Ecological Stability Emerges at the Level of Strains in the Human Gut Microbiome

Richard Wolff [1], William Shoemaker[1], Nandita Garud[1,2,*],

1 Department of Ecology and Evolutionary, UCLA
2 Department of Human Genetics, UCLA

*** ngarud@g.ucla.edu**

Abstract

The human gut microbiome is a complex community that harbors substantial ecological diversity at the species level, as well as at the strain level within species. In healthy hosts, species abundance fluctuations in the microbiome community are thought to be stable, and these fluctuations can be described by macroecological laws. However, it is less clear how strain abundances change over time. An open question is whether individual strains behave like species themselves, exhibiting stability and following the macroecological relationships known to hold at the species level, or whether strains have different dynamics, perhaps due to the relatively close phylogenetic relatedness of co-colonizing lineages. In this study, we sought to characterize the typical strain-level dynamics of the healthy human gut microbiome on timescales ranging from days to years. We show that genetic diversity within almost all species is stationary, tending towards a long-term typical value within hosts over time scales of several years, despite fluctuations on shorter timescales. Moreover, the abundance fluctuations of strains can be sufficiently described by a stochastic logistic model (SLM) – a model previously used to describe abundance fluctuations among species around a fixed carrying capacity – in the vast majority of cases, suggesting that strains are dynamically stable. Lastly, we find that strain abundances follow the same macroecological laws known to hold at the species level. Together, our results suggest that macroecological properties of the human gut microbiome, including its stability, emerge at the level of strains.

Introduction

The human gut microbiome is composed of a diverse array of microbial species. While a typical gut microbial species harbors considerable genetic variation both within and across hosts, the ecological and functional consequences of this diversity remain largely unknown. Although recent efforts have begun to characterize how genotypic diversity changes within healthy hosts over months to years, these trends are not, at present, quantified on the short time frames most relevant for microbial ecology – that is, over periods of days [5, 10, 15, 24, 25, 28, 30]. Understanding the typical scale of daily fluctuations in genetic variation is critical to assessing both the long-term stability of the genetic composition of the gut

microbiome, as well the effects of occasional large perturbations resulting from changes in host diet, medication, travel, illness, and other factors.

Two kinds of processes drive within-host changes in genetic variation in the human gut. First, there is the evolutionary modification of resident lineages, which can result in small numbers ($O(1) - O(10)$) of single nucleotide variants (SNVs) sweeping from low to high frequency on timescales of weeks to months. Second, fluctuations in the abundance of strains, which have a typical nucleotide divergence of 1%, can result in large numbers ($\sim O(10^4)$) of SNV frequency changes over time [10]..The most dramatic manifestation of this second process is strain replacement, when one strain of a species invades and drives the resident to extinction, though such events are infrequent over ~ 1 year timescales [24, 25]. Thus, strain abundance fluctuations have several orders of magnitude greater impact on intraspecies genetic variation over time than evolutionary changes.

Prior analyses have demonstrated that the majority of strains persist within hosts over a period of at least several years [5, 15, 24, 25]. Moreover, strains can be resilient to even large perturbations of the gut community, such as antibiotics [23] and fecal microbiome transplants (FMT) [11]. Interestingly, strains in the gut microbiome frequently co-exist with a handful of other strains belonging to the same species. This "oligo-colonization" model – in which a species is made up of $\sim 1 - 4$ strains [10, 54] – has been observed in a number of other host-associated microbiota, both at different human body sites [42] as well as in other organisms [8, 16].

The coexistence of multiple strains within an individual gut for periods of years contrasts starkly with the rapid evolution known to occur regularly at individual SNVs. This suggests that while competitive exclusion and directional selection may frequently prevail among closely related lineages, highly diverged lineages are generally subject to different eco-evolutionary forces. That is, while SNVs are known to frequently arise and fix within populations, strains, which are far more genetically diverged, seem much less likely to drive other strains extinct.

To understand the typical abundance fluctuations of strains in the microbiome, we leverage concepts from macroecology. Macroecology focuses on elucidating the statistical and ecological properties of communities. There is an increasing body of work which demonstrates that patterns of microbial species abundance and diversity follow macroecological laws across disparate environments, including the human gut [1, 6, 14, 45]. Surprisingly, many of these macroecological laws can be recapitulated through intuitive ecological models containing few if any free parameters [1, 6, 44]. Among these successful models is the Stochastic Logistic Model (SLM), which describes the dynamics of a population experiencing rapid environmental fluctuations around a fixed carrying capacity. Whether the strains making up a community exhibit regular, statistically quantifiable dynamics, and if so, whether these dynamics can be explained using simple models, are fundamentally macroecological questions.

In this study, we examine whether the macroecological dynamics observed at the species level hold at the strain level. We investigate the temporal dynamics and macroecology of strains in a densely sampled cohort of four healthy, adult hosts (*am*, *an*, *ao*, and *ae*) from a previously published data set [21]. We find that the vast majority of strains in the human gut are stable in these healthy hosts on ~ 1 year time scales, and that they exhibit some of the same macroecological patterns as species. We approached the problem of intraspecies stability first by quantifying the change in genetic polymorphism through time, and showed that levels of intra-species genetic variation (as measured by the nucleotide diversity π) fluctuate around long-term steady state values. Next, we connected the lack of directionality exhibited by genetic diversity through time with an underlying model of stable population dynamics among—the SLM, first applied

by [1, 6] to characterize microbial diversity at the species level. We find that this model provides a sufficient description of strain dynamics in almost all cases, and that it fails in the only case of a clear strain "replacement" in our cohort. Lastly, we demonstrated that several macroecological laws initially shown to hold at the species level also hold among strains. Together, this work indicates that the stability of the gut microbiome emerges at the level of individual strains.

Results

Stability of intraspecies diversity

Fluctuations in intraspecies genetic diversity reflect both the population genetic forces affecting lineages within a population and the ecological forces affecting the relative abundances of different strains. To investigate these forces, we first analyzed the nucleotide diversity of species in our cohort.

An illustrative example of these temporal dynamics is provided by *Bacteroides vulgatus*, the most abundant species in host *am*. In **Figure 1A**, nucleotide diversity π for this population is plotted over the two-year sampling period. While π undergoes relatively large fluctuations (varying by more than a factor of two), there is no apparent trend for diversity to systematically either increase or decrease. Rather, π seems to fluctuate around a characteristic value of about 1.5×10^{-3}, with periods of elevated or decreased diversity followed by a return to the "steady" state. Across all species examined, nucleotide diversity typically ranged from approximately $10^{-4} - 10^{-2}$ per basepair.

We implemented a permutation test to determine quantitatively if levels of diversity were indeed constant through time for species in our data (Methods). Applying this test, we found that 55 of the 64 species for which SNPs could reliably be inferred (including *B. vulgatus* in *am*) showed no trend in π throughout the sampling period at a 5% significance level, confirming our initial qualitative assessment of the overall stationarity of genetic diversity (Supplementary Figure 1).

Levels of nucleotide diversity in many species in our data are inconsistent with the presence of a single strain, or of a group of lineages which diversified since entering the host. Using conservatively high estimates for per-site mutation rates, generation times, and time since colonization, the authors of [**?**] estimated that genetic polymorphism could reach values of at most about 10^{-3} per basepair if lineages diverged within a host. If however, a species is made up of multiple strains that accumulated mutations for many generations before colonizing the same gut community, nucleotide diversity can easily surpass this value.

The nucleotide diversity of *B. vulgatus* hovers just above this upper threshold, indicating that the species may be made up of multiple diverged strains. To assess this, we inferred the underlying strain structure *B. vulgatus* (**Figure 1B**). While accurate strain inference is inherently limited in cases where read depth is low, distinct strains can be confidently inferred when SNV data is available at many timepoints. Using a previously published algorithm which leverages the correlations in allele frequency trajectories between linked variants to detect strains in dense longitudinal data, we were able to separate out two distinct clusters of allele frequency trajectories, strongly indicating that *B. vulgatus* is oligo-colonized by a mixture of three strains (Supplementary Figure 2). As with π, the relative abundances of the strains seem centered around a fixed steady state, to which the strains return following transient increases or decreases in abundance. Indeed, variations in the relative abundances of the strains away from their steady state (e.g. around timepoint 100, when one strain rises markedly in abundance) correspond

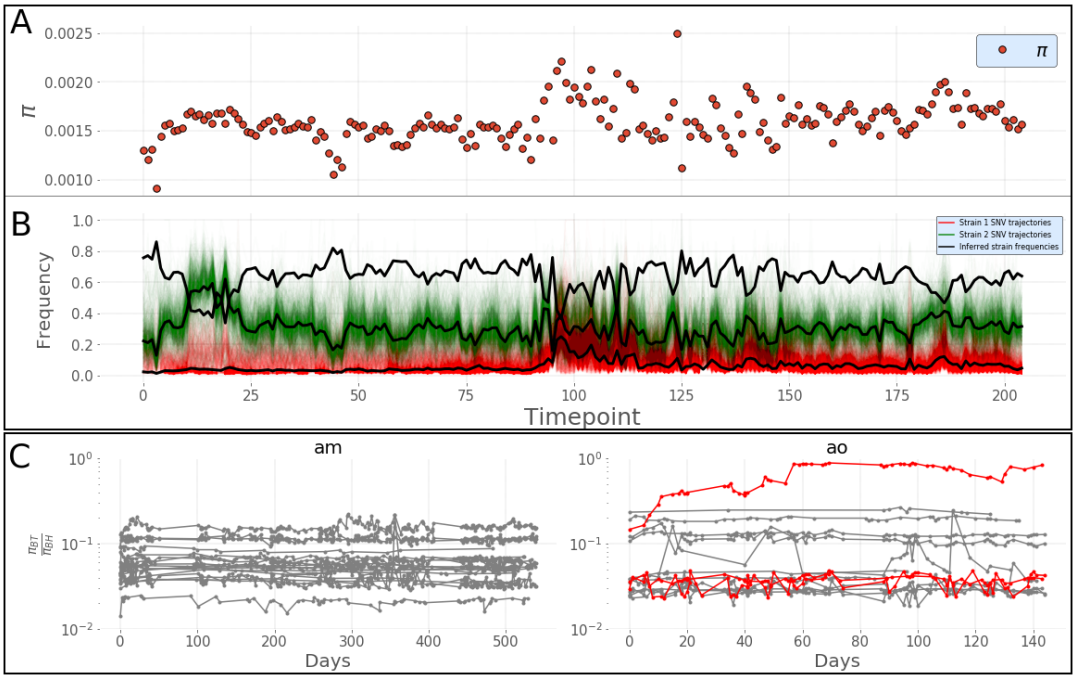


Figure 1. **A**) Nucleotide diversity π of *Bacteroides vulgatus* in host *am*. Short term fluctuations in diversity tend ultimately to revert to a long term average value. **B**) Results of strain inference for *B. vulgatus*. Each colored line represents the trajectory of a single phased allele, while the color itself denotes the strain cluster the allele was identified as belong to (the unphased allelic trajectories can be seen in Supplementary Figure 2.) Black lines are centroids of the allele trajectory clusters, and are an estimator for the true relative abundance of the strain. Fluctuations in strain frequency are mirrored by fluctuations in genetic diversity. Note the x-axis shows the ordinal timepoint number, not number of days. **C**) Trajectories of the ratio of π_{BT} to π_{BH} for each species in hosts *am* and *ao*. The bulk of trajectories are grey, indicating no significant trend across the sampling period; red trajectories are those which show a significant positive trend. π_{BT} approaches π_{BH} for *F. prausnitzii* in host *ao*, indicating a possible partial displacement.

exactly to perturbations in π. It is clear in this case that the changes in strain abundance are in fact the leading contribution to fluctuations in intraspecies genetic diversity.

Nucleotide diversity π can, in theory, remain relatively constant even when there are dramatic changes in the strain-level composition of a species—this might happen if, for instance, a species colonized by a single strain was replaced by another single strain from a different host. However, such changes would manifest in the genetic composition of the species changing dramatically, with the level of diversity between timepoints approaching or exceeding that between hosts, as unrelated hosts typically contain distinct, highly diverged sets of strains [24]. To understand how the genetic composition of species changes in time, we computed the nucleotide diversity between timepoints π_{BT}, normalizing this quantity by the average genetic diversity between hosts π_{BH} (Methods). We again conducted a permutation test on π_{BT} time series, and found significant temporal trends in diversity change in just four species. Intriguingly, three of these four species were found in a single host (*ao*). Among all species in all hosts, the ratio $\frac{\pi_{BT}}{\pi_{BH}}$ approached one only for *Faecalibacterium prausnitzii* in host *ao* (**Figure 1C**, right, in red).

Stochastic logistic model

Next, we sought to determine if the temporal dynamics of strains could be captured using a naive model. Recent work in microbial ecology has repeatedly demonstrated the power of such models to reproduce qualitative and quantitative features of natural microbial community dynamics [1, 6, 43, 44]. We show that the stochastic logistic model (SLM), a minimal model itself requiring the fit of no free parameters, is a good fit for nearly all the strain time series in our cohort.

We first obtained time series of strain abundances. To do so, we determined the relative frequencies of strains using the technique described above, and then multiplied these by the frequencies of the species to which they belong, excluding species and samples with low abundance (Methods).

If x_i is the abundance of strain i which follows an SLM, then:

$$\frac{dx_i}{dt} = \frac{x_i}{\tau_i}\left(1 - \frac{x_i}{K_i}\right) + \sqrt{\frac{\sigma_i}{\tau_i}} x_i \eta(t)$$

here τ_i is the intrinsic growth rate of the strain and $\eta(t)$ is a Brownian noise term. Under the assumptions of the model, each population has a long-term carrying capacity K_i, and temporal fluctuations in abundance around this value are driven entirely by environmental noise with amplitude determined by σ_i. Populations may experience large fluctuations in abundance over short timescales, and may even be temporarily found far from their long-term average value, but these fluctuations will be transient. Over long timescales, the stationary distribution in abundances predicted by the SLM is the following Gamma distribution [1]:

$$\rho(x_i) = \frac{1}{\Gamma(2\sigma_i^{-1} - 1)} \left(\frac{2}{K_i \sigma_i}\right)^{2\sigma_i^{-1} - 1} x^{2\sigma_i^{-1} - 2} \exp\left(-\frac{2}{K_i \sigma_i} x\right) \tag{2}$$

In Figure 2A, simulations from the SLM are compared with the actual time series of two strains. The qualitative agreement between data and model is evident, and is further reflected in the close match of the empirical distributions of abundances over the whole sampling period with the predicted stationary Gamma distribution.

To assess quantitatively whether the time series of populations in our cohort could be adequately described by an SLM, we developed and implemented a goodness-of-fit test (Methods). The test determines whether the transitions between subsequent timepoints are consistent with an SLM.

The SLM fit the data in the overwhelming majority of cases: 94% across all hosts combined (Figure 2B), with similar percentages of strains fitting the model in each host. We emphasize that this test did not require the fit of any free parameters, as the parameters of the SLM associated with each population were estimated only from the mean and variance in its abundance (Methods). The agreement of data and model is thus unlikely to be an artifact of model over-specification.

The SLM was rejected in some instances. Notably, the model was rejected for one of the two strains of *F. prausnitzii* in *ao*, the same species which was shown above to have experienced a large, directional change in its genetic diversity and composition. This strain experienced a rapid decrease in abundance midway through the sampling period, and thereafter never fully recovered to its previous state (Supplemental Figure 3). The other strain fit the SLM. We conclude that ecological processes happening at the level of strains drove the observed change in the genetic composition of this species; and in particular, the change can be attributed to a dramatic shift in abundance of just one of the two populations initially

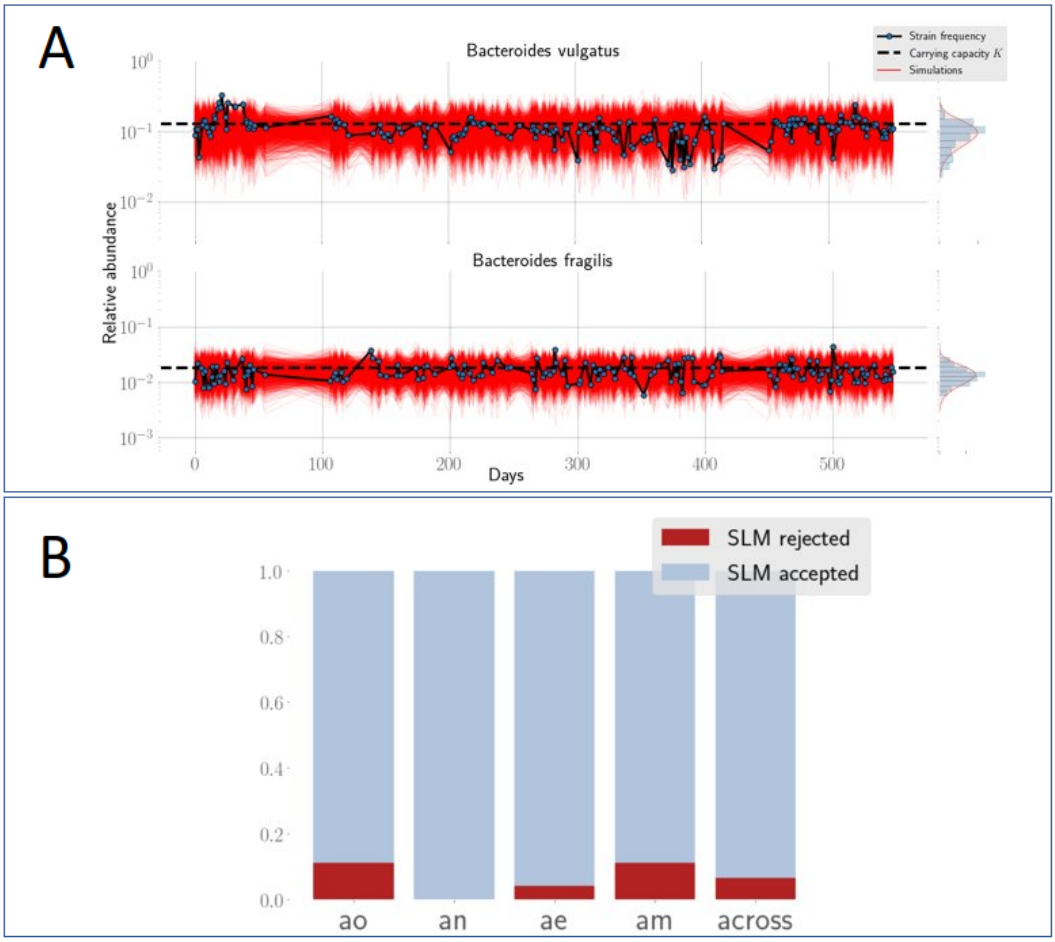


Figure 2. A) Simulations of the SLM associated with each strain, in red, capture the behavior of the real data at multiple time scales. Short-term increases and decreases are followed by returns to the long-term carrying capacity K. At far right, in blue, are histograms of the strains' abundance, and overlaid in red is the Gamma stationary distribution of abundances predicted by the SLM, described in equation 1. **B)** The percentage (y-axis) of strains for hosts *am*,*ae*,*ao*, and *an* passing the SLM goodness of fit test. A high proportion of strains pass the test—89% in *ao*, 100% in *an*, 96% in *ae*, and 89% in *am*. Across all four hosts, 94% of strains pass the test.

present. Despite this shift, both strains were detectable at all time points. Though the SLM was rejected in several other populations, *F. prausnitzii* was the only species which simultaneously experienced a statistically significant change in genetic diversity.

Together, these results indicate that the great majority of strains in the gut microbial community fluctuate around fixed average carrying capacities for periods of years, at least.

Macroecology of strains

While we have already seen that individual strains tend to be well described by an SLM, we show in the following section that the size of fluctuations across strains are strongly constrained. In many kinds of

microbial ecosystems, including the human gut, species have been shown to broadly obey macroecological laws [1, 45]. We show here that these patterns equally well characterize patterns of variation in the abundance of strains across our cohort.

The first pattern examined is a power law scaling between the mean and variance of abundance, known in ecology as Taylor's Law, and can be stated:

$$\sigma^2_{x_i} \propto \langle x_i \rangle^{\alpha} \tag{3}$$

where $\langle x_i \rangle$ and $\sigma^2_{x_i}$ are the mean and variance of population x_i, respectively, and α is the scaling exponent of the power law. In communities where the relative scale of fluctuations is independent of population size—constant per-capita fluctuations—α will equal 2 [45]. We observed a Taylor's Law scaling with an exponent of $\alpha = 1.63$ among all strains (**Figure 3A**), closely mirroring previous findings at the species level [45].

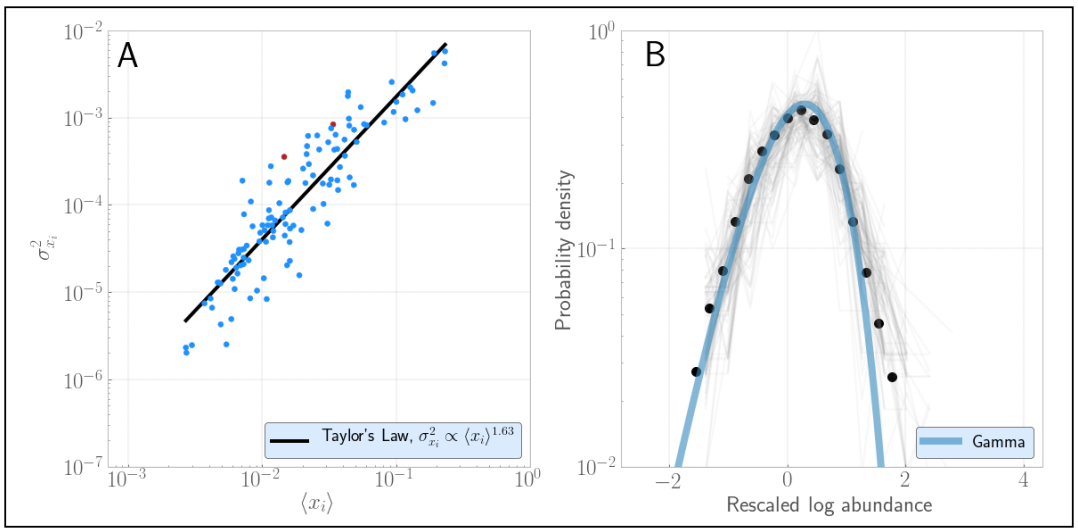


Figure 3. A) Strains obey Taylor's Law with exponent $\alpha = 1.63$. All strains are colored blue except the previously discussed strains belonging to *Faecalibacterium prausnitzii* in ao, which are colored red. **B)** The Gamma distribution describes the fluctuations in abundances of strains through time. To enable comparison across strains which have mean abundances ranging across several orders of magnitude, the AFD of each individual strain (grey lines) has been rescaled to have mean zero and unit variance. The black dots show the mean probability of a given rescaled abundance across strains.

The next pattern considered is the Gamma abundance fluctuation distribution (AFD), the overall distribution of abundances of a population through time. As discussed above, a population governed by stochastic logistic dynamics will tends towards a Gamma distribution of abundances over long timescales. Given the generally excellent fit of the SLM to the population time series, the abundances of strains might generically be expected to each individually follow a Gamma distribution. In (**Figure 3B**), we see that the distributions of strain abundances are indeed well described by a Gamma distribution; however, this Gamma is also conserved across strains—that is, all strains approximately lie (up to a scaling factor) along the same Gamma. Recalling that each SLM is uniquely determined by the mean and variance of population, it is apparent that the collapse of the AFDs to a single Gamma is in fact a consequence of the strong constraint Taylor's Law places on these quantities across strains.

The observed macroecological patterns continue to hold even when limiting our attention only to strains for which another strain of the same species is present (Supplemental Figure 4). Thus, two very broad macroecological laws observed at the species level in the human gut are also observed among strains, suggesting that the biotic and abiotic factors driving these patterns may act at the level of strains.

Discussion

In this study, we sought to characterize the typical within-species population dynamics in the human gut microbiome. Previous efforts have demonstrated that the genetic diversity within a host persists for multiple years for most species [24, 25]. We build on this result by demonstrating that intraspecies diversity tends to fluctuate around a long-term average value within a typical host on a time scale of several years. We show, crucially, that the abundance fluctuations of the vast majority of strains can be sufficiently described by a stochastic logistic model (SLM) of growth, a model which also recapitulates fluctuations at the species level [1, 6]. Furthermore, empirical patterns of strain abundance follow macroecological laws which have been previously demonstrated to hold at the species level [1, 6]. Together, our results suggest that the macroecological dynamics exhibited by species are recapitulated at the strain level.

While the SLM was able to sufficiently describe strain dynamics for the vast majority of strains across species, its success is not universal. One strain of *F. prausnitzii* is a noted exception, as its abundance decreased dramatically in abundance in a way that could not be explained by the stationary dynamics of the SLM. The directional change in abundance of this strain throws into relief the stability of the majority of the other strains—it is an exception that proves the rule, illustrating that strain dynamics might be quite different across strains, and that these differences can be detected by our test of the SLM. This example also brings to light an interesting tension in the interpretation of our results. Under the SLM, strains are expected to persist indefinitely. However, over the course of decades, much of the strain content of the adult gut is replaced [10, 15], suggesting that there is an additional timescale which is relevant for strain replacement. One hypothesis is that this timescale reflects a waiting time for large environmental perturbations, such as antibiotics [23, 32] or bowel cleanse [49], but this is just one of many hypotheses. Indeed, this hypothesis is partially challenged by [23], where the strain content of an adult gut was initially perturbed during a course of antibiotics, but ultimately recovered to its pretreatment state. This antibiotics study is a powerful demonstration of the stability of strains as ecological units, even in the face of large perturbations. Testing the possible explanations for the apparent discrepancy between years and decades-long population dynamics is an important problem to address with broad cohorts and extended timescales of observation.

Beyond the success of the SLM as an ecological model of strain dynamics in our cohort, the existence of many linked mutations segregating at intermediate frequencies across multiple species is qualitatively inconsistent with most standard population genetic models of microbial evolution [18]—particularly, models emphasizing directional selection or neutral evolution. However, the stability of both total genetic diversity and strain abundance belies the fact that SNVs likely continue to arise and fix within these populations even on the timescales examined here. While strain dynamics may often be suitably described by a time invariant model, these populations are not genetically static. Recent work has shown that variants arise and fix within populations in the gut microbiome regularly over months to years [10, 15, 30]

How evolution impacts the ecological dynamics of strains and how in turn these ecological dynamics constrain and channel evolution, is an active area of research [18]. In the context of the SLM, these

eco-evolutionary feedbacks can viewed as tuning a strain's carrying capacity K, growth rate $\frac{1}{\tau}$, and sensitivity of the growth rate to environmental perturbation σ. Naively, it is expected that evolution would tend to increase carrying capacity while minimizing the sensitivity of growth to abiotic fluctuations, but evolutionary modifications driving changes in one quantity may affect the other. The observed power-law scaling between the mean and variance in abundance (Taylor's Law) is, in essence, a constraint on K given σ, and vice versa. The SLM thus not only describes ecological dynamics, but also, in conjunction with the empirical observation of macroecological laws, provides a useful framework for investigating the ecological effects of adaptation.

How and why closely related strains stably coexist in the human gut is one of the central biological questions raised by these results. Spatial segregation between strains, perhaps occupying different colonic crypts, could underlie the observed pattern of strain coexistence [20, 27], much as it does among *Cutibacterium acnes* strains inhabiting different pores on the facial microbiome [42]. However, spatial structure is far from the only mechanism that can foster coexistence between strains. Coexisting strains have been reported in well-mixed laboratory evolution experiments [4, 8, 55]. In these experiments, strains coexist by finely partitioning some aspect of the abiotic environment or by engaging in ecological interactions (e.g. cross-feeding), or by some combination of both. Recent work has shown that consumer-resource models, which describe the flux of metabolites through a community and the growth of community members on these metabolites, can recapitulate a large number of species-level macroecological properties of microbial communities with only a small number of input parameters [29, 31, 44]. Investigating which of these scenarios promotes strain coexistence will be an interesting avenue for future research.

Finally, the success of the SLM at the strain and species level raises questions regarding which scale ought to be the focus of ecological investigations. The ambiguity surrounding the bacterial species concept is well known [34] and reasonable alternatives have been proposed [19], but operationally species are, nonetheless, the predominant focus of attention in microbial ecology. This focus is reasonable, as within-host strain structure is a comparatively recent discovery [10, 28, 54] and 16S rRNA sequencing provides an inexpensive high-throughput means to examine community dynamics through OTUs. Regardless, the recapitulation of species-level macroecological dynamics at the level of strains calls into question the disproportionate focus on species as the primary locus of attention in characterizing community structure and dynamics. Instead, it is reasonable to propose that for the human gut, and perhaps other microbial ecosystems, strains are an ecologically relevant unit.

Methods

Data and metagenomic pipeline

We analyzed shotgun metagenomic sequence data from a panel of stool samples from 4 healthy human subjects [21]. The four hosts examined—*ae, am, an* and *ao*—were sampled longitudinally over the course of between six months and two years, and none of the hosts experienced any disturbances such as antibiotics or bowel cleanse. We excluded one sample from host ae which appeared to be mislabelled and/or contaminated (Supplemental Figure 5). We used a reference-based approach to analyze metagenomic sequences, calling SNVs and gene content, as well as species abundances, using the software MIDAS [13].

Calculating diversity statistics

Nucleotide diversity π is a classical population genetic measure of polymorphism, representing the average number of pairwise difference between randomly chosen members of a population. To determine π, we used the estimator:

$$\pi = \frac{1}{|G|} \sum_{i=1}^{|G|} 2p_i(1 - p_i) \tag{4}$$

where p_i is the frequency of the reference allele at site i, and $|G|$ is the total number of sites in the genome. This quantity was calculated after first excluding sites with low read depth (< 5x), as reliable estimates of true allele frequency cannot be made for such sites.

Similarly, π_{BT}, the diversity between timepoints, was calculate

$$\pi_{BT} = \frac{1}{|G|} \sum_{i=1}^{|G|} (p_{1i}(1 - p_{2i}) + p_{2i}(1 - p_{1i})) \tag{5}$$

where p_{1i} is the frequency of the reference allele at site i in sample 1, and p_{2i} its frequency in sample 2.

Lastly, to determine π_{BH}, the diversity between hosts for a give species, we used shotgun data from the Human Microbiome Project (HMP), processed through the same metagenomic pipeline as above [17]. HMP was chosen as an due to the large number of samples (469, in total) and high coverage. π_{BH} was calculated as the mean pairwise diversity (Equation 5) between all pairs of samples of a species found in different hosts.

Permutation test

To assess whether diversity tended to systematically increase or decrease through time for species in our cohort, we performed a standard permutation test [51]. First, a linear regression was performed on β_i. The observed time series of π and π_{BT} were permuted with respect to temporal order 1000 times, and for each permutation, a linear regression was fit. The slopes of these regressions—$\beta_i^1, \beta_i^2, ..., \beta_i^{1000}$—are centered around 0, and form a null distribution for the true slope under the assumption that there are no long-term temporal trends in the data. We rejected the null hypothesis at a signficance level of 5%.

Strain inference

To infer strains, we used a recently published algorithm [23] developed specifically to detect strains in metagenomic timecourse data. At a high level, this algorithm identifies clusters of SNVs that have similar allele frequency trajectories across a longitudinal panel of samples, modulo binomial sampling noise at each timepoint. Such clusters are expected when alleles at different loci are linked on the same genetic background (and therefore have the same true frequency at any timepoint), but differ in their observed frequencies due to finite sampling. Once SNVs have been clustered, the centroid of each cluster of trajectories is taken to be an estimator of the underlying relative frequency of the strain.

SLM

To simulate the SLM, we used the Euler-Maruyama method:

$$X(t+\delta t) = X(t) + \frac{x(t)}{\tau_i}\left(1 - \frac{x(t)}{K}\right)\delta t + \sqrt{\frac{\sigma}{\tau}}x(t)Z_t\sqrt{\delta t} \tag{6}$$

where Z_t is a standard normal random variable. In simulations, we set $\delta t = \frac{1}{1000}$.

The SLM associated with population i depends on three parameters: K_i, σ_i, and τ_i. K_i and σ_i are not fit, but rather are determined directly from the mean and variance of the actual time series using the formulae:

$$\sigma_i = \frac{2}{\frac{\langle x_i \rangle^2}{\sigma^2_{x_i}} - 1}, \quad K_i = \frac{\langle x_i \rangle}{1 - \frac{\sigma_i}{2}} \tag{7}$$

where $\langle x_i \rangle$ is the mean abundance of the population and $\sigma^2_{x_i}$ is its variance. The parameter τ_i was held constant ($\tau_i = 1$) for all strains to avoid overfitting.

To calculate $\sigma^2_{x_i}$, we used the sampling-corrected estimate of the true variance as done in [1] and [43]:

$$\sigma^2_{x_i} = \frac{1}{|T|}\sum_{t\in T}\frac{x_i(t)(x_i(t)-1)}{N(t)(N(t)-1)} - \left(\frac{1}{|T|}\sum_{t\in T}\frac{x_i(t)}{N(t)}\right)^2. \tag{8}$$

where T is the set of timepoints for which strain i is present, and $N(t)$ is the total abundance of all species present in the sample at timepoint t, as determined by MIDAS.

Goodness of fit test

The goodness of fit test for the SLM was adopted from the test described in [50]. The null hypothesis of this test is that the SLM with the parameters determined in the previous section generated the observed time series.

The test is performed as follows. Suppose that $x(t_0), x(t_1), ..., x(t_{T-1})$ are the T observations of strain's abundance, at times $t_0, t_1, ...t_{T-1}$. M simulations are performed using the Euler-Maruyama procedure described in equation (6), above, from time t_{i-1} until time t_i, starting at initial abundance x_{i-1}.

Let $X_i^{(m)}$ be the the mth simulated value at time t_i for $m = 1, 2, ..., M$. Define r_i to be the number of $X_i^{(m)} < x(t_i)$—that is, the number of simulations of the process from t_{i-1} to t_i in which the final simulated value was less than the true abundance.

Under the null hypothesis, the r_i are equally likely to take any value between 0 and M. Therefore, we perform a χ^2 goodness-of-fit test to determine if the r_i follow a uniform distribution on $0, 1, ..., M$, obtaining a p-value. We repeat this whole process 1000 times for each time series, and take the true p-value to be the median p-value across all runs. We rejected the null hypothesis at a significance level of 5%.

Acknowledgments

We thank Colin Kremer for his critical feedback on this manuscript, and Van Savage for early discussions on this work. We also thank members of the Garud lab for their feedback. This work was supported by the NSF Postdoctoral Research Fellowships in Biology Program under Grant No. 2010885 (W.R.S.). as well as the the Paul Allen Foundation, the UCLA Hellman Fellowship, and the Research Corporation for Science Advancement (N.R.G).

References

1. Grilli, J. Macroecological laws describe variation and diversity in microbial communities. *Nat Commun* 11,4743 (2020).

2. Smith, J. Maynard, et al. "How clonal are bacteria?" *Proceedings of the National Academy of Sciences* 90.10 (1993): 4384-4388.

3. Shapiro, B. Jesse. "How clonal are bacteria over time?." *Current opinion in microbiology* 31 (2016): 116-123.

4. Lenski, R. "Experimental evolution and the dynamics of adaptation and genome evolution in microbial populations." *ISME* J 11, 2181–2194 (2017).

5. Chen, Lianmin, et al. "The long-term genetic stability and individual specificity of the human gut microbiome." Cell 184.9 (2021): 2302-2315.

6. Descheemaeker L., De Buyl S. "Stochastic logistic models reproduce experimental time series of microbial communities." *Elife* (2020)

7. Ho, Po-Yi, Benjamin H. Good, and Kerwyn Huang. "Competition for fluctuating resources reproduces statistics of species abundance over time across wide-ranging microbiotas." *bioRxiv* (2021).

8. Goyal, A., Bittleson, L., Leventhal, G., Lu, L., Cordero, X. "Interactions between strains govern the eco-evolutionary dynamics of microbial communities." *bioRxiv* (2021).

9. Verster A; Ross B; Radey M; Bao Y; Goodman A; Mougous J; Borenstein E. "The Landscape of Type VI Secretion across Human Gut Microbiomes Reveals Its Role in Community Composition." *Cell Host Microbe* (2017).

10. Garud N; Good B; Hallatschek O; Pollard K. "Evolutionary dynamics of bacteria in the gut microbiome within and across hosts." *PLOS Biology* (2019).

11. Li, Simone S., et al. "Durable coexistence of donor and recipient strains after fecal microbiota transplantation." *Science* 352.6285 (2016): 586-589.

12. Smillie C. *et al.* "Strain Tracking Reveals the Determinants of Bacterial Engraftment in the Human Gut Following Fecal Microbiota Transplantation." *Cell Host Microbe.* (2018).

13. Nayfach, S., Rodriguez-Mueller, B., Garud, N., & Pollard, K. S. "An integrated metagenomics pipeline for strain profiling reveals novel patterns of bacterial transmission and biogeography." *Genome research*, 26(11), 1612-1625. (2016).

14. Shoemaker, W. R., Locey, K. J., & Lennon, J. T. "A macroecological theory of microbial biodiversity." *Nature ecology & evolution*, 1(5), 1-6. (2017)

15. Yaffe, Eitan, and David A. Relman. "Tracking microbial evolution in the human gut using Hi-C reveals extensive horizontal gene transfer, persistence and adaptation." *Nature microbiology* 5.2 (2020): 343-353.

16. Russell, Shelbi L., and Colleen M. Cavanaugh. "Intrahost genetic diversity of bacterial symbionts exhibits evidence of mixed infections and recombinant haplotypes." *Molecular Biology and Evolution* 34.11 (2017): 2747-2761.

17. Lloyd-Price, J. *et al.* "The integrative human microbiome project." *Nature* (2019).

18. Good, Benjamin H., Stephen Martis, and Oskar Hallatschek. "Adaptation limits ecological diversification and promotes ecological tinkering during the competition for substitutable resources." *Proceedings of the National Academy of Sciences* 115.44 (2018): E10407-E10416.

19. Tikhonov, Mikhail. "Theoretical microbial ecology without species." *Physical Review E* 96.3 (2017): 032410.

20. Tropini, Carolina, et al. "The gut microbiome: connecting spatial organization to function." *Cell host & microbe* 21.4 (2017): 433-442.

21. Poyet, M., Groussin, M., Gibbons, S.M. *et al.* "A library of human gut bacterial isolates paired with longitudinal multiomics data enables mechanistic microbiome research." *Nat Med* 25, 1442–1452 (2019).

22. Voigt, A.Y., Costea, P.I., Kultima, J.R. *et al.* "Temporal and technical variability of human gut metagenomes." *Genome Biol* 16, 73 (2015).

23. Roodgar, M., Good, B. H., Garud, N. R., Martis, S., Avula, M., Zhou, W., Snyder, M. P. *et al.* "Longitudinal linked read sequencing reveals ecological and evolutionary responses of a human gut microbiome during antibiotic treatment." *Genome Research*. (2021).

24. Schloissnig, S., Arumugam, M., Sunagawa, S., Mitreva, M., Tap, J., Zhu, A., Bork, P. "Genomic variation landscape of the human gut microbiome." *Nature*, 493(7430), 45-50. (2013).

25. Faith, J. J., Guruge, J. L., Charbonneau, M., Subramanian, S., Seedorf, H., Goodman, A. L., Gordon, J. I. "The long-term stability of the human gut microbiota." *Science*, 341(6141). (2013)

26. Hubbell, Stephen P. "The unified neutral theory of biodiversity and biogeography" (MPB-32). Princeton University Press, 2011.

27. Karita, Yuya, David T. Limmer, and Oskar Hallatschek. "Scale-dependent tipping points of bacterial colonization resistance." bioRxiv (2021).

28. Zheng, W, et al. “Microbe-seq: high-throughput, single-microbe genomics with strain resolution, applied to a human gut microbiome.” *bioRxiv* (2020)

29. Goldford, Joshua E., et al. “Emergent simplicity in microbial community assembly.” Science 361.6401 (2018): 469-474.

30. Zhao, S., *et al.* “Adaptive evolution within gut microbiomes of healthy people.” *Cell host microbe*, 25(5), 656-667. (2019).

31. Marsland III, Robert, et al. “Available energy fluxes drive a transition in the diversity, stability, and functional structure of microbial communities.” *PLoS computational biology* 15.2 (2019): e1006793.

32. Korpela, Katri, et al. “Selective maternal seeding and environment shape the human gut microbiome.” *Genome research* 28.4 (2018): 561-568.

33. Lloyd-Price, J., Abu-Ali, G. Huttenhower, C. “The healthy human microbiome.” *Genome Med* 8, 51 (2016). https://doi.org/10.1186/s13073-016-0307-y

34. Tang, Le et al. “Defining natural species of bacteria: clear-cut genomic boundaries revealed by a turning point in nucleotide sequence divergence.” *BMC genomics* vol. 14 489. 18 Jul. 2013, doi:10.1186/1471-2164-14-489

35. Cordero, Otto X., and Martin F. Polz. “Explaining microbial genomic diversity in light of evolutionary ecology.” *Nature Reviews Microbiology* 12.4 (2014): 263-273.

36. Lewontin, Richard C. The genetic basis of evolutionary change. Vol. 560. New York: Columbia University Press, 1974.

37. Murray, C.S., Gao, Y. Wu, M. “Re-evaluating the evidence for a universal genetic boundary among microbial species.” *Nat Commun* 12, 4059 (2021). https://doi.org/10.1038/s41467-021-24128-2

38. Cordero, O., Polz, M. Explaining microbial genomic diversity in light of evolutionary ecology. Nat Rev Microbiol 12, 263–273 (2014). https://doi.org/10.1038/nrmicro3218

39. Utter, Daniel R., Colleen M. Cavanaugh, and Gary G. Borisy. ”Gene-and genome-centric dynamics shape the diversity of oral bacterial populations.” bioRxiv (2021).

40. Leventhal, Gabriel E., et al. “Strain-level diversity drives alternative community types in millimetre-scale granular biofilms.” *Nature microbiology] 3.11 (2018): 1295-1303.*

41. Mark Welch, Jessica L., et al. “Dynamics of tongue microbial communities with single-nucleotide resolution using oligotyping.” *Frontiers in microbiology* 5 (2014): 568.

42. Conwill, Arolyn, et al. “Anatomy promotes neutral coexistence of strains in the human skin microbiome.” *bioRxiv* (2021).

43. Zaoli, Silvia, and Jacopo Grilli. “A macroecological description of alternative stable states reproduces intra-and inter-host variability of gut microbiome.” *bioRxiv* (2021).

44. Ho, Po-Yi, Benjamin H. Good, and Kerwyn Huang. "Competition for fluctuating resources reproduces statistics of species abundance over time across wide-ranging microbiotas." *bioRxiv* (2021).

45. Ji, Brian W., et al. "Macroecological dynamics of gut microbiota." *Nature microbiology* 5.5 (2020): 768-775.

46. Rosen, Michael J., et al. "Fine-scale diversity and extensive recombination in a quasisexual bacterial population occupying a broad niche." Science 348.6238 (2015): 1019-1023.

47. Kashtan, Nadav, et al. "Single-cell genomics reveals hundreds of coexisting subpopulations in wild Prochlorococcus." *Science* 344.6182 (2014): 416-420.

48. Bak, Jakob, Henrik Madsen, and Henrik Aalborg Nielsen. "Goodness of fit of stochastic differential equations." *Symposium i Anvendt Statistik*. Copenhagen Business School Copenhagen, Denmark, 1999.

49. Tropini, Carolina, et al. "Transient osmotic perturbation causes long-term alteration to the gut microbiota." *Cell* 173.7 (2018): 1742-1754.

50. Allen, Edward. Modeling with Itô stochastic differential equations. Vol. 22. Springer Science Business Media, 2007.

51. Anderson, Marti J., and John Robinson. "Permutation tests for linear models." Australian New Zealand Journal of Statistics 43.1 (2001): 75-88.

52. Martinson JNV, Pinkham NV, Peters GW, Cho H, Heng J, Rauch M, et al. "Rethinking gut microbiome residency and the Enterobacteriaceae in healthy human adults." *ISME J.* 2019

53. Franzosa, Eric A., et al. "Identifying personal microbiomes using metagenomic codes." *Proceedings of the National Academy of Sciences* 112.22 (2015): E2930-E2938.

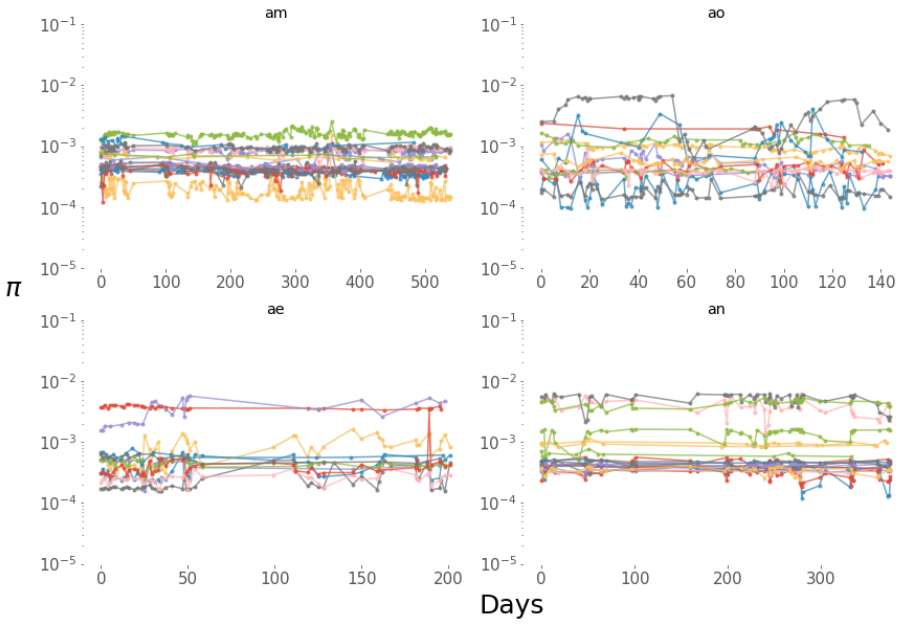
54. Truong, Duy Tin, et al. "Microbial strain-level population structure and genetic diversity from metagenomes." Genome research 27.4 (2017): 626-638.

55. Helling, Robert B., Christopher N. Vargas, and Julian Adams. "Evolution of Escherichia coli during growth in a constant environment." *Genetics* 116.3 (1987): 349-358.

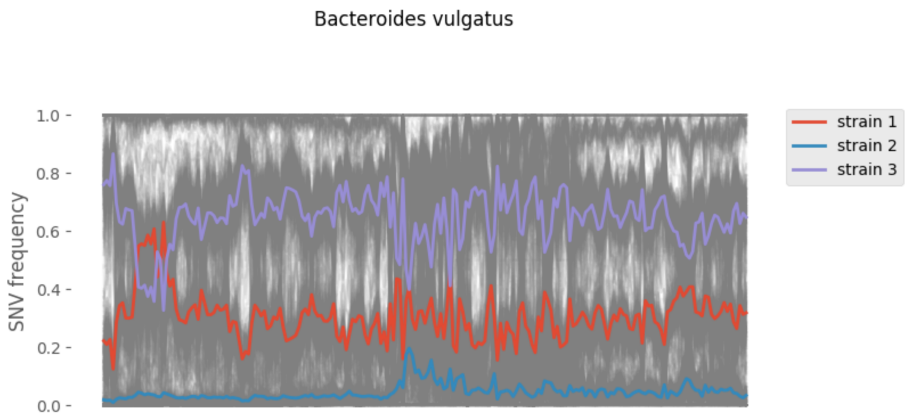
56. Ghoul, Melanie, and Sara Mitri. "The ecology and evolution of microbial competition." *Trends in microbiology* 24.10 (2016): 833-845.

57. Verster, Adrian J., and Elhanan Borenstein. "Competitive lottery-based assembly of selected clades in the human gut microbiome." *Microbiome* 6.1 (2018): 1-17.

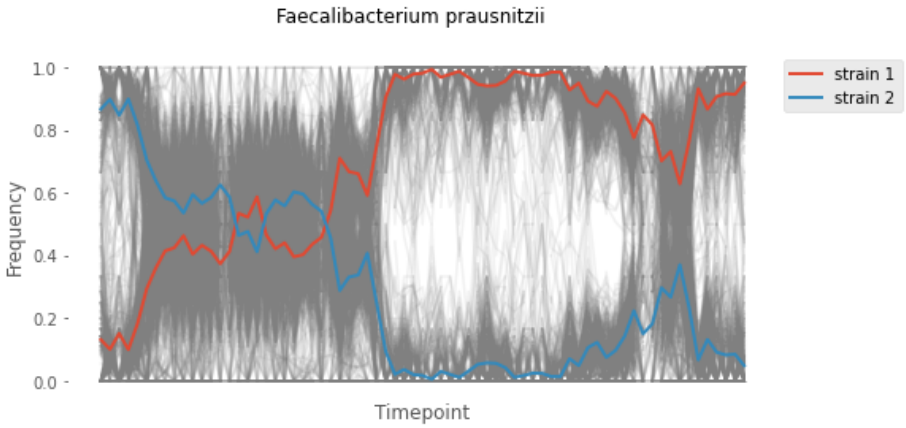
Supplementary Information



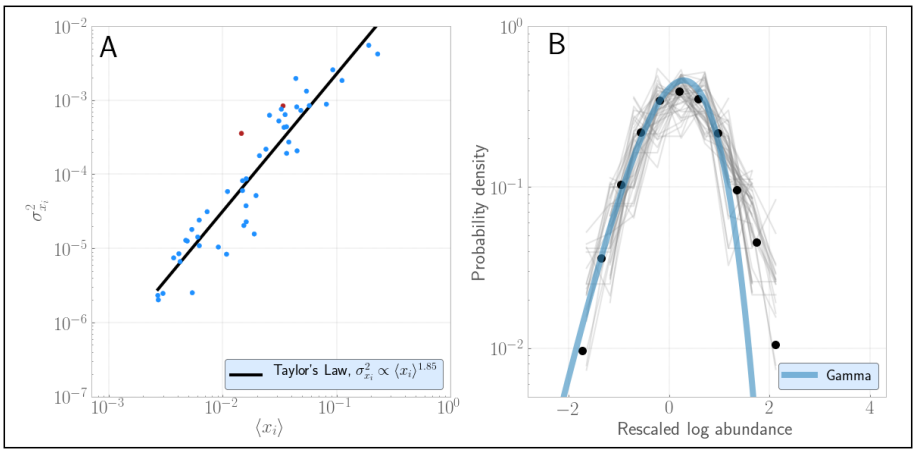
Supplementary Figure 1: Nucleotide diversity π for each species for which SNVs could be inferred, grouped by host. Each color represents a different species. Throughout the course of the sampling period, π undergoes large fluctuations in some species. However, in 55 of 64 cases, these fluctuations show no directional trend significantly different from from 0, according to our permutation test.



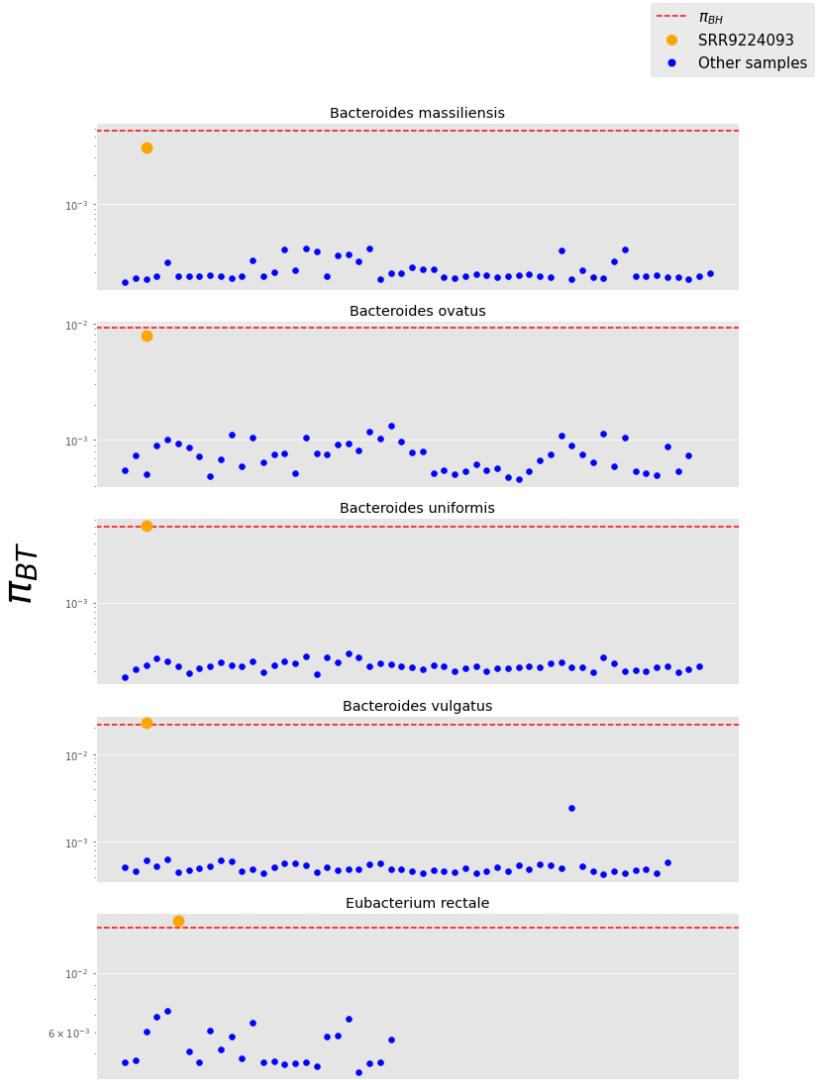
Supplementary Figure 2: Unphased allelic trajectories in *B. vulgatus* in host *am* (grey lines), and imputed strain frequencies. After phasing alleles to strains, only two of the clusters of SNVs remain visible.



Supplementary Figure 3: Two strains of *F. prausnitzii* present in host *ao* underwent a partial strain displacement event. In grey are the unphased trajectories of all SNPs detected throughout the sampling period. In red and blue are the inferred underlying strain frequencies detected after clustering SNP trajectories (Methods).



Supplementary Figure 4: The macroecological laws plotted only for strains belonging to species which harbor more than one strain. The red points in **A**, corresponding to the two strains of *F. prausnitzii* in host *ao*, are clear outliers, as expected given the partial strain replacement event these strains underwent. While the slope of the Taylor's Law exponent increased from 1.63 to 1.85 when considering only these strains, the power law scaling of the variance in abundance with the mean still holds. Similarly, the Gamma AFD still provides a good approximation to the true (rescaled) strain AFDs.



Supplementary Figure 5: One sample, SRR9224093, was removed from our analysis due to a possible mislabelling or contamination, which was noticed while performing this analysis. This sample was labelled to have come from host *ae*. Here, each blue dot represents π_{BT} between each sample from that host and a reference sample from host *ae* (the earliest sample detected for that species in this host). Only those species with more than three samples are plotted. SRR9224093 (orange dot) shows anomalously high π_{BT} compared to other samples. In particular, π_{BT} approaches π_{BH} only for this sample across multiple species, indicating that the genetic content of this sample is as different from the others in this host as a sample chosen from a random other individual would be. The consistent elevation of π_{BT} across multiple species at only this sample suggests this sample may have been mislabelled, or else may have been partially contaminated.