ecological theory of phylogenetic signatures in community assembly (O’Dwyer *et al.*, 2012; Goberna *et al.*, 2014). Patterns and rules of community assembly are particularly important in human systems, including the primary succession of microbes on a human host following birth (Koenig *et al.*, 2011), secondary successions following disease, disturbances caused by host lifestyle or antibiotic use (Peterfreund *et al.*, 2012; David *et al.*, 2014; Kennedy *et al.*, 2016), and the natural turnover of microbial communities over time (Caporaso *et al.*, 2011).

Insights into these difficult-to-observe community assembly processes can be gained via the comparison of microbial communities using high-throughput DNA sequencing (Nemergut *et al.*, 2016; Caporaso *et al.*, 2011; Nemergut *et al.*, 2013), especially in time-series studies (Caporaso *et al.*, 2011; Koenig *et al.*, 2011; Kennedy *et al.*, 2016; Nemergut *et al.*, 2016; Darcy *et al.*, 2017). But one major drawback of culture-independent sequencing studies is that it is difficult to link species with their traits (Martiny *et al.*, 2015). Phylogenetic approaches can compensate for this drawback somewhat, by using the shared evolutionary history of species as a proxy for their ecological (trait-based) similarity (Webb *et al.*, 2002; Wiens *et al.*, 2010). Indeed, phylogenies constructed using 16S rDNA recapitulate core-genome differences among bacteria (Langille *et al.*, 2013; Zaneveld *et al.*, 2010). Within this framework of phylogenetic niche conservatism (Wiens *et al.*, 2010), microbial communities can be phylogenetically over- or underdispersed; “overdispersed” and “underdispersed” communities respectively add more or less phylogenetic diversity during assembly relative to a neutral model of random community assembly (Webb, 2000; Webb *et al.*, 2002; Letcher and Chazdon, 2012; Gerhold *et al.*, 2015).

Here, we use the phylogenetic relationships among species within a time-series to understand the community’s change over time in a new way. Instead of analyzing broad patterns of community change via beta-diversity statistics (UniFrac (Lozupone and Knight, 2005)) or analyzing patterns of select clades within the community (PhyloFactor (Washburne *et al.*, 2017), Edge PCA (Matsen IV *et al.*, 2013)), we model the probability of each new species’ immigration into the community as a monotonic function of its phylogenetic distance to members of the community that have already arrived. Our model allows us to perform two-tailed tests of the null hypothesis that phylogenetic distance to species within an existing community does not affect which species can colonize and the order of colonization (neutral model); rejection of this null hypothesis, and whether the communities are overdispersed or underdispersed relative to the null model, carries information about assembly mechanisms at play.

Since the arrival of species into a community likely signifies the arrival of new traits, this paradigm provides a window into the assembly of ecosystem function. Arriving species that have not previously arrived (first-time arrivals) constitute new phylogenetic diversity being added to a community. Any organism of interest in a human microbiome data set, from the pathogenic to the probiotic, will at some point be a first-time arrival, and the order in which these organisms arrive in the community is determined by community assembly processes (Nemergut *et al.*, 2013). Predicting which lineages of organisms can arrive and establish in a given environment has far-reaching implications for ecosystem remediation and management, especially in microbial communities where the functional traits (Violle *et al.*, 2007) of many microbes are still largely unknown (Martiny *et al.*, 2015). Identifying conditions under which assembly mechanisms change, or under which non-neutral assembly is particular strong, may facilitate microbial community rehabilitation by understanding when and how microbial communities can be colonized by close/distant relatives.

The model we present here can be used to estimate the extent to which first-time arrivals are phylogenetically over- or underdispersed during a time-series microbiome dataset. Said another way, we estimate the extent to which first-time arrivals are “nepotistic”, meaning they are more likely to arrive when a close relative is already present (phylogenetic underdispersion). We fit our model (described in our methods section, below) to several time-series human microbiome data sets (Caporaso *et al.*, 2011; Koenig *et al.*, 2011), to compare patterns of microbial community assembly between subjects, sample sites, or time periods. All microbial community assemblies analyzed here showed a nepotistic arrival pattern – microbes with low phylogenetic distance to the existing community had a higher probability of arriving than microbes with a high phylogenetic distance to the existing community. The universality of non-neutral assembly in the microbial communities analyzed here suggests that the phylogeny carries relevant information for microbial community assembly, with implications for microbiome perturbation and rehabilitation.

Materials and Methods

Overview

Here, we describe a statistical model of phylogenetic microbial community assembly, and how we apply that model to time-series microbiome data. Our goal is to estimate the degree to which species arriving into a community for the first time are phylogenetically over- or underdispersed. Said another way, we want to estimate the extent to which recruitment of species into the community is driven by the arriving species’ phylogenetic similarity to (or distance from) species that already arrived. Our *statistical model* describes the probabilities of those species arriving into a community over time. We use our model with empirical data via S*imulations*, where we re-sample the empirically observed species using our model with known parameter values, to produce surrogate data sets. Specifically, we fix and record the model’s dispersion parameter (*D*), which determines the extent to which species with a close relative are preferentially added to the surrogate community (or, conversely, if species without a close relative are preferred). Our *Parameter Estimation* compares the empirical pattern of species arrival to that of the surrogate data sets (which have known *D* values), in order to determine which value of *D* best describes the empirical data. *Hypothesis* testing is done by comparing the 95% confidence intervals of the *D* estimate to the neutral model, which is *D =* 0. We describe the bioinformatic and technical details of this process in our *Analysis* section, and make our code available to others in the *Code and data accessibility* section.

Statistical model

At any point in time, a community is composed of many species, and other species are not present but are available to be added. Species not yet added (“species pool”) represent organisms present within the metacommunity but not the local community. Our model parameterizes the probability of species arriving in a local community for the first time, based on their phylogenetic distances from species that have already arrived.

In a species-neutral model of community assembly, each species *i* in the species pool has the same probability of arrival at time *t*, irrespective of how different it is from species that are already present at time *t*. Thus, the neutral model for first-time arrivals is a random draw without replacement of species from the species pool. We extend the species-neutral model by modeling the probability *pi,t* of species *i* being observed for the first time at time *t* as

Equation 1

where *di,t* is the phylogenetic distance from species *i* to its closest relative that has already been observed prior to time point *t*, and *D* is a dispersion parameter.

When *D* = 0, our model functions as a neutral model; all species have the same probability of arriving in the community for the first time, since *pi,t* is the same for every species. When *D* < 0, *pi,t* decreases with *di,t* meaning that species from the species pool have higher probabilities of arriving when they are more closely related to species that have already been observed in the local community (underdispersion; phylogenetically constrained). When *D* > 0, the opposite is true (overdispersion; phylogenetically divergent). Our hypothesis testing and parameter estimation focus on the dispersion parameter, *D*.

Simulations

Our analysis of a data set relies on re-constructing that data set via simulation of our statistical model using known values of *D*, allowing for hypothesis testing and parameter estimation. Using the empirical data as a starting point, we simulate many surrogate data sets with *D* values ranging from *D* < 0 (underdispersed) to *D* = 0 (neutral) to *D* > 0 (overdispersed). This is done so that the empirical data can later be compared to the surrogate data sets, to estimate the empirical value of *D.*

We start each surrogate data set with the same species present in the first sample of its corresponding empirical data set. Then, surrogate data sets are constructed forward in time by randomly drawing *Mt* new arrivals from the species pool (all species observed in the empirical data set that have not yet been sampled by this process), where the probabilities of species arriving at any given time are given by Equation 1, and *Mt* is the number of new arrivals in the empirical dataset from times *t*-1 to *t*. The number of arrivals from the empirical data set is used so that species richness is kept constant between the empirical data set and all surrogate data sets. The species pool is updated to exclude those species drawn at previous time points, and the newly sampled species are recorded. This process is repeated for all time points. Surrogate data sets are produced for many different *D* values, ranging from underdispersed to overdispersed models. In the analyses we present here, we performed 500 simulations (as described above) for each data set analyzed.

Parameter Estimation

Our main goal is to estimate the empirical dispersion parameter *D* (Equation 1), which quantifies the degree to which first-time arrivals are phylogenetically underdispersed (*D* > 0), neutral (*D* = 0), or overdispersed (*D* > 0). To this end, we use Faith's phylodiversity (Faith, 1992) to compare each of the 500 surrogate data sets (described above) to the empirical data set. Phylodiversity is the sum of branch-lengths on a phylogenetic tree for a set of species, so phylodiversity of a set of highly related species is low (*i.e*. phylogenetically constrained) because there are no long branch lengths in the tree, but phylodiversity is higher (phylogenetically divergent) for a set of more distantly related species (Faith, 1992). If *D* ≠ 0, then species are preferentially added if they have relatively low (*D* < 0) or relatively high (*D* > 0) phylogenetic distance to the resident community (*dit*, Equation 1), yielding accumulations of total phylodiversity that are relatively slow (*D* < 0) or relatively fast (*D* > 0) compared to the neutral model (Figure 1). In other words, at any time point *t*, the phylogenetic diversity of species that have already been observed is *PDt*, and the extent to which *PDt* accelerates or decelerates over a sampling effort depends on *D*. Because of this, we can estimate *D* by comparing the empirical phylodiversity curve to our surrogate phylodiversity curves, which have known *D* values (Figure 1).

For the comparison of an empirical phylodiversity accumulation curve to curves for corresponding surrogate data sets, we re-scale the absolute phylodiversity accumulation curves, *PDt*, to measure the relative phylodiversity accumulation. In particular, we used relative phylodiversity accumulation curves,

Equation 2

where *PDt* is the accumulated phylodiversity at time point t, *PD0* is the phylodiversity at the initial time point, and *PDtmax* is the phylodiversity at the final time point. We then estimate the empirical value of *D* by minimizing KS distance (Massey, 1951) between surrogate phylodiversity accumulations and the empirical phylodiversity accumulation. Specifically, we create and minimize an error model for KS distances using segmented regression (“broken stick” model) with the R package SiZeR (Sonderegger, 2012). This procedure estimates the empirical *D* value, as well as 95% confidence intervals for that estimate. Segmented regression was used because plots of KS distances are linear (Figure 2).

Hypothesis Testing

For this test, our null hypothesis is the neutral model, where *D* = 0, since this model represents the absence of the effect we are testing (over- or underdispersion, a 2-tailed test). We test against the null hypothesis *D* = 0 by comparing the estimate for the empirical *D* value and its 95% confidence interval to 0. If 0 is not within the 95% confidence interval, we reject the null hypothesis. Evidence of either overdispersion (*D* > 0) or underdispersion (*D* < 0) allows us to reject the null hypothesis. An example of an empirical *D* estimate (black line), 95% confidence intervals for that estimate (dashed lines) and the null hypothesis (red line) can be seen in Figure 2.

Analysis

16S rDNA sequencing data from Koenig et al. (Koenig *et al.*, 2011) were downloaded from the NCBI Short Read Archive (SRA) website (http://www.ncbi.nlm.nih.gov/sra) along with their metadata. These data are a time-series of fecal bacterial communities from an infant subject, over the first 500 days of life. QIIME (Caporaso *et al.*, 2010) was used to trim primer regions from these data. Clustering was performed using the unoise3 pipeline (Edgar, 2016); sequences were de-replicated at 100% identity using vsearch (Rognes *et al.*, 2016), zOTU centroid sequences were picked and chimeric sequences were removed using unoise3 (Edgar, 2016), then all sequences were mapped onto zOTU seeds to create a zOTU table using vsearch. zOTU stands for “zero-radius operational taxonomic unit” (Edgar, 2016). Unlike traditional *de novo* clustered OTUs, zOTUs are exact sequence variants (ESVs) which are consistent and easily comparable across data sets much like a closed-reference OTU approach (Callahan *et al.*, 2017). However, ESVs are superior to closed-reference OTUs in that they still allow for the detection of novel diversity (Callahan *et al.*, 2017). The SINA aligner (Pruesse *et al.*, 2012) was used to align zOTU centroid sequences to the SILVA SSU Ref 128 database (available from https://www.arb-silva.de/download/arb-files/). We then used FastTree (Price *et al.*, 2010) to build a phylogenetic tree.

The resulting zOTU table was rarefied to 1000 sequences per sample, and samples with fewer sequences were excluded. The last five time points were excluded as well because they were sampled at a much lower temporal resolution. This left 52 time points spread over the first 469 days of the infant subject's life. The zOTU table was then split into two zOTU tables, one for time points before the infant started using baby formula, and one for those after. The “pre-formula” zOTU table contained ages 4 days through 146, and the “post-formula” zOTU table contained ages 161 days through 469 days. Each zOTU table was used to run our model as described above using 500 *D* values (Equation 1), ranging from underdispersed (*D* = -1) to overdispersed (*D* = 0.5), using zOTUs in lieu of species. zOTUs with zero phylogenetic distance between them were combined, because these zOTUs were uninformative for our statistical model (Equation 1; zero raised to a negative exponent is undefined). The 500 resulting phylodiversity accumulation curves were compared to the empirical curve using the KS-statistic as described above. Minimization of the resulting KS-statistic values was used to determine which value D best fit the data using the R package SiZer (Sonderegger, 2012), which fit a segmented regression (“broken stick”) model to the KS distances. 1000 bootstrap replicates of the segmented regression generated 95% confidence intervals for the D parameter estimates.

Sequence data from Caporaso et al. (Caporaso *et al.*, 2011) were downloaded from the MG-RAST database (http://metagenomics.anl.gov/). These are time-series data from one adult male subject and one adult female subject, over a period of several hundred days, across multiple sample sites (feces, both palms, tongue). Time points were excluded which did not have sequence data for each of the 8 environments (L hand, R hand, mouth, and feces of the male and female subjects), and rarefied to 5000 sequences per sample. This left 107 time-points, ranging from day 1 to day 185. Analysis for each environment within the data set was carried out as described above, except raw sequences were trimmed to a length of 91 bp after the end of the forward PCR primer site in order to ensure that all raw sequences spanned the same region of the 16S rRNA gene. 91 bp was chosen as a length cutoff in order to keep 95% of the sequence data (5% of sequences were discarded because they were shorter).

Code and data accessibility

R code and data to replicate our analysis, or to perform a similar analysis on other data, are available at <https://figshare.com/s/922b268891f1945c1944> (temporary private link, please do not share until publication).

Results

Figure 1 shows that by varying *D*, we changed the rate at which phylodiversity is added to surrogate (i.e. resampled) microbial communities over time. Compared to the neutral model where *D =* 0, higher *D* values result in phylodiversity accumulating quickly, since in the overdispersed model, species that contribute more phylodiversity are preferentially sampled. Conversely, lower *D* values result in phylodiversity accumulating slowly, since in the underdispersed model, species that contribute less phylodiversity (since they are very similar to species that are already present) are preferentially sampled. These results show that the *D* parameter in our model does in fact correspond to over- and underdispersion relative to the neutral model.

Results from “moving pictures” data

All eight time-series from adult microbiomes (Caporaso *et al.*, 2011) how phylogenetic underdispersion of first-time arrivals, since their D estimates were negative (Figure 3) and their 95% confidence intervals did not overlap with 0 (neutral model). This means that when a zOTU was observed for the first time in one of these communities, it was more likely to be phylogenetically similar to a zOTU that had previously arrived in that community. For both the male and female subject, *D* estimates were highest (closer to neutral) in the tongues, and lowest (more underdispersed) in the feces. Palm time-series had *D* estimates that were between those of the feces and tongues, but were consistent across all 4 palm data sets. *D* estimates for tongue time series were lower in the female subject than in the male subject.

Results from infant gut data

The phylodiversity accumulation pattern in the infant gut microbiome showed a sharp increase in phylodiversity after day 161 (Figure 4), the same date that the infant began consuming baby formula. This suggests that baby formula changed the phylogenetic colonization patterns of the developing infant gut. We analyzed this data set as two separate time-series, one before formula use and one after, and both had negative *D* estimates with 95% confidence intervals that did not overlap with 0 (neutral model). The pre-formula time-series was more underdispersed than the post-formula time-series, and the 95% confidence intervals for the *D* parameter estimates did not overlap with each other, indicating a change in the phylogenetic dispersion of arrival order post-disturbance. Because the post-formula D estimate is higher than the pre-formula *D* estimate, arrivals were less underdispersed (closer to neutral model) after the subject began consuming formula.

Discussion

The world is a mosaic of communities undergoing assembly following different magnitudes and types of disturbance. If there are patterns or general rules for how diversity accumulates during community assembly or succession, and which taxa have higher probabilities of arriving, these rules can guide habitat restoration projects, help us better design probiotics for colonization, and better exploit disturbance as a tool for managing microbial systems related to human health and disease. We found that assembly during primary succession of the infant gut and during turnover of the microbial communities on the adult palms, tongue, and gut, follows a predictable pattern: OTUs are more likely to arrive for the first time if a close relative has already arrived.

This “nepotistic” pattern in arrivals suggests that traits are driving community assembly in these human environments. In our model, each arrival is a new branch of a phylogenetic tree (alpha diversity accumulation), and these branches are indicators of trait differences in OTUs. Although the phylogenetic tree we use here is constructed with 16S rDNA sequences, such phylogenies have been shown to track genomic differences in bacteria (Langille *et al.*, 2013; Zaneveld *et al.*, 2010). This suggests that the phylodiversity accumulation we model here is also trait accumulation. Indeed, in previous frameworks of phylogenetic over/underdispersion (Gerhold *et al.*, 2015; Letcher and Chazdon, 2012; Webb *et al.*, 2002; Webb, 2000), non-neutral assembly has been interpreted to mean that traits are under ecological selection (Nemergut *et al.*, 2013). If traits are not driving community assembly (Hubbell, 2001) or if the traits driving community assembly are largely horizontally transferred between taxa independent of their relatedness (as estimated by a 16S rDNA phylogeny), we would expect no phylogenetic signature in first-time arrivals. Instead, we observed very a strong phylogenetic signal in arrival order for all data sets we analyzed.

However, even if non-neutral assembly is trait-based, selection on traits may not occur within the host environment. An alternative explanation for the underdispersion we observed is that selection is external to the host environment (i.e. selection occurs within the neighboring species pool from which emigration occurs), causing change in the community entering the host to already be underdispersed. Similarly, phylogenetic dispersion has been unable to distinguish between selection and differences in migration rates (Emerson and Gillespie, 2008), so a pre-underdispersed community entering the host is a plausible mechanism for phylogenetic underdispersion of arrivals. But selection of microbial communities within the host has been shown by multiple studies (Peterfreund *et al.*, 2012; David *et al.*, 2014; Kennedy *et al.*, 2016), so we think that trait-based selection within the host is a more likely scenario.

The burst of newly observed phylogenetic diversity in the infant gut microbiome after the subject began consuming baby formula (Figure 4) indicates that this particular disturbance resulted in a re-assembly of a community that was approaching island-biogeographic equilibrium (MacArthur and Wilson, 1967; Jean *et al.*, 2016). The pre-formula community had reached a state where phylogenetic diversity was not being added (Figure 4), but post-formula, there was a significant jump in novel phylogenetic diversity arriving in the community. After a few months post-formula, the community again moved toward a state where novel diversity was no longer being added.

Our model shows that both of these time series showed underdispersion (Figure 5) as did each of the 8 data sets from adult microbiomes (Figure 3), which suggests that microbial systems, as surveyed through 16S rDNA sequencing, have underdispersed phylodiversity accumulation over time. If true, our findings have major implications for the management and restoration of human-associated microbial systems (Shooner *et al.*, 2015), in particular for probiotic development and remediation of pathological microbial communities. Without any other information, our finding of consistent phylogenetic under-dispersion in arrivals suggests that probiotics for sustained colonization of the human gut should be close relatives to the microbes already present. Indeed, recent research has shown that for fecal transplants, donor strains are able to integrate into the recipient's gut community when a conspecific strain is already present, but novel donor strains are unlikely to successfully integrate into the recipient (Li *et al.*, 2016). Different body sites - as we saw with the skin – may have qualitatively similar patterns of underdispersion, yet quantitatively different magnitudes of this effect. Thus the efficacy of an engineered probiotic based on similarity to organisms already present in the community for which it was engineered may largely depend on the body site for which it's intended.

Microbial communities provide a unique opportunity to study community assembly in primary and secondary succession. In addition to standard cross-sectional studies of communities of different ages or successional stages, the short timescales of microbial community dynamics allows longitudinal studies of community assembly over manageable time frames (David *et al.*, 2014; Kennedy *et al.*, 2016). Microbial communities allow large sample sizes, longitudinal studies, and experimental manipulations that enable us to identify general rules and statistical patterns of community assembly. The model presented here makes use of such data, and to facilitate further discovery both in the human microbiome and in other environments, we have made our R code available at FigShare: <https://figshare.com/s/922b268891f1945c1944> (temporary private link, please do not share until publication).

Conflict of interest statement

The authors declare that no conflict of interest exists.

Acknowledgements

The authors thank D.R. Nemergut for her help and support, and also thank J.P. O'Dwyer, P. Sommers, E.M. Gendron, A. Solon, and S. Sauce for many helpful discussions. Funding was provided by an NSF grant for studying microbial community assembly following disturbance (DEB-1258160).

References

Bengtsson J, Nilsson SG, Franc A, Menozzi P. (2000). Biodiversity, disturbances, ecosystem function and management of European forests. *For Ecol Manage* **132**: 39–50.

Callahan BJ, McMurdie PJ, Holmes SP. (2017). Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J*. e-pub ahead of print, doi: 10.1038/ismej.2017.119.

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, *et al.* (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**: 335–6.

Caporaso JG, Lauber CL, Costello EK, Berg-Lyons D, Gonzalez A, Stombaugh J, *et al.* (2011). Moving pictures of the human microbiome. *Genome Biol* **12**: R50.

Darcy JL, King AJ, Gendron EMS, Schmidt SK. (2017). Spatial autocorrelation of microbial communities atop a debris-covered glacier is evidence of a supraglacial chronosequence. *FEMS Microbiol Ecol* **93**: 7287–90.

David LA, Materna AC, Friedman J, Campos-Baptista MI, Blackburn MC, Perrotta A, *et al.* (2014). Host lifestyle affects human microbiota on daily timescales. *Genome Biol* **15**: R89.

Edgar RC. (2016). UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing. *bioRxiv*. http://www.biorxiv.org/content/early/2016/10/15/081257 (Accessed July 26, 2017).

Emerson BC, Gillespie RG. (2008). Phylogenetic analysis of community assembly and structure over space and time. *Trends Ecol Evol* **23**: 619–30.

Faith DP. (1992). Conservation evaluation and phylogenetic diversity. *Biol Conserv* **61**: 1–10.

Gerhold P, Cahill JF, Winter M, Bartish I V., Prinzing A. (2015). Phylogenetic patterns are not proxies of community assembly mechanisms (they are far better) Venail P (ed). *Funct Ecol* **29**: 600–614.

Goberna M, Navarro-Cano JA, Valiente-Banuet A, García C, Verdú M. (2014). Abiotic stress tolerance and competition-related traits underlie phylogenetic clustering in soil bacterial communities. *Ecol Lett* **17**: 1191–201.

Hubbell SP. (2001). The Unified Neutral Theory of Biodiversity and Biogeography (MPB-32). Princeton University Press.

Jean K, Burnside WR, Carlson L, Smith K, Guégan J-F. (2016). An equilibrium theory signature in the island biogeography of human parasites and pathogens. *Glob Ecol Biogeogr* **25**: 107–116.

Kennedy RC, Fling RR, Robeson MS, Saxton AM, Donnell RL, Darcy JL, *et al.* (2016). Temporal Development of Gut Microbiota in Triclocarban Exposed Pregnant and Neonatal Rats. *Sci Rep* **6**: 33430.

Koenig JE, Spor A, Scalfone N, Fricker AD, Stombaugh J, Knight R, *et al.* (2011). Succession of microbial consortia in the developing infant gut microbiome. *Proc Natl Acad Sci U S A* **108 Suppl**: 4578–85.

Langille MGI, Zaneveld J, Caporaso JG, McDonald D, Knights D, Reyes JA, *et al.* (2013). Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat Biotechnol* **31**: 814–821.

Letcher SG, Chazdon RL. (2012). Life History Traits of Lianas During Tropical Forest Succession. *Biotropica* **44**: 720–727.

Li SS, Zhu A, Benes V, Costea PI, Hercog R, Hildebrand F, *et al.* (2016). Durable coexistence of donor and recipient strains after fecal microbiota transplantation. *Science* **352**: 586–9.

Lozupone C, Knight R. (2005). UniFrac : a New Phylogenetic Method for Comparing Microbial Communities UniFrac : a New Phylogenetic Method for Comparing Microbial Communities. *Appl Environ Microbiol* **71**: 8228–8235.

MacArthur RH, Wilson EO. (1967). Theory of Island Biogeography. Princeton University Press.

Martiny JBH, Jones SE, Lennon JT, Martiny AC. (2015). Microbiomes in light of traits: A phylogenetic perspective. *Science* **350**: aac9323.

Massey FJ. (1951). The Kolmogorov-Smirnov test for goodness of fit. *Joirnal Am Stat Assoc* **56**: 68–78.

Matsen IV FA, Evans SN, Gilks W, Ghodsi M, Kingsford C. (2013). Edge Principal Components and Squash Clustering: Using the Special Structure of Phylogenetic Placement Data for Sample Comparison Moustafa A (ed). *PLoS One* **8**: e56859.

Nemergut DR, Knelman JE, Ferrenberg S, Bilinski T, Melbourne B, Jiang L, *et al.* (2016). Decreases in average bacterial community rRNA operon copy number during succession. *ISME J* **10**: 1147–1156.

Nemergut DR, Schmidt SK, Fukami T, O’Neill SP, Bilinski TM, Stanish LF, *et al.* (2013). Patterns and processes of microbial community assembly. *Microbiol Mol Biol Rev* **77**: 342–56.

O’Dwyer JP, Kembel SW, Green JL. (2012). Phylogenetic diversity theory sheds light on the structure of microbial communities. *PLoS Comput Biol* **8**: e1002832.

Palmer MA, Ambrose RF, Poff NL. (1997). Ecological Theory and Community Restoration Ecology. *Restor Ecol* **5**: 291–300.

Peterfreund GL, Vandivier LE, Sinha R, Marozsan AJ, Olson WC, Zhu J, *et al.* (2012). Succession in the gut microbiome following antibiotic and antibody therapies for Clostridium difficile. *PLoS One* **7**: e46966.

Price MN, Dehal PS, Arkin AP. (2010). FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments Poon AFY (ed). *PLoS One* **5**: e9490.

Pruesse E, Peplies J, Glöckner FO. (2012). SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* **28**: 1823–1829.

Richards SA, Possingham HP, Tizard J. (1999). Optimal fire management for maintaining community diversity. *Ecol Appl* **9**: 880–892.

Rognes T, Flouri T, Nichols B, Quince C, Mahé F. (2016). VSEARCH: a versatile open source tool for metagenomics. *PeerJ Prepr* **4**: e2409v1.

Shooner S, Chisholm C, Davies TJ. (2015). The phylogenetics of succession can guide restoration: an example from abandoned mine sites in the subarctic Kardol P (ed). *J Appl Ecol* **52**: 1509–1517.

Sonderegger D. (2012). SiZer: Significant Zero Crossings. https://cran.r-project.org/package=SiZer.

Temperton VM. (2004). Assembly Rules and Restoration Ecology: Bridging the Gap Between Theory and Practice. Island Press.

Violle C, Navas ML, Vile D, Kazakou E, Fortunel C, Hummel I, *et al.* (2007). Let the concept of trait be functional! *Oikos* **116**: 882–892.

Washburne AD, Silverman JD, Leff JW, Bennett DJ, Darcy JL, Mukherjee S, *et al.* (2017). Phylogenetic factorization of compositional data yields lineage-level associations in microbiome datasets. *PeerJ* **5**: e2969.

Webb CO. (2000). Exploring the Phylogenetic Structure of Ecological Communities: An Example for Rain Forest Trees. *Am Nat* **156**: 145–155.

Webb CO, Ackerly DD, McPeek MA, Donoghue MJ. (2002). Phylogenies and Community Ecology. *Annu Rev Ecol Syst* **33**: 475–505.

Wiens JJ, Ackerly DD, Allen AP, Anacker BL, Buckley LB, Cornell H V., *et al.* (2010). Niche conservatism as an emerging principle in ecology and conservation biology. *Ecol Lett* **13**: 1310–1324.

Zaneveld JR, Lozupone C, Gordon JI, Knight R. (2010). Ribosomal RNA diversity predicts genome diversity in gut bacteria and their relatives. *Nucleic Acids Res* **38**: 3869–3879.

Figure Captions

**Figure 1**: Phylodiversity accumulation in the female left palm. The dashed line is the empirical phylodiversity accumulation curve, meaning that each time point's phylodiversity value is the cumulative sum of all branch lengths observed up to that time point (Faith 1992). This curve has been scaled to the 0-1 scale using Equation 2. The colored lines are 500 surrogate (i.e. resampled) phylodiversity curves with different *D* values (Equation 1). These lines are color-coded by their *D* value (see key at right). The empirical model (dashed) is below the neutral model (purple), signifying underdispersion in the order of first-time arrivals.

**Figure 2**: Model fitting and hypothesis testing for the female left palm. Each surrogate (i.e. resampled) data set from Figure 1 is compared to empirical data using the KS statistic, with a lower KS statistic indicating a closer fit. Each dot on this figure is colored according to the surrogate’s *D* value (Equation 1), and corresponds to its line color on Figure 1. The black solid vertical line is the estimate for the best-fit model, and the vertical dashed lines are the 95% confidence intervals for that estimate. The red solid vertical line is placed at *D*=0, which is the neutral model.

**Figure 3**: Dispersion parameter (*D*) estimates for Caporaso et al. (2011) data sets. All environments analyzed exhibited underdispersion. For both the female (left) and male (right) subjects, the feces (brown) had the lowest *D* estimates (most phylogenetically underdispersed), and the tongues (red) had the highest *D* estimates, although they were still underdispersed compared to the neutral model. For both subjects, the *D* estimates for the palm (tan) communities were similar.

**Figure 4**: Empirical phylodiversity accumulation in the infant gut microbiome. Phylodiversity increases sharply after day 161 of the infant’s life, then plateaus. This timing coincides with the day the subject began consuming baby formula. The times of sampling points are shown as vertical blue lines below the X-axis.

**Figure 5**: Dispersion parameter (*D*) estimates in the infant gut, pre- and post-formula. Formula use began on day 161, thus the first 160 days of the subject's life were analyzed separately. Community assembly was underdispersed both pre- and post-formula, but moreso pre-formula.