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 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 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. Here, we analyze the daily dynamics of intra-specific genetic variation in the gut microbiomes of four healthy, densely longitudinally sampled hosts. First, we find that overall genetic diversity in a large majority of species is stationary over time, despite short-term fluctuations. Next, we show that fluctuations in abundances in approximately 80% of strains analyzed can be predicted with a stochastic logistic model (SLM)—an ecological model of a population experiencing environmental fluctuations around a fixed carrying capacity which has previously been shown to capture statistical properties of species abundance fluctuations. The success of this model indicates that strain abundances typically fluctuate around a fixed carrying capacity, suggesting that most strains are dynamically stable. Finally, we find that the strain abundances follow several empirical 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 a complex ecological community, composed of tens of trillions of cells that interact directly and indirectly with one another and the host [20, 44, 45]. Though the precise species composition of the gut microbiome differs between hosts, healthy adult guts tend to be both ecologically diverse and temporally stable at the species level under normal circumstances [21, 23, 29, 40]. The ecological stability of the gut community is critical to the preservation of its functional capacity through time, and periods of instability and heightened variability are often associated with environmental perturbation or disease states [9, 22, 24].

Much as the gut community as a whole is made up of a diverse array of species, within species, populations of gut microbes harbor many genetic variants [38]. A growing body of literature highlights the importance of inter-host differences in microbiome genetic composition for various aspects human health, with specific microbial genotypes and strains associated with the digestion

of certain foods [25], a range of host disease risk factors [7], bile and lipid composition [8], and antibiotic resistance [10]. Recent studies have begun to characterize how this genotypic diversity changes through time within hosts [15, 46, 38, 40], and have linked longitudinal changes in genetic composition to specific host phenotypes and metabolite levels [6].

Broadly, dynamic changes in genetic variation in the gut microbiome occur at two distinct levels. First, there are changes in the frequencies of lineages that have clonally diverged since their common ancestor colonized the host due to the evolutionary forces of mutation, drift, selection, and recombination. Typically, such lineages differ from one another at a small number $(\mathcal{O}(1) - \mathcal{O}(10^2))$ of single nucleotide variants (SNVs) [46, 15]. Second, there are fluctuations in the relative abundances of con-specific strains that do not share a common ancestor within the host. When levels of recombination between such strains are sufficiently low, clonal descendants of the initial colonizers may persist within the host as genetically distinguishable populations differing from one another at a similar number of sites in their shared, core genome $(\mathcal{O}(10^3) - \mathcal{O}(10^4))$ as strains drawn from unrelated hosts [79, 38, 73, 78, 15]. Multiple colonization by con-specific strains is evidently under some degree of ecological constraint, as only a small handful of strains (typically between one and four) are ever observed within a host at any one time—a phenomenon dubbed "oligo-colonization" [36, 15, 69, 14]. The mechanisms enabling a small number of strains to colonize a host and rise to high frequency, but preventing a large number of exogenous strains from doing the same, are not known as yet. Interestingly, similar colonization patterns have been observed in a number of other host-associated microbiota, both at different human body sites [57], and in other organisms [27, 13].

In healthy, adult hosts, a large majority of strains persist over periods of months to years [38, 40, 31, 6, 26, 15]. Moreover, strains within the gut can remain resilient in the face of large perturbations, such as antibiotics [36] and fecal microbiome transplants [16]. However, little is known about the magnitude of daily fluctuations in genetic composition at either the strain or lineage level under ordinary conditions, or how such fluctuations ultimately affect the stability properties of the gut community, as longitudinal studies of genetic diversity in the gut have tended to focus on samples collected at multi-month intervals.

In this work, we seek to understand how the genetic composition of the gut changes through time, from daily to multi-year timescales in four healthy, adult hosts sampled over the course of six to 18 months [34]. To do so, we leverage concepts from macroecology to examine the dynamics of strains in the these four hosts. Macroecology focuses on characterizing statistical regularities in patterns of abundance and diversity within and between ecological communities. A growing body of work has demonstrated that species-level patterns of diversity in a variety of natural microbial communities are well-described by macroecological laws [1, 60, 19, 11]. Many of these macroecological laws can be recapitulated through intuitive ecological models containing few if any free parameters [1, 11, 59]. Among these successful models is the Stochastic Logistic Model (SLM), which describes the dynamics of a population experiencing rapid stochastic fluctuations induced by environmental noise around a fixed carrying capacity [80]. Whether the populations 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 work, we find not only that a large majority of strains in these healthy hosts exhibit abundance dynamics consistent with an SLM, but also that strain abundance fluctuations follow several macroecological laws known to hold among species [1, 11, 59, 60].

Together, our results indicate that daily fluctuations in overall genetic composition within the gut microbiome are largely stationary, and that these fluctuations follow broad macroecological

patterns. Thus, several macroecological properties of the human gut known to hold at higher levels of taxonomic organization, including its stability, appear to emerge at the level of strains.

Results

To explore how daily fluctuations in nucleotide diversity and strain abundances translate into stability over periods of months to years, we used high-resolution temporal data from four hosts sampled in the BIO-ML project [34] (see Materials and Methods for further details on sampling).

Temporal stability of intraspecific genetic variation

Within hosts, allele frequencies change through time in gut microbial populations due to mutation, drift, selection, and fluctuations in the relative abundances of strains. While studies examining broad cohorts of sparsely longitudinally sampled individuals indicate that the magnitude of intra-host fluctuations only infrequently approaches that of inter-host differences over timescales of months to years [38, 40, 15], more finely-resolved temporal trends are less well characterized. To evaluate the stability of the gut community, it is crucial to determine whether temporal fluctuations are stationary or directional.

To assess the stability of intra-specific genetic variation through time, we examined temporal trends in patterns of nucleotide diversity within our four hosts using F_{ST}, a standard measure of genetic differentiation between populations (see Materials and Methods for further details on F_{ST} calculations). If the genetic composition of a species changes directionally, we expect that samples drawn later in the timecourse will have greater F_{ST} relative to the initial timepoint than earlier samples. If, by contrast, fluctuations are stationary, then later timepoints should on average be no more diverged from the initial sample than earlier timepoints.

To contextualize the magnitude of variation in genetic composition through time, we normalized our longitudinal measurements of F_{ST} within hosts by the mean F_{ST} value of the species across hosts. We calculated this species-wide mean F_{ST} using shotgun metagenomic data from 250 North American hosts sampled in the Human Microbiome Project [28, 29], allowing us to better capture the extent of inter-host diversity. We refer to the resulting normalized F_{ST} statistic, obtained by dividing each intra-host F_{ST} measurement by the mean F_{ST} across hosts, as F'_{ST}. F'_{ST} will approach or exceed one when intra-host fluctuations are of the same magnitude as inter-host differences, and will remain close to zero when the genetic composition of the population is constant over time.

In line with previous work [38, 40], we observe that changes in genetic composition within the hosts we examine (*am*, *ao*, *an*, and *ae*) rarely approach the magnitude of inter-host differences, as F'_{ST} remained well below one at all timepoints for all but one species examined (Figure S5 in S1 Text). For the one aberrant species, *Faecalibacterium prausnitzii* in host *ao*, F'_{ST} increases steadily before a rapid rise above one around timepoint 60 (Figure 1B). More typical, however, is the example of *Bacteroides vulgatus* in host *am*, for which F'_{ST} fluctuates but appears to remain near a long-term steady state (Figure 1A).

To determine quantitatively whether fluctuations in genetic composition were stationary or directional, we implemented an Augmented Dickey-Fuller test (ADF) for each F'_{ST} time series [77]. The ADF tests the null hypothesis that a time series is non-stationary against the alternate hypothesis that the time series is stationary. Rejecting the null hypothesis with the ADF is thus evidence that the mean and variance in a temporally varying quantity are time invariant. For 34 of the 45

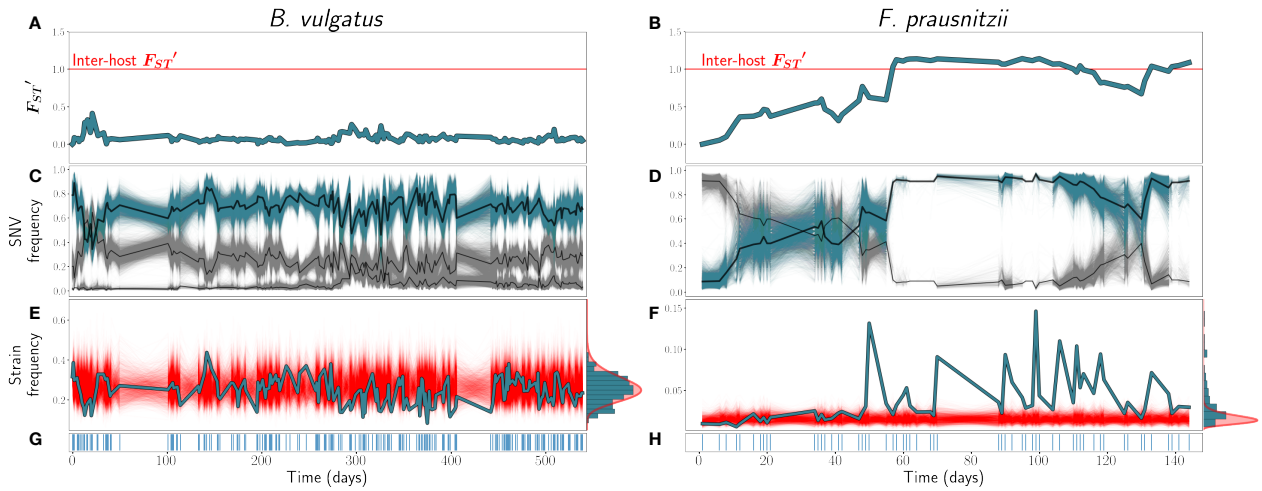

Figure 1: A & B) $F_{ST}{}'$ trajectories for *B. vulgatus* (*am*) and *F. prausnitzii* (*ao*), respectively. **C & D**): SNV frequencies of three inferred strains for B. vulgatus (C) and two inferred strains for F. prauznitzii (D). In black are the inferred strain trajectories. Highlighted in blue are example strains featured further in (E) and (F). **E & F**) Frequencies of the example strains in blue, with simulations of the corresponding SLM overlaid in red. At right, the empirical distribution of strain abundances is plotted in blue, and the stationary Gamma distribution of abundances (see Equation 2) predicted by the SLM in red. **G & H**) Sampling time points. Blue lines indicate that a sample was taken on that day.

species (76%) considered, we rejected the null hypothesis at a significance level of $p = 0.05$, indicating that the majority of species exhibit stationary F'_{ST} trends.

Our results suggest that fluctuations in allele frequencies are typically stable, at least when coarse-grained across the whole genome, in these hosts. However, large-scale, rapid changes in allele frequency, such as observed in *F. prausnitzii*, also occur.

Strain frequencies

Recent empirical studies using isolate, single-cell, and shotgun sequencing data have demonstrated that at any one time, the human gut microbiome is colonized by at most a handful of distinct conspecific strains (between one and four) which do not share a common clonal ancestor within the host [31, 34, 15, 78, 69]. Much less is known, however, about the dynamics of strains once they have colonized the host.

To investigate these dynamics, we infer strain genotypes and frequencies using an algorithm adapted from by Roodgar *et al*. This algorithm identifies large clusters of SNVs with tightly correlated frequency trajectories [36], indicative of linkage on a common genomic background. By identifying large clusters, we expect to distinguish the trajectories of deeply diverged strains (for further information on our strain phasing, see S1 Text section 3.2). When no large cluster of tightly linked SNVs was detected within a species, we inferred that only a single strain was present.

Of the 45 species/host pairs examined, 15 (33.3%) harbored multiple strains. Of these 15, 13

were colonized by two strains, while in two separate hosts (*am* and *an*), the species *B. vulgatus* was composed of three strains. The number of fixed differences between strains varied between $\mathcal{O}(10^3) - \mathcal{O}(10^4)$, though due to our conservative filters for both calling SNVs from reads and assigning SNVs to strains, these likely represent underestimates of the true divergence between these strains. For further discussion of inter-strain genetic divergence, see S1 Text section 3.1.

As may be expected given our F'_{ST} results, most strains, by visual inspection, exhibit frequency dynamics that are heuristically consistent with stationarity (see Supplementary Text S2—S5 for strain trajectories of all species). The relative frequencies of most strains appear to fluctuate around a constant value throughout the sampling period. In a typical example of this kind of behavior, the dominant strain of *B. vulgatus* in host *am* fluctuates around 60% frequency for more than 500 days (Figure 1C). However, in a minority of cases, strain frequencies shifted dramatically throughout the sampling period. The most striking example of this is *F. prausnitzii* in *ao*, which we have already seen underwent fluctuations in F'_{ST} (Figure 1B) of the magnitude of inter-host differences in genetic composition. In this species, an initially rare strain almost fully supplants the initially dominant strain within the span of 60 days, before a partial reversion later in the timecourse (Figure 1D). In another case, *P. distasonis* in host *ao*, a single strain which initially falls below the detection threshold rises to detectable abundance midway through the sampling timecourse (S3 Text). Similarly, the minor strain of *B. vulgatus* in *am* is initially very rare before rapidly increasing in frequency around day 300, and subsequently reverting to an intermediate steady state below its maximum frequency. This subtle shift in strain frequencies is not, interestingly, detected by our ADF test of F'_{ST} as a departure from stationarity—likely due to the fact that the absolute magnitude of the shift in allele frequencies between the beginning and end of the timecourse is small (less than 5%).

We further note the striking visual correspondence between the F'_{ST} and strain frequency trajectories, as is evident in Figure 1 for both *B. vulgatus* and *F. prausnitzii*. Because fluctuations in the relative frequencies of co-colonizing strains with respect to one another determine allele frequencies at a very large fraction of all polymorphic sites, genome-wide average diversity statistics like F'_{ST} will reflect strain dynamics. This correspondence is therefore supporting evidence that the overwhelming majority of genetic variation in these species is due to fixed differences between strains, rather than among lineages belonging to strains. However, as is evident from the example of the minor strain of *B. vulgatus* in *am*, subtle but important strain dynamics can also be obscured when only considering genome-wide average patterns of diversity.

Stochastic logistic model

In the preceding sections, we have seen that the genetic composition of most species examined exhibit stationary dynamics over the timescale of observation. We hypothesized that this behavior might result from the underlying strains fluctuating around fixed absolute carrying capacities.

To test this hypothesis, we assessed the fit and predictive capacity of the stochastic logistic model (SLM)—a model of a population experiencing stochastic fluctuations around a fixed abundance. Recent work in microbial ecology has demonstrated the power of minimal models like the SLM, requiring the fit of no free parameters, to reproduce qualitative and quantitative features of natural microbial community dynamics [1, 59, 58, 11, 80]. Here, we tested the capacity of the SLM to forecast future strain behavior when trained on an initial subset of timepoints. By training only on a subset of initial points, we can assess whether strain dynamics are consistent through time.

Under the assumptions of the SLM, each population i has a long-term carrying capacity K_i, and temporal fluctuations in abundance around this value are driven by environmental noise with am-

plitude σ_i. The dynamics of a population governed by an SLM can be expressed with the following stochastic differential equation:

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

where τ_i^{-1} is the growth rate, and $\eta(t)$ is a Brownian noise term.

Populations following an SLM may experience large fluctuations in abundance over short timescales, and may even be temporarily found far from their long-term average value, but these deviations will be transient. Over long timescales, the observed distribution of abundances will converge to a stationary Gamma distribution [1]:

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

To determine whether strain trajectories could be described by an SLM, we first obtained time series of strain abundances by multiplying the relative frequencies of the strains inferred in the previous section by the relative abundance of the species to which they belong.

Next, we estimated K_i and σ_i from the first third of timepoints for each strain. K_i and σ_i are not free parameters, but rather are functions of the mean and variance of observed abundances (for details, see S1 Text section 4). To assess quantitatively whether the time series of strains in our cohort could be adequately described by an SLM, we developed and implemented a goodness-of-fit test. The test determines whether the transitions between subsequent timepoints are consistent with an SLM (for further details, see S1 Text section 5). Qualitatively, if a strain follows an SLM, its average and variance in abundance in the latter two-thirds of the time series should match that of the former third, and the strain should have a tendency to revert to its carrying capacity K_i.

Returning to our case studies, we see that the dominant strain of *B. vulgatus* is well described by the SLM (Figure 1E). The true abundance trajectory (in blue) explores largely the same space as the SLM simulations (in red). Moreover, the empirical distribution of abundances across the entire timecourse appears to approach the stationary Gamma distribution (Equation 2) predicted by the SLM (Figure 1E, at right). By contrast, the dynamics of the invasive strain of *F. prausnitzii* deviate strongly from the SLM (Figure 1F).

Overall, 79% (49/62) of strains passed our goodness of fit test, with 83% passing in *am*, 86% passing in *ao*, 73% passing in *an*, and 69% passing *ae* (Figure 2). While the dominant strain of *B. vulgatus* passes, both strains of *F. prausnitzii* in *ao*, as well as the minor strain of *am*, fail the SLM. Thus, the SLM recapitulates the qualitative behavior of a large majority of strains even when trained on only a subset of initial points, while also having power to discriminate instances in which the dynamics of strains are evidently quite non-stationary.

Moreover, the likelihood of exhibiting SLM dynamics is independent of the presence of other strains. Of the 32 strains for which another con-specific strain was present, 24 (75%) passed the SLM test, while among the 30 singly colonizing strains, 25 (83%) passed test. Thus, while singly colonizing strains tend to be moderately more likely to exhibit stable dynamics, the difference in pass rate is not statistically significant ($\chi^2 = 0.24$, p-value $= 0.62$), indicating that the presence of con-specific strains is not *prima facie* destabilizing for a focal strain.

Next, we conducted an identical test of the SLM at the species level. Overall, 84% (38/45) of species exhibited dynamics consistent with a SLM. In the case of *F. prausnitizii*, for instance, the

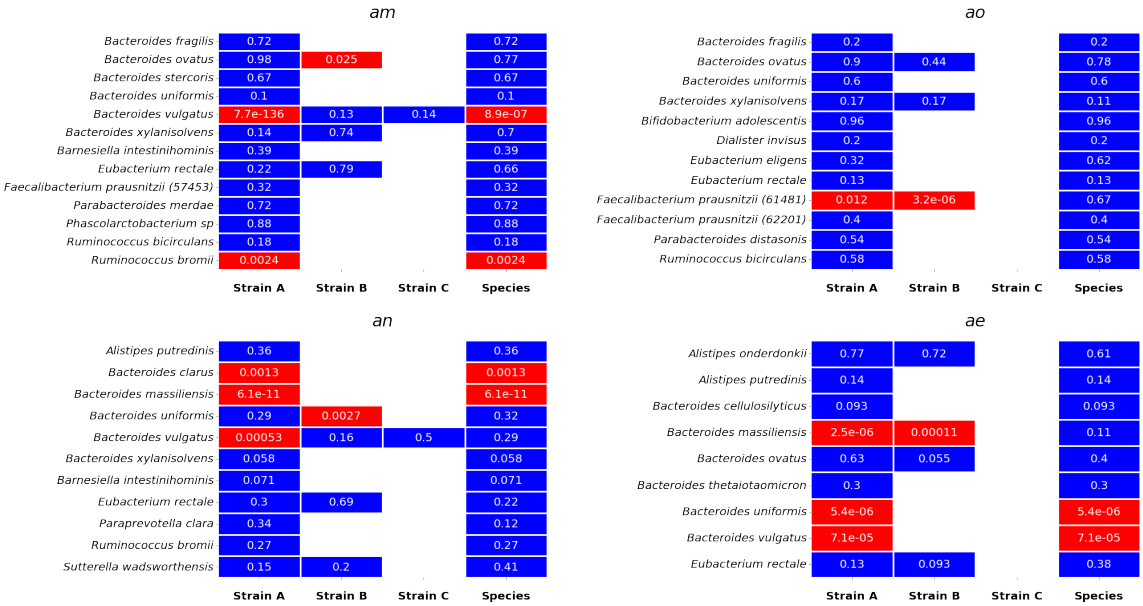

Figure 2: Results of the SLM goodness of fit test, by host. 79% (49/62) of strains, and 86.6% (39/45) of species, exhibit stochastic logistic dynamics across the sampling interval. The percentage of strains passing the test varied between hosts, with 83% passing in *am*, 87% passing in *ao*, 75% passing in *an*, and 69% passing in *ae*. The p-value associated with each strain or species is shown in white within each cell.

.

abundance of the species overall fluctuated stably, obeying the SLM. Thus, despite a partial replacement event, the total abundance of the species remained roughly constant. One interpretation of this observation is that these strains strongly compete with one another for the same species' niche. Interestingly, *F. prausnitizii* is known to experience higher rates of replacement over multi-year timescales that other gut commensals, and these replacements are associated with alterations in levels of plasma metabolites which affect host immunity [6]. However, we emphasize that the replacement was only partial, and the "displaced" strain recovered temporarily to intermediate abundance. This example highlights the complexity of strain dynamics, as well as their potential relevance for host phenotypes.

Macroecology of strains

Much as the dynamics of individual strains can largely be recapitulated with a single relatively simple model, we can also attempt to parsimoniously characterize patterns of variability across strains collectively. Such low-dimensional representations of complex community dynamics are the natural purview of macroecology, which attempts to characterize variation within and among communities by observing the statistical patterns of abundance, distribution and diversity across their constituent members. In many kinds of microbial ecosystems, including the human gut, patterns of species abundance and distribution have been shown to broadly follow a number of macroecological laws, including Taylor's Law and a Gamma distribution of abundance fluctuations [1, 60, 11]. We show here that these macroecological relationships also 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, which can be stated as:

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

where $\langle x_i \rangle$ and $\sigma^2_{x_i}$ are the mean and variance of x_i, respectively, and α is the scaling exponent of the power law.

Many mechanisms can give rise to Taylor's Law. For instance, when the only source of variability between communities (or, in our case, longitudinal samples) is due to sampling noise, the Taylor's Law exponent α will equal one. By contrast, in communities where the scale of fluctuations is independent of abundance—that is, where all populations have identical per-capita fluctuations—α will equal two [60, 1] (for further details, see S1 Text section 6). We observed a Taylor's Law scaling with an exponent of $\alpha = 1.8$ among all strains (Figure 3A), mirroring previous findings at the species level [60]. This Taylor's Law exponent indicates that higher abundance strains are proportionally less variable than lower abundance strains ($\alpha < 2$), and that variation in strain abundance is not driven solely by sampling noise ($\alpha > 1$), but rather reflects true, underlying biological variability. However, the existence of a power-law scaling between mean and variance cannot, by itself, conclusively prove that any specific ecological model governs community dynamics. Indeed, the fit of the SLM does not depend on the existence of a Taylor's Law scaling, or vice versa, as the SLM can hold with arbitrary mean and variance.

The next pattern considered is the abundance fluctuation distribution (AFD), the overall distribution of abundances of a population through time. It is known that in a variety of microbial ecosystems, the AFD of many species tends to approach a Gamma distribution [1]. As discussed above, a population governed by stochastic logistic dynamics will tends towards a Gamma distribution of abundances over long timescales (see, for instance, the histogram of abundances for the

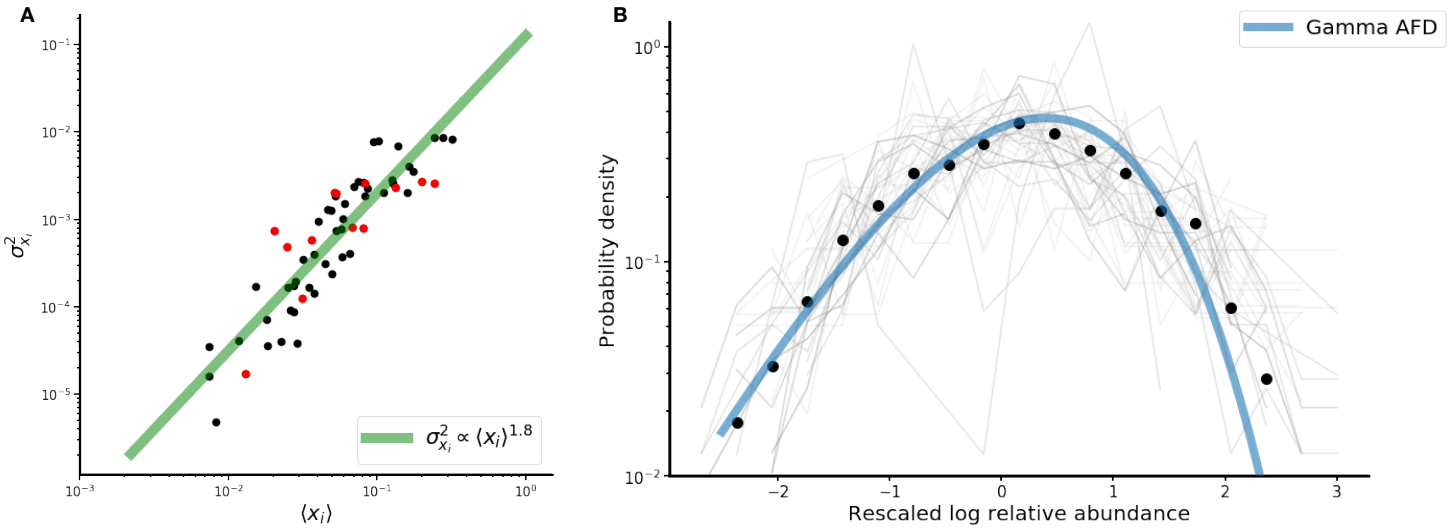

Figure 3: **A**) Scaling of variance in strain abundance with mean abundance obeys Taylor's Law with exponent of 1.8. Black dots are strains passing the SLM, while red dots are strains failing the SLM. **B**) Strain abundances approximately follow a Gamma distribution, which is the stationary distribution of the SLM. Black circles are the average probability densities of the rescaled abundance across all strains, and the blue line is the Gamma fit of the bin means of the rescaled distributions. Light gray lines are the individual rescaled abundance distributions for each strain individually (62, in total).

dominant strain of *B. vulgatus* in Figure 1E, at right). 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, on average, well-described by a Gamma distribution (black dots, blue line), though some individual strains (grey lines) deviate somewhat from the Gamma stationary distribution. Recalling that the SLM of a given strain is uniquely determined by its mean and variance, 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.

Discussion

In this study, we sought to characterize the within-species population dynamics of the human gut microbiomes of four healthy hosts. Previous efforts have shown that within hosts changes in the genetic composition of gut microbial populations through time tend to be small compared with inter-host differences [38, 40]. We build on this result by demonstrating that at daily temporal resolution, intra-specific diversity tends to fluctuate around a long-term average value within the hosts examined over periods of years. We show, crucially, that the abundance fluctuations of a large majority of strains we detect can be predicted by the stochastic logistic model (SLM) of growth, a model that also recapitulates fluctuations at the species level [1, 11, 80]. Lastly, we find that empirical patterns of strain abundance variation in these hosts follow macroecological laws which have also previously been demonstrated to hold at the species level, including Taylor's Law and a Gamma abundance fluctuation distribution [1, 11, 60]. Together, our results indicate that many of the broad properties of the gut observed at higher taxonomic levels of organization—such as its ecological and functional stability—may in fact emerge at the level of strains.

While the SLM was able to sufficiently describe strain dynamics for a majority of strains across species, its success was not universal, and deviations from this typical pattern were also informative. In one host, for instance, two strains of *F. prausnitzii* appear to undergo a rapid strain replacement, and fail the test. Whether this replacement was due to shifting environmental conditions, or direct inter-strain competition, is unclear. Regardless, our work indicates any successful description of gut microbial dynamics must incorporate the possibilities of both coexistence and rapid replacement. Over very long timescales, in fact, strain replacement may dominate the stable dynamics we observe here. Prior work [15, 26] suggests that over the course of decades, a large fraction of strains are ultimately replaced. One hypothesis is that this timescale reflects a waiting time for large environmental perturbations, such as antibiotics [36, 48] or bowel cleanses [64], but this is just one of many hypotheses. Indeed, this hypothesis is partially challenged by [36], where the strain content of an adult gut was perturbed during a course of antibiotics, but ultimately largely recovered to its pre-treatment state. This antibiotics study is a powerful demonstration of the stability of strains, even in the face of large perturbations. Investigating the possible explanations for the discrepancy between years-long and decades-long population dynamics at the strain level is an important problem that can be addressed with more extended timescales of observation.

While in this work we characterize the population dynamics of strains as ecological units, strain are by no means internally genetically homogeneous. The deep divergences we detect between conspecific strains ($\mathcal{O}(10^3) - \mathcal{O}(10^4)$ SNVs) are in fact genetic backgrounds, representing timescales of divergence likely far preceding colonization of the host. However, individual lineages bearing these backgrounds can differ from one another both at sites in the core genome and in gene con-

tent, and the relative frequencies of these different lineages with respect to one another can change due to evolution. Previous studies have shown that, at the level of lineages belonging to a strain, populations of gut microbes can experience both rapid selective sweeps [15, 34] and diversification into stably coexisting lineages [46].

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 [30]. In the context of the SLM, these eco-evolutionary feedbacks can viewed as tuning a strain's carrying capacity K_i, growth rate τ_i^{-1}, and sensitivity of the growth rate to environmental perturbation σ_i. 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_i given σ_i, 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.

The SLM is ultimately a phenomenological model, not a mechanistic one, and its success at the strain level does not explain why strains coexist. How and why closely related strains coexist in the human gut is one of the central biological questions raised by our results. Spatial segregation between strains, perhaps occupying different colonic crypts, or partitioning luminal and mucosal niches, could contribute to the observed pattern of strain coexistence [33, 42, 82], much as it does among *Cutibacterium acnes* strains inhabiting different pores on the facial microbiome [57]. However, spatial structure is far from the only mechanism that can foster coexistence between strains. For instance, genetic differences at polysaccharide utilization loci content may contribute to intra-host metabolic niche differentiation, potentially favoring coexistence of closely related strains [81]. Moreover, stably coexisting strains have been reported under laboratory conditions, as well [5, 13]. In these experiments, strains may 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. Indeed, recent theoretical work suggests that even subtle differences in resource uptake rates in high and low nutrient conditions may, in the presence of a temporally variable environment, lead to the coexistence of small numbers of closely related strains [76]. Investigating which of these mechanisms promotes strain coexistence in the human gut microbiome, and identifying the relevant genomic architectures, is an important avenue for future research.

We note that the four hosts examined here are not a representative sample of the full diversity of human lifestyles. For instance, all hosts were between the ages of 21 and 37, and residing in the United States at the time of sampling. The proportion of strains exhibiting stable or unstable dynamics may vary in different cohorts, but the tests we developed here will nonetheless be useful in identifying such differences. A future avenue of work will be to assess the generality of these findings in different cohorts—using new time series which are at least as long and densely temporally sampled as the BIO-ML data analyzed here, and of similar quality—including diseased or perturbed cohorts which may exhibit quite different dynamics.

Finally, our work highlights the importance of strains in understanding community structure and dynamics in the human gut microbiome [13, 76]. The ambiguity surrounding the bacterial species concept is well known [50] and reasonable alternatives have been proposed [32], but operationally species are, nonetheless, the predominant focus of attention in gut microbiome ecology. This focus is reasonable, as within-host strain structure is a comparatively recent discovery [69, 15, 31] and 16S rRNA sequencing provides an inexpensive, high-throughput means to examine community dynamics. However, it is reasonable to propose that for the human gut, and perhaps

other microbial ecosystems, many higher-level macroecological patterns of abundance and diversity may originate at the level of strains.

Materials and Methods

Longitudinal data

To investigate the temporal dynamics of the human gut microbiome, we analyzed densely sampled shotgun metagenomic time series data from four hosts from the BIO-ML project [34]. By analyzing shotgun metagenomic sequences, we capture longitudinal patterns of intra-specific genetic variation for many bacterial species in these communities simultaneously. A total of 402 samples were drawn from these four individuals (aliases *am*, *ao*, *an*, and *ae*), with 206 samples coming from *am*, 74 from *ao*, 63 from *an*, and 59 from *ae* (see Table S1 and Figure S2 in S1 Text for further details on sampling). All four hosts were healthy adults between the ages of 21 and 37, of whom three were male and one female, and all were residing in the USA at the time of sampling. Crucially, for our purposes, these hosts were sampled at very fine temporal resolution, with a median interval between successive samples of either one or two days in each host, over a period of five (*ao*) to 18 (*am*) months.

Aligning reads

To call single nucleotide variants (SNVs) and gene content, we aligned shotgun metagenomic reads to a panel of species which were prevalent and abundant within each host using MIDAS [18] (see S1 Text section 2 for further details on the bioinformatic pipeline employed). In total, we detected 45 species across the four hosts which met our coverage and prevalence criteria.

Stationarity of intra-specific genetic variation

To calculate F_{ST}, a measure of sub-population differentiation, we used the estimator:

$$F_{ST} = \frac{\pi_{BT} - \pi}{\pi_{BT}} \tag{4}$$

where π is the nucleotide diversity within a population and π_{BT} is the level of diversity between populations.

The nucleotide diversity π is a classical population genetic measure of polymorphism, representing the average number of pairwise SNV differences between randomly chosen members of a population. To determine π for a given species within a sample, we used the estimator:

$$\pi = \frac{1}{|G|} \sum_{i=1}^{|G|} \frac{r_i}{d_i} \frac{a_i}{d_i - 1} + \frac{a_i}{d_i} \frac{r_i}{d_i - 1} \tag{5}$$

where r_i is the count of the reference allele at site i, a_i is the count of the alternate allele, d_i is the depth of coverage, 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. Our π calculations follow the same methodology as [38]—an early,

foundational work characterizing patterns of genetic diversity in gut microbial populations within and across hosts—and are meant to be directly comparable.

Similarly, π_{BT}, the diversity between timepoints, was calculated:

$$\pi_{BT} = \frac{1}{|G|} \sum_{i=1}^{|G|} \left(\frac{r_{1i}}{d_{1i}} \frac{a_{2i}}{d_{2i}} + \frac{a_{1i}}{d_{1i}} \frac{r_{2i}}{d_{2i}} \right) \tag{6}$$

where r_{ji}, a_{ji}, and d_{ji} are the reference allele count, alternate allele count, and depth of coverage of site i in sample j, and $|G|$ is the total number of sites in the genome.

We calculated π for each species in each sample in each host, and π_{BT} for each sample relative to the initial timepoint.

To calculate F_{ST} between the initial timepoint (sample i) and sample j, we used the formula:

$$F_{ST}^{(ij)} = \frac{\pi_{BT} - \frac{(\pi_i + \pi_j)}{2}}{\pi_{BT}} \tag{7}$$

Lastly, we obtained our normalized statistic F'_{ST} by dividing $F_{ST}^{(ij)}$ by the species mean inter-host F_{ST}, estimated from a panel of 250 North American subjects sequenced by the Human Microbiome Project [28, 29]. To determine this species mean F_{ST}, we first calculated the pairwise F_{ST} for each pair of samples in which a species appeared, and then took the mean of these values.

To implement the Augmented Dickey-Fuller test on each time series of F'_{ST} values, we used the `adfuller` function from the python `statsmodels` library [83].

Strain inference

To phase strains, we use a modified version of the allele frequency trajectory clustering algorithm developed by Roodgar *et al.* [36]. While the approach of Roodgar *et al.* was appropriate for their purposes—namely, detecting selective sweeps of linked variants that deviated substantially from the overall background—our clustering scheme is designed to detect only large clusters of SNVs (minimally, greater than 1000 SNVs within a cluster) whose linkage patterns are consistent with perfect linkage on a single haplotype background. Our choice of 1000 SNVs as a cutoff was informed by previous work estimating the typical scale of genetic divergence between strains found in different hosts (See S1 Text section 3.1 for further information). While lineages can and do diverge as a result of diversifying evolution within hosts [34, 46], by imposing the minimum cluster size of 1000 core genome SNVs, we expect largely to exclude cases of within-host diversification. When no cluster of greater than 1000 SNVs was detected, only a single strain was inferred to be present.

Macroecology

To fit Taylor's Law (Figure 3A), we used the `polyfit` function from the python `numpy` library to fit a power law regression between the mean and variance of each strain's abundance distribution. To obtain the log rescaled Gamma AFD (Figure 3B), each strain's abundance distribution was first log rescaled, and then normalized to have 0 mean and unit variance. We then binned each strain's rescaled abundance distribution into 20 evenly spaced bins, and fit the Gamma AFD to the bin-wise mean across strains. To perform the Gamma AFD fit, we adapted code from [1].

Competing interests

The authors declare no competing interests.

Code availability

All necessary metadata, as well as the source code for the MIDAS metagenomic pipeline, downstream analyses, and figures, are available on GitHub: https://github.com/richardwolff/strain_stability.

Acknowledgments

We thank Colin Kremer for his crucial comments on this manuscript, Van Savage for early discussions on this work, Jacopo Grilli for code and important discussions, as well as members of the Garud lab for continued feedback. Lastly, we thank two anonymous reviewers for their constructive 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] Pearson, K. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. Philosophical Magazine. Series 5. 50 (1900), pp. 157-175

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

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

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

[7] Zeevi, David, et al. Structural variation in the gut microbiome associates with host health. Nature 568.7750 (2019): 43-48.

[8] Lynch, Jonathan B., et al. Turicibacter modifies host bile acids and lipids in a strain-specific manner. bioRxiv (2022).

[9] Gonze, Didier, et al. Multi-stability and the origin of microbial community types. The ISME journal 11.10 (2017): 2159-2166.

[10] Benz, Fabienne, and Alex R. Hall. Host-specific plasmid evolution explains the variable spread of clinical antibiotic-resistance plasmids. bioRxiv (2022).

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

[12] 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).

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

[14] 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).

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

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

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

[18] 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).

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

[20] Sender, Ron, Shai Fuchs, and Ron Milo. Revised estimates for the number of human and bacteria cells in the body. PLoS biology 14.8 (2016): e1002533.

[21] Coyte, Katharine Z., Jonas Schluter, and Kevin R. Foster. The ecology of the microbiome: networks, competition, and stability. Science 350.6261 (2015): 663-666.

[22] Relman, David A. The human microbiome: ecosystem resilience and health. Nutrition reviews (2012): S2-S9.

[23] Lozupone, Catherine A., et al. Diversity, stability and resilience of the human gut microbiota. Nature 489.7415 (2012): 220-230.

[24] Levy, Roie, et al. Longitudinal analysis reveals transition barriers between dominant ecological states in the gut microbiome. Proceedings of the National Academy of Sciences 117.24 (2020): 13839-13845.

[25] Hehemann, Jan-Hendrik, et al. Transfer of carbohydrate-active enzymes from marine bacteria to Japanese gut microbiota. Nature 464.7290 (2010): 908-912.

[26] 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.

[27] 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.

[28] Consortium HMP. A framework for human microbiome research. nature 486, no. 7402 (

[29] Lloyd-Price, J. *et al.* The integrative human microbiome project. *Nature* (2019).

[30] 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.

[31] Zheng, Wenshan, et al. High-throughput, single-microbe genomics with strain resolution, applied to a human gut microbiome. Science 376.6597 (2022): eabm1483.

[32] Tikhonov, Mikhail. Theoretical microbial ecology without species. *Physical Review E* 96.3 (2017): 032410.

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

[34] 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).

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

[36] 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).

[37] Fang, Xin, et al. Metagenomics-based, strain-level analysis of Escherichia coli from a time-series of microbiome samples from a Crohn's disease patient. Frontiers in microbiology (2018): 2559.

[38] 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).

[39] Kang, Jonathan TL, et al. Long-term ecological and evolutionary dynamics in the gut microbiomes of carbapenemase-producing Enterobacteriaceae colonized subjects. bioRxiv (2022).

[40] 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)

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

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

[43] Goldford, Joshua E., et al. Emergent simplicity in microbial community assembly. Science 361.6401 (2018): 469-474.

[44] Venturelli, Ophelia S., et al. Deciphering microbial interactions in synthetic human gut microbiome communities. Molecular systems biology 14.6 (2018): e8157.

[45] Daniel, Noëmie, Emelyne Lécuyer, and Benoit Chassaing. Host/microbiota interactions in health and diseases—time for mucosal microbiology!. Mucosal Immunology 14.5 (2021): 1006-1016.

[46] Zhao, S., *et al*. Adaptive evolution within gut microbiomes of healthy people. Cell host & microbe, 25(5), 656-667. (2019).

[47] 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.

[48] Korpela, Katri, et al. Selective maternal seeding and environment shape the human gut microbiome. Genome research 28.4 (2018): 561-568.

[49] 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

[50] 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

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

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

[53] 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

[54] 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

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

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

[57] Conwill, Arolyn, et al. Anatomy promotes neutral coexistence of strains in the human skin microbiome. bioRxiv (2021).

[58] Zaoli, Silvia, and Jacopo Grilli. A macroecological description of alternative stable states reproduces intra-and inter-host variability of gut microbiome. bioRxiv (2021).

[59] 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).

[60] Ji, Brian W., et al. Macroecological dynamics of gut microbiota. Nature microbiology 5.5 (2020): 768-775.

[61] 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.

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

[63] 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.

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

[65] Allen, Edward. Modeling with Itô stochastic differential equations. Vol. 22. Springer Science & Business Media, 2007.

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

[67] 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

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

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

[70] 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.

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

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

[73] Van Rossum, Thea, *et al.* Diversity within species: interpreting strains in microbiomes. Nature Reviews Microbiology 18.9 (2020): 491-506.

[74] Olm, Matthew R., et al. Identical bacterial populations colonize premature infant gut, skin, and oral microbiomes and exhibit different *in situ* growth rates. Genome research 27.4 (2017): 601-612.

[75] Kang, Jonathan TL, et al. Long-term ecological and evolutionary dynamics in the gut microbiomes of carbapenemase-producing Enterobacteriaceae colonized subjects. bioRxiv (2022).

[76] Wang, Zihan, et al. Fine-scale diversity of microbial communities due to satellite niches in boom-and-bust environments. bioRxiv (2022).

[77] Fuller, Wayne A. Introduction to statistical time series. John Wiley & Sons, 2009.

[78] Yassour, Moran, et al. Natural history of the infant gut microbiome and impact of antibiotic treatment on bacterial strain diversity and stability. Science translational medicine 8.343 (2016): 343ra81-343ra81.

[79] Quince, Christopher, et al. STRONG: metagenomics strain resolution on assembly graphs. Genome biology 22.1 (2021): 1-34.

[80] Lim, Joe J., Christian Diener, and Sean M. Gibbons. Growth phase estimation for abundant bacterial populations sampled longitudinally from human stool metagenomes. bioRxiv (2022).

[81] Feng, Jun, et al. Polysaccharide utilization loci in Bacteroides determine population fitness and community-level interactions. Cell Host & Microbe 30.2 (2022): 200-215.

[82] Jin, Shen, Daniela Wetzel, and Melanie Schirmer. Deciphering mechanisms and implications of bacterial translocation in human health and disease. Current Opinion in Microbiology 67 (2022): 102147.

[83] Seabold, S., Perktold, J. statsmodels: Econometric and statistical modeling with python. 9th Python in Science Conference (2010).

[84] Harris, C.R., Millman, K.J., van der Walt, S.J. et al. Array programming with NumPy. Nature 585, 357–362 (2020). DOI: 10.1038/s41586-020-2649-2