Molecular evolutionary dynamics of energy limited microorganisms

William R. Shoemaker,¹,²,∗ Evgeniya Polezhaeva,¹ Kenzie B. Givens,¹,³ and Jay T. Lennon¹,∗
¹Department of Biology, Indiana University, Bloomington, IN, 47405, USA
²Current affiliation: Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA, 90095, USA
³Current affiliation: Luddy School of Informatics, Computing, and Engineering, Indiana University, Bloomington IN, 47408, USA
∗Corresponding author: E-mail: williamrshoemaker@gmail.com, lennonj@iu.edu.
Associate Editor: Unassigned

Abstract

Microorganisms have the unique ability to survive extended periods of time in environments with extremely low levels of exploitable energy. To determine the extent that energy limitation affects microbial evolution, we examined the molecular evolutionary dynamics of a phylogenetically diverse set of taxa over the course of 1,000 days. We found that periodic exposure to energy limitation affected the rate of molecular evolution, the accumulation of genetic diversity, and the rate of extinction. We then determined the degree that energy limitation affected the spectrum of mutations as well as the direction of evolution at the gene level. Our results suggest that the initial depletion of energy altered the direction and rate of molecular evolution within each taxon, though after the initial depletion the rate and direction did not substantially change. However, this consistent pattern became diminished when comparisons were performed across phylogenetically distant taxa, suggesting that while the dynamics of molecular evolution under energy limitation are highly generalizable across the microbial tree of life, the targets of adaptation are specific to a given taxon.

Key words: Experimental evolution, microbial evolution, energy limitation, comparative evolution.

Introduction

Organisms frequently encounter environments where growth and reproduction cannot be maintained. This is particularly true for microorganisms, where growth and reproduction are often limited by the flux of exploitable energy (Hobbie and Hobbie, 2013; Hoehler and Jørgensen, 2013; Lever et al., 2015).

The occurrence of microorganisms in these environments is frequent enough that it has been argued that energy limitation constitutes a general constraint on microbial metabolism and growth (Maitra and Dill, 2015; Morita, 1988, 1993). Because evolution is fundamentally a birth-death process (Doebeli et al., 2017), if a microbial population does not go extinct the extent that energy limitation constrains reproduction would in turn affect its subsequent...
However, the evolutionary effect of energy limitation is not limited to the quantity and rate of change of genetic diversity. Rather, certain genes repeatedly acquire mutations in energy limited environments (Finkel, 2006; Finkel and Kolter, 1999; Katz et al., 2020; Kram et al., 2017; Orsi et al., 2021; Zambrano and Kolter, 1996; Zambrano et al., 1993; Zinser and Kolter, 2004). These repeated evolutionary outcomes (i.e., parallel evolution) consistently occur across phylogenetically distant taxa (Bacundruzdina et al., 2007; Bruno and Freitag, 2011; Helmus et al., 2012; Martínez-García et al., 2003; Ratib et al., 2021; Sewell et al., 2011), suggesting that adaptation continues as the environment becomes increasingly depleted of exploitable energy. Though such studies rarely manipulate the degree of energy limitation, contrast their observations to those from populations in energy replete environments, or perform comparisons across multiple taxa. These shortcomings prevent us from determining whether the targets of molecular evolution change as the environment transitions from an energy replete to depleted state (i.e., divergent evolution) as well as whether any such change in the direction of evolution can be generalized across the microbial tree of life.

In this study we examined the molecular evolutionary dynamics of energy limited microbial populations from six bacterial taxa that were propagated for ~1,000 days, which were chosen to maximize phylogenetic diversity subject to...
the condition that they all grew in the same environmental conditions (Bacillus, Caulobacter, Deinococcus, Janthinobacterium, Pedobacter, and Pseudomonas). We manipulated energy limitation by extending the time between transfer events, corresponding to transfer times of 1, 10, and 100 days, which allowed us to examine evolutionary outcomes across a range of energy limitation regimes as well as taxa. We examined how genetic diversity, the spectrum of mutations, and the distribution of extinction events were affected by energy limitation. We then examined the degree that energy limitation affected the direction of molecular evolution across all taxa. Finally, we determined whether phylogenetically diverse taxa converged on similar molecular evolutionary outcomes within and across energy limitation regimes.

Results

All 90 replicate populations from six phylogenetically diverse taxa were transferred every 1, 10, or 100 days for 1,000 days (Fig. 1). While extinction events occurred, we were able to reconstruct the molecular fossil record of virtually all populations through pooled population sequencing, allowing us to examine how mutations arose and changed in frequency over time. We found that molecular evolutionary dynamics were consistently affected by transfer time, our proxy for energy limitation, across phylogenetically diverse taxa. In energy depleted environments, mutations accumulated at a faster rate than expected based on the timescale of the experiment across all taxa, though the molecular evolutionary dynamics of 10-day and 100-day regimes were more similar to one another than either were to the 1-day regime. We found that the spectrum of mutations was consistently affected by energy limitation across all taxa, though the effect on the proportion of nonsynonymous and synonymous polymorphisms (pN/pS) was less consistent, suggesting that the strength of negative selection did not change as energy limitation increased. By comparing the genes that acquired excess nonsynonymous mutations, we found a consistent pattern of divergent evolution among all energy limitation regimes and taxa, where, again, the 10 and 100-day transfer regimes had the highest degree of similarity. However, this pattern of divergent evolution was absent when comparisons were made at a coarser phylogenetic scale, suggesting that it is unlikely that the targets of molecular evolution under energy limitation within a taxon can be generalized across the tree of life. The recurrent patterns we observed within a given taxon lead us to conclude that the rate and direction of molecular evolution consistently shifted across all taxa after energy depletion occurred but remained relatively constant thereafter.

Dynamics and fates of mutations under energy limitation

Microbial populations continue to acquire genetic diversity under energy limited conditions (Finkel
and Kolter, 1999), though their quantitative dynamics are rarely examined. To address this, we constructed a molecular fossil record for \( \sim 100 \) populations that evolved for \( \sim 1,000 \) days (Figs. S1-6). By examining the accumulation of mutations by time \( t \) as the sum of derived allele frequencies \( M(t) \), we found that the accumulation of mutations was fairly consistent across energy limitation regimes, where \( M(t) \) typically saturated by day 500 (Figs. S7-12,a-c). While population genetic models have not been proposed to describe the trajectory of \( M(t) \) over time, we can provide an intuitive explanation for the saturating behavior we observe. All populations were grown from single clonal isolates, meaning that \( M(t) \) was initially zero and we would expect it to rapidly increase as \emph{de novo} mutations are acquired. Because the population is finite, a finite number of mutations can segregate at a given point in time, suggesting that saturation occurs when the increase in \( M(t) \) due to mutation and positive selection is roughly balanced by the decrease in \( M(t) \) due to negative selection and drift. A similar consistency across treatments and taxa was found for the change in \( M(t) \) between timepoints (\( \Delta M(t) \); Figs. S7-12,d-f), suggesting that these general molecular evolutionary patterns were not affected by energy limitation.

By examining the relationship between \( M(t) \) and the degree of energy limitation, we were able to determine the degree that different taxa continued to acquire mutations under energy limitation. If cell division completely stopped once the environment was exhausted of energy and the majority of \emph{de novo} mutations were acquired due to genome replication errors, then the number of mutations \( M(t) \) should decay as the time between transfers increased. All taxa deviated from this prediction, where the slope was consistently greater than the null expectation \( \beta_1 > -1 \); Fig. 2a-f). This deviation was extreme in certain cases, as the slopes of \emph{Pseudomonas} and \emph{Pedobacter} were virtually zero, indicating that the amount of accumulated genetic diversity remained highly similar despite the wide range of energy limitation regimes. Conversely, the slope of \emph{Bacillus} was closest to the null expectation. Surprisingly, very few mutations appeared to fix over the course of 1,000 days, even in populations that were transferred daily (Figs. S1-6). Using a previously developed Hidden Markov Model-based approach (Good et al., 2017), we confirmed this observation by assigning each mutation a final state (i.e., fixed, extinct, or polymorphic; Figs. 2g, S7-12). The comparative absence of fixations made it difficult to compare substitution rates between treatments and taxa due to the excess number of zeros. However, some general conclusions could still be made. Beyond the low number of fixation events, only \emph{Pedobacter} and \emph{Bacillus} experienced a sharp decline in fixation events as the time between transfers increased (Fig. 2g). While there is no obvious explanation for the disparity in the former, the sharp decline
Molecular evolutionary dynamics of energy limited microorganisms

in *Bacillus* was likely driven, again, by its ability to form protective endospores.

Despite the lack of fixations, it was clear that few mutations reached frequencies $\geq 0.4$ in 10 and 100-day transfer regimes, even if the change in $M(t)$ remained constant (Figs. S7-12). Leveraging this observation, we examined the distribution of the maximum observed frequency of all mutations ($f_{\text{max}}$) to infer whether the trajectories of mutations varied among energy limitation regimes. We found that the distribution of $f_{\text{max}}$ in the 1-day transfer regime was clearly distinct from the 10 and 100-day regimes, while the latter two regimes largely overlapped (Fig. 2h). Direct comparison of $f_{\text{max}}$ cumulative distributions confirmed this, as the mean Kolmogorov–Smirnov distance between all combinations of energy limitation regimes was significant (permutational ANOVA: $F = 1.63, P = 0.0430$). The mean distance remained high for 1 versus 10-days and 1 versus 100-days comparisons, but notably dropped for 10 versus 100-days comparisons (Fig. 2h). This drop in $f_{\text{max}}$ suggests that the trajectory of *de novo* mutations was altered after the initial depletion of exploitable energy, where nominally beneficial mutations could not continue to increase in frequency.

We note the absence of any segregation of intermediate-frequency mutations (i.e., clade formation) across all taxa and transfer regimes. This observation seemingly contrasts with the repeated emergence of quasi-stable clades in an experiment with similar energy limitation regimes conducted over a similar evolutionary timescale with *E. coli* (Behringer et al., 2018, 2019) as well as the proposal that the emergence of clades is a general feature of microbial evolution (Good and Hallatschek, 2018). The lack of structure in our experiment was likely caused by differences in experimental design, as our experiment was performed in well-mixed flasks rather than test tubes with a lack of headspace (Behringer et al., 2018, 2019), meaning that there was no spatial structure that could promote the emergence of population genetic structure. Though experimental design does not entirely account for the lack of structure, as other evolution experiments using well-mixed flasks have documented the emergence of co-existing clades, the Long-term Evolution Experiment (LTEE) with *E. coli* being the most notable example (Good et al., 2017). However, the average timescale for the emergence of new clades in the LTEE was $\sim 16,000$ generations, far

Table 1. The distribution of extinction events among replicate populations suggest historical contingency due to the acquisition of *de novo* mutations in certain taxa within the 10-day transfer regime. By comparing the coefficient of variation (CV) of the number of extinction events to a null distribution calculated from simulations using the mean number of extinction events ($\pi_{\text{ext}}$) we determined whether historical contingency occurred within a given taxon-treatment combination. All taxon-treatment combinations where at least three replicate populations went extinct were examined.

| Taxon              | Transfer (d.) | $\pi_{\text{ext}}$ | CV   | $P_{\text{BH}}$ |
|--------------------|---------------|---------------------|------|---------------|
| Caulobacter        | 10            | 1.2                 | 0.97 | 0.409         |
|                    | 100           | 0.6                 | 0.82 | 0.837         |
| Pedobacter         | 10            | 5.4                 | 0.74 | 0.03          |
|                    | 100           | 3.2                 | 0.31 | 0.837         |
| Janthinobacterium  | 10            | 7.6                 | 0.75 | 0.005         |
|                    | 100           | 1.0                 | 1.1  | 0.409         |


longer than our longest experimental timescale of \( \sim 3,000 \) generations. This disparity in evolutionary timescales suggests that while energy limitation alone is insufficient to promote the emergence of multiple clades on short evolutionary timescales, we cannot rule out the possibility of clades eventually emerging if the experiment had continued.

The molecular spectra of \textit{de novo} mutations

We determined the degree that energy limitation affected the spectrum of mutations by examining the transition rates of all mutations at four-fold redundant sites. By performing dimensionality reduction on a population-by-spectra matrix (i.e., PCA), we found that replicate populations largely grouped by their energy limitation regime (Fig. S13). A permutational multivariate analysis of variance (PERMANOVA) test using a permutation structure that controlled for taxon identity confirmed this observation \((F = 5.77, P < 10^{-4})\). By examining the factor loadings, we found the nucleotide spectra that are correlated with the first and second principal components (Table S1). The transversion A:T\( \rightarrow \)C:G was highly correlated with the first principle component \((\rho_{\text{Pearson}} = 0.91, P < 10^{-4})\) while the transition A:T\( \rightarrow \)G:C was correlated with the second \((\rho_{\text{Pearson}} = 0.76, P = 0.002)\), suggesting that these mutational spectra were affected by energy limitation.

Expanding our scope from synonymous mutations, we examined the spectrum of nonsynonymous and synonymous mutations to determine the degree that positive and negative selection change as energy limitation increases. Given the lack of fixation events within populations in 10 and 100-day transfer regimes, our comparative analyses were limited to polymorphic mutations. However, we were able to leverage the large number of mutations we observed to compare the ratio of nonsynonymous to synonymous polymorphisms \((pN/pS)\) between transfer regimes, which we examined as a function of \(f_{\text{max}}\) rather than as a single snapshot so that we could control for the qualitatively different mutation trajectories found in each energy limitation regime. In general, we found that \(pN/pS\) remained considerably less than 1 as \(f_{\text{max}}\) increased, 1-day \textit{Janthinobacterium} being the sole exception and only for \(f_{\text{max}} > 0.4\) (Fig. 3c). This result is consistent with the hypothesis that negative selection predominantly occurs across microbial genomes (Garud \textit{et al.}, 2019; Hahn, 2018). To determine whether the relationship between \(f_{\text{max}}\) and \(pN/pS\) varied by energy limitation regime, we calculated the mean absolute deviation (MAD) for each taxon across all pairs of energy limitation regimes. We found that the mean absolute deviation (± standard error) in \(pN/pS\) was greater between the 1-day transfer regime and the 10 and 100-day transfer regimes \((\text{MAD}_{1,10} = 0.130 \pm 0.0262, \text{MAD}_{1,100} = 0.133 \pm 0.0251)\) than for the latter two regimes \((\text{MAD}_{10,100} = 0.0785 \pm 0.0166)\). This
pattern is consistent with the observation that the molecular evolutionary dynamics were affected by the initial depletion of energy and remained relatively constant as the degree of energy limitation increased. However, through a permutational ANOVA, we found that the difference in MAD over our three treatment comparisons was not significant ($F = 1.12$, $P = 0.446$), suggesting that any change in the strength of negative selection as energy limitation increased was weak relative to the overall strength of negative selection operating across all regimes.

While a complete explanation for the expected relationship between $f_{max}$ and $pN/pS$ is beyond the goals of this study, it is worth discussing the notable "U" shape exhibited by many taxon-treatment combinations (Fig. 3). Assuming that a typical nonsynonymous mutation confers a fitness benefit, we would expect nonsynonymous mutations to reach a higher frequency than those that are neutral. Suppose that nonsynonymous mutations are primarily beneficial. In this scenario, as we increase $f_{max}$, we would expect $pN/pS$ to also increase. Contrastingly, suppose instead that nonsynonymous mutations are primarily deleterious. We would then expect that the typical nonsynonymous mutation would reach a lower frequency than the typical neutral mutation, as it cannot reach a frequency greater than the inverse of the population scaled strength of selection $1/N|s|$ known as the drift barrier (Cvijović et al., 2018; Lynch et al., 2016). In this scenario, as we increase $f_{max}$ we expect $pN/pS$ to decrease. Putting these two pieces together, we have constructed a heuristic explanation for the U-shape of $f_{max}$ vs. $pN/pS$, where the lower point of the curve reflects a transition from deleterious to beneficial effects among nonsynonymous mutations. A rigorous explanation for this empirical pattern is a potential avenue of future research.

The majority of populations survived the experiment, though many Caulobacter, Janthinobacterium, and Pedobacter replicate populations repeatedly went extinct. Because all replicate populations were initially genetically identical, the dispersion of extinction events within a treatment contained information about whether the rate of extinction was contingent on a population’s evolutionary history. If the per-generation rate of extinction was sufficiently rare and an independent process, we would expect that the distribution of extinction events would follow a Poisson distribution. However, if mutations acquired by a given population over the course of the experiment affected the probability of extinction (i.e., historical contingency), the distribution should be overdispersed relative to a Poisson. By comparing the coefficient of variation calculated from empirical data (CV) to a null distribution generated from Poisson sampling, we found that extinctions in the 10-day transfer regime was overdispersed for Pedobacter and Janthinobacterium (Table 1). This result suggests...
that mutations acquired over the course of the experiment contributed to repeated extinction events in at least two taxa. It is worth noting that these two taxon-treatment combinations had the highest mean number of extinction events ($\pi_{\text{ext}}$), suggesting that historically contingent extinction events regularly occurred under energy limitation across phylogenetically distant taxa, but that an insufficient number of extinction events occurred in our experiment to detect said signal in all taxon-treatment combinations.

Parallel, divergent, and convergent evolution within each taxon

To identify potential targets of adaptation, we examined the set of genes that acquired more nonsynonymous mutations than expected based on gene size within each energy limitation regime for each taxon (i.e., genes that contributed to parallel evolution). We first calculated the number of nonsynonymous mutations pooled across replicate populations at each gene after adjusting for gene size (i.e., multiplicity (Good et al., 2017)) and tested whether nonsynonymous mutations were non-randomly distributed across genes (Supplementary Information). We easily rejected this null hypothesis for all taxon-treatment combinations ($P < 10^{-4}$; Figs. S15-20, a). We note the presence of a small cluster of genes where the mean maximum observed allele frequency was fairly high ($\bar{f}_{\text{max}} > 0.5$), but multiplicity was low (Figs. S14-19, b). The formation of this cluster was likely driven by the fact that few mutations reached appreciable frequencies, which would decrease multiplicity given that it is weighted by total number of mutations in the population (Supplementary Information). However, few genes acquired more than one high frequency mutation, meaning that we were unable to identify the set of significantly enriched genes conditioned on a minimum $f_{\text{max}}$. Instead, we determined whether parallelism at the gene level was dependent on allele frequency by selecting for mutations with a minimum value of $f_{\text{max}}$ and calculating the likelihood difference of the enrichment of nonsynonymous mutations at certain genes across the genome versus no such enrichment ($\Delta \ell$). We found that $\Delta \ell$ typically increased with $f_{\text{max}}$ across energy limitation regimes for the majority of taxa (Fig. S14), which is consistent with the hypothesis that higher frequency mutations were primarily driven by positive selection across energy limitation regimes. A list of genes with at least one fixed mutation and the number of mutations with $f_{\text{max}} \geq 0.8$ (excluding synonymous mutations) for each taxon-treatment combination is provided in Table S2.

Using the set of genes that contributed to parallel evolution within each energy limitation regime for each taxon, for a given pair of energy limitation regimes we asked whether the evolved features of these genes were more similar or different than expected by chance. Operating under this broad definition of convergent and
Molecular evolutionary dynamics of energy limited microorganisms · doi:XX.XXX

MBE

divergent evolution at the molecular level, we first examined whether the same genes were consistently enriched. Across taxa, the vast majority of genes that were enriched within a given treatment were enriched across all treatments (Figs. S21-25), which would suggest that convergent rather than divergent evolution occurred. Simulations modeled on the probability distribution describing the null expectation (Eq.1) confirmed this observation, as the proportion of genes that were enriched within both energy limitation regimes for a given pair was consistently greater than expected by chance for all pairwise combinations of energy limitation regimes across all taxa (Fig. 4a; Table S3). However, the degree of convergent evolution across pairs of energy limitation regimes was not equal, as the 10 versus 100-day comparison had a stronger degree of convergence than all other comparisons (Fig. 4a). This result builds on our above observations, providing evidence that the largest shift in the direction of evolution occurred shortly after the environment was depleted of exploitable energy.

While our definition of convergent/divergent evolution as the proportion of genes that were enriched across energy limitation regimes provided insight, it was likely too coarse an approach as it only considered gene identity. Assuming that each energy limitation manipulation acted as a slight perturbation on the rate of molecular evolution at each gene, gene identity alone would likely be insufficient to detect divergent evolution and would generally suggest convergent evolution of varying degrees. A direct comparison of mutation counts supports this claim, as the set of genes that were enriched across multiple energy limitation regimes consistently differed in the number of mutations acquired (Figs. S15-20, d-f). To quantify this potential divergence, we defined convergent and divergent evolution as the degree that the correlation of mutation counts between energy limitation regimes (ρ) was greater or less than expected by chance. To determine whether a given value of ρ constituted convergent or divergent evolution, we generated a null distribution by randomizing combinations of mutation counts while controlling for the total number of mutations acquired within each treatment and gene. By examining the standardized ρ (Zρ) of each taxon, we found that divergent evolution consistently occurred across all combinations of energy limitation regimes (Fig. 4a-b). Similar to our analysis on gene identity, we found that the mean Zρ across taxa was closest to the null expectation for comparisons between 10 and 100-day transfers (Fig. 4c). The difference in the degree of divergent evolution was statistically significant (F = 3.13, P = 0.0439; Fig. 4d) and what we observed was consistent with the reoccurring pattern of the evolutionary dynamics in the 10 and 100-day energy limitation regimes having greater similarity to one another than either one did to the 1-day regime.
Divergent and convergent evolution across taxa

While divergent and convergent evolution could readily be examined among populations within the same taxon, a similar analysis could not be performed among all phylogenetically diverse taxa due to the lack of shared gene content. Instead, it was necessary to map genes that were enriched for mutations to higher levels of biological organization that were present in all taxa. Therefore, we mapped genes to their functional modules using the Metabolic And Physiological potentiaL Evaluator (MAPLE) (Arai et al., 2018), where we found several modules that were consistently hit across three or more taxa. To test whether this observation was significant, we simulated random draws of genes that were then mapped to their respective modules, from which we generated null rank curves for the number of modules that were hit across all possible combinations of taxa. We found that the observed decay curve deviated from our null expectation (Fig. 5a), but only up until four taxa for 1-day and 100-day transfers. This pattern suggests that convergent evolution did occur among taxa within the limits of energy repletion (1-day) and depletion (100-days), though the ability to detect this convergence dissipated as additional taxa were included. This reduced rate of decay was driven, at least in-part, by shared evolutionary history, as the similarity in enriched MAPLE modules decays with phylogenetic distance for 1 and 100-day transfer regimes (Fig. 5d). We note that while this trend was only significant for 1-day transfers, the slope for the 10-day treatment was virtually zero and the lack of a significant slope in the 100-day treatment can be explained by the low number of nonsynonymous mutations acquired by populations within that treatment, limiting our ability to identify enriched MAPLE modules and inflating the number of observations with zero overlap.

While there were consistent patterns of convergent evolution across phylogenetically diverse taxa within the 1-day and 100-day energy limitation regimes, there remained the question of whether each regime acquired nonsynonymous mutations in a distinct set of modules. In other words, it was necessary to determine whether energy limitation continued to affect the direction of molecular evolution when all taxa are examined together instead of on an individual basis. By constructing a module-by-taxon/treatment matrix and performing PCA, we found that taxon-treatment combinations tended to cluster together by taxon identity rather than their energy limitation regime (Fig. 5e). A PERMANOVA test supports this observation (Fig. 5e). While there was no evidence of genome-wide divergent evolution between energy limitation regimes when comparisons were made at a higher phylogenetic level, we identified individual modules that may be of interest to future research efforts (Table 2). Overall,
Table 2. The number of taxa with genes within a MAPLE module that were enriched for nonsynonymous mutations. Only modules that were enriched in at least two taxa within a given treatment regime were included. There is a large amount of overlap between transfer-regimes, suggesting that highly conserved genes contributed towards adaptation in similar ways in energy deplete and replete environments.

| ModuleID | Module type | Description                              | 1-day Number of taxa enriched/ Total number of taxa | 10-days | 100-days |
|----------|-------------|------------------------------------------|-----------------------------------------------------|---------|---------|
| M00307   | Pathway     | Pyruvate oxidation                       | 2/6/1/5                                             | 2/5     | 3/6     |
| M00350   | FuncSet     | Aminoacyl-tRNA biosynthesis               | 3/6/1/5                                             | 2/5     | 3/6     |
| M00258   | Complex     | Putative ABC transport system             | 4/6/2/5                                             | 2/5     | 3/6     |
| M00049   | Pathway     | Adenine ribonucleotide biosynthesis       | 4/6/2/5                                             | 2/5     | 3/6     |
| M00083   | Pathway     | Fatty acid biosynthesis                   | 3/6/1/5                                             | 1/5     | 1/6     |
| M00125   | Pathway     | Riboflavin biosynthesis                   | 3/6/1/5                                             | 1/5     | 1/6     |
| M00048   | Pathway     | Fatty acid biosynthesis                   | 3/6/1/5                                             | 1/5     | 3/6     |
| M00120   | Pathway     | Coenzyme A biosynthesis                   | 2/6/0/5                                             | 0/5     | 0/6     |
| M00240   | Complex     | Iron complex transport system             | 2/6/2/5                                             | 2/5     | 3/6     |

there was a high degree of overlap among energy limitation regimes, where Coenzyme A biosynthesis was the only module enriched in at least two taxa within a single transfer regime. Iron complex transport systems also appeared to be slightly more enriched in 100-day transfer regimes, a micronutrient that has been shown to be essential for the survival of microorganisms in energy limited conditions (Helmus et al., 2012; Lacour and Landini, 2004).

Discussion

We found that the molecular evolutionary dynamics of phylogenetically distant taxa were consistently altered as energy limitation increased. In general, genetic diversity accumulated to a higher level than what we would expect if cellular division stopped after populations exited their exponential phase of growth after a single day. This consistent deviation suggests that some degree of cellular division continued while the population remained in an energy limited state, as it is doubtful that mutation alone would be capable of driving \textit{de novo} mutations to detectable frequencies. While the biological mechanism responsible for this continued reproduction cannot be identified from mutation trajectory data, it is likely due to surviving cells scavenging dead cells, the only available resource, for energy (Bradley et al., 2018; Shoemaker et al., 2021). This hypothesis is supported by the observation that the slope of \textit{Bacillus} was closest to the null out of all taxa examined, which was likely due to \textit{Bacillus} being the only taxon capable of forming non-reproductive endospores that buffer the accumulation of \textit{de novo} genetic diversity (Blath et al., 2016; Shoemaker and Lennon, 2018). Though \textit{Pseudomonas} and \textit{Pedobacter} were noted outliers, where energy-limited populations accumulated similar quantities of genetic diversity as energy-replete populations. A prior study found evidence that these taxa are capable of surviving energy-limitation and reproducing by recycling dead cells, which may explain how genetic diversity continued to be acquired after
the populations exited their exponential phase of growth and the extreme values exhibited in *Pseudomonas* and *Pedobacter* populations (Shoemaker et al., 2021). This dynamic may be further compounded by the accumulation of mutations between replication events (Gangloff and Arcangioli, 2017; Gangloff et al., 2017).

Comparisons of the molecular dynamics of different energy limitation regimes revealed that the frequency trajectories of mutations and the spectrum of nonsynonymous and synonymous polymorphic mutations $pN/pS$ had higher similarity between 10 and 100-day regimes than either one did to the 1-day regime. A similar pattern was also found in our analyses of the direction of molecular evolution at the gene-level, where the degree of divergent evolution was at its lowest for comparisons made between 10 and 100-day regimes across taxa. Furthermore, the nucleotide mutation spectra of energy-limited populations was consistently altered across taxa, suggesting that mutations acquired between transfer events contributed towards the patterns of diversity we observed. These altered dynamics are likely partially due to the effects of stress in energy-depleted environments on the mutation spectrum, which has been noted in prior studies (Chib et al., 2017; Maharjan and Ferenci, 2015, 2017, 2018; Saumaa et al., 2002; Shoemaker and Lennon, 2018). This recurring pattern of populations in 10-day and 100-day treatments being more similar to one another than either were to the 1-day treatment suggests that the main effect of energy limitation on molecular evolution occurred shortly after the populations exited their exponential phase of growth, an observation consistent with recent findings documenting the physiological and cellular changes that microorganisms incur after energy depletion and may be relevant to clinical contexts (Kram et al., 2020).

Despite the generality of the effect of energy limitation on the molecular evolutionary dynamics within a given taxon, the signal of this pattern dissipated when comparisons were made between phylogenetically distant taxa. By coarse-graining genes to a higher level of functional annotation, we found that convergent evolution occurred among phylogenetically distant taxa within certain energy limitation regimes. However, the same biological functions contributed to these signals of convergent evolution in repleted and depleted environments. Alternatively stated, while phylogenetically distant taxa acquired mutations at the same molecular targets within certain energy limitation regimes and divergent evolution occurred between energy limitation regimes within each taxon, the statistical signal of divergent evolution was lost when comparisons were made among all taxa. This lack of signal suggests that while energy limitation can consistently alter the rate and direction of molecular evolution in predictable ways,
the targets of adaptation likely vary across phylogenetically distant taxa.

The lack of biological functions enriched for mutations within a specific energy limitation regime casts doubt on whether universal contributors towards adaptation can be identified by comparing evolve-and-resequence experiments from phylogenetically distant taxa. The absence of such contributors contrasts with the consistent effect that energy limitation has on evolutionary dynamics within a given taxon. This supposed contradiction can be resolved by accounting for the observation that energy limitation effects birth-death dynamics in a consistent manner across the microbial tree of life (Shoemaker et al., 2021), while proposing that different cellular mechanisms contribute towards adaptation within a given taxon. To compensate for this lack of signal, we argue that future efforts to investigate adaptation in energy limited environments should focus on characterizing the cellular effects of mutations that reach high frequencies. This approach, redolent of the framework of evolutionary cell biology (Lynch and Trickovic, 2020; Lynch et al., 2014), would circumnavigate the issue of annotation conservation and allow researchers to instead examine conserved cellular features that contribute to adaptation across phylogenetically distant taxa.

Methods

Long-term evolution under energy limitation

Energy limitation was manipulated by extending the time between transfers for microbial populations. We performed our energy limited evolution experiments using the following microbial taxa: *Bacillus subtilis* NCIB 3610 (ASM205596v1), *Caulobacter crescentus* NA1000 (ASM2200v1), *Deinococcus radiodurans* BAA-816 (ASM856v1), *Janthinobacterium* sp. KBS0711 (ASM593795v2), *Pedobacter* sp. KBS0701 (ASM593864v2), *Pseudomonas* sp. KBS0710 (ASM593804v2). A single colony was isolated from each taxon and grown in 10 mL of PYE media with 0.2% glucose and 0.1% casamino acids (2 g bactopeptone, 1 g yeast extract, 0.3 g MgSO₄ x 7 H₂O, 2 g dextrose, 1 g casamino acids in 1 L E-Pure™ H₂O with 50 mL of a 10 g/mL cycloheximide solution added after autoclaving to prevent fungal contamination) in a 50 mL Erlenmeyer flask. A cultured flask of each taxon was transferred once a day for two days before being split between 15 flasks. Five replicates of each taxon were transferred as 1 mL aliquots into 9 mL of medium every 1, 10, or 100 days for ~1,000 days. All flasks were organized in a 25°C shaker at 250 RPM such that no two populations of the same taxon were adjacent. One and 10 day treatments were cryopreserved in 20% glycerol at -80°C on days 30, 60, and 100 in 100-day increments, while 100-day treatments were cryopreserved every 100 days. Every 100
days biomass was cryopreserved for sequencing, where biomass was centrifuged, decanted and flash frozen with liquid nitrogen and placed into a -80°C freezer until DNA extraction was performed. We tested for contamination by plating media and examining colonies every 100 days, as all taxa had distinct colony coloration, colony morphology, and media coloration. The sole contamination we found was the growth of airborne fungi during the first 100 days of our experiment, which we prevented by adding cycloheximide to the media and adding a UV light to the air filter of the shaker.

Given that $\log_2(10) \approx 3.3$ generations per-transfer, over a maximum of 1,000 days this experiment amounted to an evolutionary timescale of 3,300, 330, and 33 generations for 1-, 10-, and 100-day transfers, respectively. While we tested all taxa to make sure that they could survive 10 and 100 days without media replenishment, two 100-day *Pedobacter* populations and four 10-day *Janthinobacterium* populations repeatedly went extinct starting at day 400. Because of this reduction in temporal resolution we excluded 10-day *Janthinobacterium* populations from our analyses.

### Signals of historical contingency in the distribution of extinction events

We compared, the coefficient of variation (CV) calculated from the observed distribution of extinction events to its null. Under a Poisson distribution, the coefficient of variation is $CV_{\text{Poi}} = \sqrt{\lambda}/\lambda = \lambda^{-1/2}$ and observed values of CV greater than this null indicate that extinction events are overdispersed. Using the mean number of extinction events $\pi_{ext}$ as an estimate of $\lambda$ for all taxon-treatment combinations where at least three replicate populations went extinct, we generated null Poisson distributed extinction events using *numpy* v1.16.4. Multiple testing correction was performed for the $P$-values using the Benjamini–Hochberg procedure at $\alpha = 0.05

### Phylogenetic reconstruction

The 16S rRNA gene of each taxon was amplified via PCR using 8F and 1492R primers and reaction conditions previously described (Lennon et al., 2012). PCR products were purified using the QIAGEN QIAquick PCR Purification Kit and Sanger sequenced at the Indiana Molecular Biology Institute (IMBI) at Indiana University Bloomington (IUB). Sequences were aligned using SILVA INcremental Aligner v1.2.11 (*SINA*) (Pruesse et al., 2012). Phylogenetic reconstruction was performed using Randomized Axelerated Maximum Likelihood v8.2.11 (*RAxML*) (Stamatakis, 2014) with the General Time Reversible model, gamma distributed rate variation, and bootstrap convergence criteria set to autoMRE. The 16S rRNA sequence of *Prochlorococcus marinus* subsp. marinus str. CCMP1375 (NC_005042) was used as an outgroup. Phylogenetic distance was calculated using the ETE Toolkit (Huerta-Cepas et al., 2016).
Library construction and sequencing

Cryopreserved biomass samples were mixed with 1 mL of a 50 mg/mL lysozyme (Dot Scientific DSL8100-10) E-PureTM H2O solution and incubated at 37°C for 60 min. DNA extractions were performed using Qiagen DNeasy UltraClean Microbial Kits. Libraries were constructed out of 1,337 samples. Approximately 30% of Illumina libraries were constructed in-house with the Nextera DNA Library Preparation Kit (Illumina, FC-121-1030) using a protocol based off of previously published protocols (Baym et al., 2015; Picelli et al., 2014). These libraries were sequenced as paired-end reads on an Illumina HiSeq 2500 at the University of New Hampshire Hubbard Center for Genome Studies. All remaining samples were constructed as Illumina Nextera DNA paired-end 2 x 150bp libraries by the Center for Genomics and Bioinformatics at Indiana University and sequenced on a Illumina NextSeq 300. A target depth of 100-fold coverage was set for all libraries, though we ultimately obtained a mean coverage of 195.0 ± 6.67 across samples.

Variant calling and analysis

The first 20 bp of all reads were trimmed and all read pairs where at least one pair had a mean Phred quality less than 20 were removed cutadapt v1.9.1 (Martin, 2011). Candidate variants were identified using a previously published approach (Good et al., 2017). We provide a brief summary of the method in the supplement, including a high-level overview of the procedure, modifications made to accommodate our experimental design, and key parameter settings. To estimate the frequency of the mth mutation in the pth population for all significant candidates, we used the naive estimator \( \hat{f}_{pm} = A_{pm}/D_{pm} \), where \( A \) is the coverage of the alternative allele and \( D \) is the total depth of coverage. We examined the accumulation of mutations by time \( t \) as the sum of derived allele frequencies \( M(t) = \sum_m \hat{f}_{mt} \) for each population. The change in \( M(t) \) between two timepoints was defined as \( \Delta M(t) = \log_{10}[M(t)/M(t-1)] \). The maximum observed frequency of a given mutation over \( T \) observations was defined as \( f_{max} = \max\{f(t):t=1,...,T\} \), where \( f(t) \) is the mutation frequency at time \( t \). Nucleotide transitions for all six spectra were counted at all mutations that occurred at four-fold redundant sites. The ratio of nonsynonymous to synonymous polymorphic mutations \( pN/pS \) was corrected by the number of nonsynonymous and synonymous sites within a genome. We examined \( pN/pS \) for values of \( f_{max} \) that were found across all taxa and treatments.

Principal component analysis was performed on the \( n \times p \) matrix \( X \) containing \( p \) relativized nucleotide transitions for all \( n \) populations using scikit-learn v0.23.1 (Pedregosa et al., 2011). Significant factor loadings were determined by creating permutation arrays \( \pi = (\pi_1,\pi_2,...,\pi_p) \), where \( \pi_j \) is a permutation of the \( n \) populations for the \( j \)th transition. This permutation array generates the permuted data matrix \( X_\pi \), on
which PCA was performed and null factor loadings were calculated. We performed 10,000 iterations and P-values were calculated by comparing observed factor loadings to their null distribution. A permutational multivariate ANOVA (PERMANOVA) was performed using a modified pseudo-F statistic to account for unequal covariance (Anderson et al., 2017). We used a permutation matrix that permuted treatment labels within each taxon, allowing us to control for taxon identity.

Within-taxon parallel and convergent/divergent evolution

For the purpose of this analysis, genetic parallelism is defined as the degree that the distribution of observed nonsynonymous mutation counts across genes deviates from the null expectation account based on gene size (i.e., multiplicity). Genes with significant multiplicity were identified using a previously published approach (Good et al., 2017). To briefly summarize, we begin by calculating the multiplicity of a gene of size \( L_i \) with \( n_i \) mutations of size as \( m_i = n_i \cdot L_i \), where \( L_i \) is the mean gene size in the genome. Under our null hypothesis, all genes have the same multiplicity \( m = \frac{n_{\text{total}}}{N_{\text{genes}}} \). We then quantify the net increase of the log-likelihood of the alternative hypothesis relative to the null across the genome as \( \Delta \ell = \sum_i n_i \log \left( \frac{m_i}{m} \right) \).

Divergence and convergence as the presence or absence of enriched genes across energy limitation regimes within a given taxon was tested by determining whether the size of the intersection of the set of genes with significant multiplicity across treatments within a taxon was greater than (convergence) or less than (divergence) expected by chance. This can be performed by extending the hypergeometric distribution to the multivariate case, where the probability that \( m \) treatments have \( k \) intersecting genes is

\[
P(|I_1 \cap I_2 \cap \cdots \cap I_m| = k) = \frac{\binom{N_{\text{genes}}}{k}}{\prod_{i=1}^m \binom{N_{\text{genes}}}{|I_i|}} \times \sum_{j=0}^{\min\{|I_1|, |I_2|, \ldots, |I_m|\} - k} (-1)^j \frac{N_{\text{genes}} - k}{j} \times \prod_{l=1}^m \frac{N_{\text{genes}} - j - k}{|I_l| - j - k}
\]

(1)
Molecular evolutionary dynamics of energy limited microorganisms

Building off of prior results (Shoemaker and Lennon, 2020), the degree of divergent versus convergent evolution among enriched genes was determined using the squared Pearson’s correlation of relative multiplicities $(\mu_i = m_i/\sum m_i)$ for each pair of energy limitation regimes. Null correlation coefficients were calculated by constructing a gene-by-regime mutation count matrix for each pair of energy limitation regimes within each taxon and randomizing combinations of mutation counts constrained on the total number of mutations acquired within each gene across treatments and the number of mutations acquired within each treatment (i.e., the row and column sums of the matrix). This randomization was performed using a Python implementation (Baak et al., 2019) of the ASA159 algorithm (Patefield, 1981). Correlation coefficients were calculated from the randomized matrices 10,000 times and the observed coefficients were standardized relative to their respective null distributions $(Z_\rho)$.

Among-taxa convergent evolution

Because few genes are present among all of the taxa we chose, it is difficult to determine whether the same gene acquired more mutations than expected by chance across multiple taxa (i.e., evolutionary convergence among taxa within the same energy limitation regime). To determine whether evolutionarily distant taxa acquired similar mutations, we grouped genes at higher levels of biological organization. To do this, we first inferred the metabolic pathway composition of all taxa using the Metabolic And Physiological potentialEvaluator (MAPLE v2.3.1; (Arai et al., 2018)). MAPLE was run using bi-directional best hit with NCBI BLAST on KEGG genes and modules version 20190318, using all bacteria and archaea in the database. MAPLE output files for the module pathways, signatures, and complexes present in all six taxa were filtered for query coverage values of 100% and merged into a single file for each taxon. A map was created using RefSeq protein IDs and KEGG annotations for each taxon and KEGG annotations were obtained for each gene with significant multiplicity. These KEGG annotations were mapped to the sub-set of MAPLE modules that passed our filtering criteria, where we finally generate a MAPLE module-by-taxon presence absence matrix. PERMANOVA was performed using the same approach, but with a modified permutation matrix that only permuted treatment labels within a given taxon. Mapping the enriched genes to a higher level of biological organization required our null statistical model to be more complex than what was described in Eq.1. First, not all genes in all genomes are able to be annotated by KEGG. Second, KEGG annotated genes can be part of multiple MAPLE modules. That is, the relationship between the set of KEGG genes...
and MAPLE modules is not, strictly speaking, a mathematical function. However, because we are only interested in the degree of overlap in MAPLE modules across taxa within a given treatment we do not need to define the map between KEGG and MAPLE annotations. Rather, we can generate a null distribution for the size of intersecting sets of MAPLE modules by randomly sampling KEGG genes, mapping them to MAPLE pathways, and generating null MAPLE pathway-by-taxon matrices. Because each taxon has a distinct annotation, we set the number of significant genes that also had KEGG annotations as the sample size. Significance of a given intersection set size was established by comparing observed data to the null distributions, where for set intersections with multiple combinations we summed over the possible combinations for k intersections among n total taxa. Significance of a relationship between phylogenetic distance and the Jaccard index of MAPLE modules was assessed by comparing the slope obtained by ordinary least squares regression to a null distribution generated by calculating the OLS slope of the permuted variables for 10,000 iterations. Slope significance of distance-decay relationships was done as previously described (Nekola and White, 1999), where the absolute difference in slope values between a given pair of transfer regimes (β_{i,j}^{+/-}) was compared to a null distribution generated by randomly permuting treatment labels and calculating the difference in slopes.

**Data availability**

Raw sequence data is available on the NCBI Sequencing Read Archive under BioProject ID PRJNA639414. Reproducible code to perform the analyses in this study is available on GitHub under the repository: https://github.com/MURI2/Phylo_Evol_Timeseries. Processed data and annotations are available on Zenodo under the DOI: 10.5281/zenodo.4517573.

**Acknowledgements**

We thank J. Weissman and R. Wolff for providing valuable feedback on an earlier version of this manuscript. We thank M. Behringer, T. Doak, D. A. Drummond, P. Foster, M. Lynch, J. McKinlay, and members of the Lennon lab for their insights at various stages of the project. We thank H. Long for his help in setting up library construction and K. Miller for assisting with sample collection. This work was supported by US Army Research Office Grant W911NF-14-1-0411. Computing resources for simulations was supported by Lilly Endowment, Inc., through its support for the Indiana University Pervasive Technology Institute, the National Science Foundation under Grant No. CNS-0521433, and Shared University Research grants from IBM, Inc., to Indiana University.

**Author contributions**

W.R.S., E.P., M.H.G., and J.T.L. designed the project; W.R.S., E.P., and M.H.G. conducted the experiments and generated the data; W.R.S.
performed all statistical analyses; W.R.S., E.P., M.H.G., and J.T.L. wrote the manuscript.

Competing interests
The authors declare no competing interests.

References
Anderson, M. J., Walsh, D. C. I., Clarke, K. R., Gorley, R. N., and Guerra-Castro, E. 2017. Some solutions to the multivariate Behrens–Fisher problem for dissimilarity-based analyses. *Australian & New Zealand Journal of Statistics*, 59(1): 57–79.

Aouizerat, T., Gelman, D., Szitenberg, A., Gutman, I., Glazer, S., Reich, E., Schoemann, M., Kaplan, R., Saragovi, A., Hazan, R., and Klutstein, M. 2019. Eukaryotic adaptation to years-long starvation resembles that of bacteria. *iScience*, 19: 545–558.

Arai, W., Taniguchi, T., Goto, S., Moriya, Y., Uehara, H., Takemoto, K., Ogata, H., and Takami, H. 2018. MAPLE 2.3.0: an improved system for evaluating the functionomes of genomes and metagenomes. *Bioscience, Biotechnology, and Biochemistry*, 82(9): 1515–1517.

Avrani, S., Bolotin, E., Katz, S., and Hershberg, R. 2017. Rapid Genetic Adaptation during the First Four Months of Survival under Resource Exhaustion. *Molecular Biology and Evolution*, 34(7): 1758–1769.

Baak, M., Koopman, R., Snoek, H., and Klous, S. 2019. A new correlation coefficient between categorical, ordinal and interval variables with pearson characteristics. *Biology and Evolution*, 34(7): 1758–1769.

Bacun-Druzina, V., Cagalj, Z., and Gjuracic, K. 2007. The growth advantage in stationary-phase (GASP) phenomenon in mixed cultures of enterobacteria. *FEMS Microbiology Letters*, 325(2): 171–179. Publisher: Oxford Academic.

Bruno, J. C. and Freitag, N. E. 2011. Listeria monocytogenes adapts to long-term stationary phase survival without compromising bacterial virulence. *FEMS Microbiology Letters*, 325(2): 171–179. Publisher: Oxford Academic.

Cvijović, I., Good, B. H., and Desai, M. M. 2018. The effect of strong purifying selection on genetic diversity. *Genetics*, 209(4): 1235–1278. Publisher: Genetics Section: Investigations.

Doebeli, M., Ispolatov, Y., and Simon, B. 2017. Towards a mechanistic foundation of evolutionary theory. *eLife*, 6: e23804. Publisher: eLife Sciences Publications, Ltd.
Finkel, S. E. 2006. Long-term survival during stationary phase: evolution and the GASP phenotype. *Nature Reviews Microbiology*, 4(2): 113–120. Number: 2 Publisher: Nature Publishing Group.

Finkel, S. E. and Kolter, R. 1999. Evolution of microbial diversity during prolonged starvation. *Proceedings of the National Academy of Sciences of the United States of America*, 96(7): 4023–4027.

Gangloff, S. and Arcangioli, B. 2017. DNA repair and mutations during quiescence in yeast. *FEMS Yeast Research*, 17.

Gangloff, S., Achaz, G., Francesconi, S., Villain, A., Miled, S., Denis, C., and Arcangioli, B. 2017. Quiescence unveils a novel mutational force in fission yeast. *eLife*, 6: e27469. Publisher: eLife Sciences Publications, Ltd.

Garud, N. R., Good, B. H., Hallatschek, O., and Pollard, K. S. 2019. Evolutionary dynamics of bacteria in the gut microbiome within and across hosts. *Genome Biology and Evolution*, 12(12): 2292–2301.

Good, B. H., McDonald, M. J., Barrick, J. E., Lenski, R. E., and Desai, M. M. 2017. The dynamics of molecular evolution over 60,000 generations. *Nature*, 551(7678): 45–50.

Gross, J., Avrani, S., Katz, S., Hilau, S., and Hershberg, R. 2020. Culture volume influences the dynamics of adaptation under long-term stationary phase. *Genome Biology and Evolution*, 12(12): 2292–2301.

Hahn, M. W. 2018. *Molecular population genetics*. Oxford University Press; Sinauer Associates.

Helmus, R. A., Liermann, L. J., Brantley, S. L., and Tien, M. 2012. Growth advantage in stationary-phase (GASP) phenotype in long-term survival strains of *Geobacter sulfurreducens*. *FEMS Microbiology Ecology*, 79(1): 218–228. Publisher: Oxford Academic.

Kram, K. E., Henderson, A. L., and Finkel, S. E. 2020. Escherichia coli has a unique transcriptional program in long-term stationary phase allowing identification of genes important for survival. *mSystems*, 5(4). Publisher: American Society for Microbiology Journals Section: Research Article.

Kunkel, T. A. 2004. DNA replication fidelity. *The Journal of Biological Chemistry*, 279(17): 16895–16898.

Lacour, S. and Landini, P. 2004. σS-dependent gene expression at the onset of stationary phase in *Escherichia coli*: Function of σS-dependent genes and identification of their promoter sequences. *Journal of Bacteriology*, 186(21): 7186–7195. Publisher: American Society for Microbiology Journals Section: GEN REGULATION.

Lennon, J. T., Aanderud, Z. T., Lehmkul, B. K., and Schoolmaster, D. R. 2012. Mapping the niche space of
soil microorganisms using taxonomy and traits. *Ecology*, 93(8): 1867–1879.

Lever, M. A., Rogers, K. L., Lloyd, K. G., Overmann, J., Schink, B., Thauer, R. K., Hoehler, T. M., and Jørgensen, B. B. 2015. Life under extreme energy limitation: a synthesis of laboratory- and field-based investigations. *FEMS Microbiology Reviews*, 39(5): 688–728. Publisher: Oxford Academic.

Lynch, M. and Trickovic, B. 2020. A theoretical framework for evolutionary cell biology. *Journal of Molecular Biology*, 432(7): 1861–1879.

Maharjan, R. P. and Ferenci, T. 2017. A shifting mutational landscape in 6 nutritional states: Stress-induced mutagenesis as a series of distinct stress input–mutation output relationships. *PLOS Biology*, 15(6): e2001477. Publisher: Public Library of Science.

Maharjan, R. P. and Ferenci, T. 2018. The impact of growth rate and environmental factors on mutation rates and spectra in *Escherichia coli*. *Environmental Microbiology Reports*, 10(6): 626–633.

Maitra, A. and Dill, K. A. 2015. Bacterial growth laws reflect the evolutionary importance of energy efficiency. *Proceedings of the National Academy of Sciences*, 112(2): 406–411.

Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, 17(1): 10–12.

Martínez-García, E., Tormo, A., and Navarro-Llorén, J. M. 2003. Gasp phenotype: presence in enterobacteria and independence of sigma s in its acquisition. *FEMS Microbiology Letters*, 225(2): 201–206. Publisher: Oxford Academic.

Morita, R. Y. 1988. Bioavailability of energy and its relationship to growth and starvation survival in nature. *Canadian Journal of Microbiology*, 34(4): 436–441.

Morita, R. Y. 1993. Bioavailability of Energy and the Starvation State. In S. Kjelleberg, editor, *Starvation in Bacteria*, pages 1–23. Springer US, Boston, MA.

Nekola, J. C. and White, P. S. 1999. The distance decay of similarity in biogeography and ecology. *Journal of Biogeography*, 26(4): 867–878.

Orsi, W. D., Magraths, T., Vargas, S., Coskun, O. K., Vuillemin, A., Höhna, S., Worheide, G., D’Hondt, S., Shapiro, B. J., and Carini, P. 2021. Genome evolution in bacteria isolated from million-year-old subseafloor sediment. *bioRxiv*, page 2020.12.19.423498. Publisher: Cold Spring Harbor Laboratory Section: New Results.

Patefield, W. M. 1981. Algorithm as 159: An efficient method of generating random r × c tables with given row and column totals. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 30(1): 91–97. Publisher: [Wiley, Royal Statistical Society].

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. 2011. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12: 2825–2830.

Picelli, S., Björklund, A. K., Reinus, B., Sagasser, S., Winberg, G., and Sandberg, R. 2014. Ta5 transposase and tagmentation procedures for massively scaled...
sequencing projects. *Genome Research*, 24(12): 2033–2040.

Pruesse, E., Peplies, J., and Glöckner, F. O. 2012. SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics (Oxford, England)*, 28(14): 1823–1829.

Ratib, N. R., Seidl, F., Ehrenreich, I. M., and Finkel, S. E. 2021. Evolution in long-term stationary-phase batch culture: Emergence of divergent escherichia coli lineages over 1,200 days. *mBio*, 12(1). Publisher: American Society for Microbiology Section: Research Article.

Ryan, F. J. 1959. Bacterial mutation in a stationary phase and the question of cell turnover. *Journal of General Microbiology*, 21(3): 530–549. Publisher: American Society for Microbiology Journals Section: GENETICS AND MOLECULAR BIOLOGY.

Saumaa, S., Tover, A., Kasak, L., and Kivisaar, M. 2002. Different Spectra of Stationary-Phase Mutations in Early-Arising versus Late-Arising Mutants of Pseudomonas putida: Involvement of the DNA Repair Enzyme MutY and the Stationary-Phase Sigma Factor RpoS. *Journal of Bacteriology*, 184(24): 6957–6965. Publisher: American Society for Microbiology Journals Section: GENETICS AND MOLECULAR BIOLOGY.

Sewell, D., Laird, K., and Phillips, C. 2011. Growth advantage in stationary phase phenomenon in Gram-positive bacteria. *Journal of Hospital Infection*, 78(1): 73–75. Publisher: Elsevier.

Shoemaker, W. R. and Lennon, J. T. 2018. Evolution with a seed bank: The population genetic consequences of microbial dormancy. *Evolutionary Applications*, 11(1): 60–75.

Shoemaker, W. R. and Lennon, J. T. 2020. Quantifying parallel evolution. *bioRxiv*, page 2020.05.13.070953. Publisher: Cold Spring Harbor Laboratory Section: New Results.

Shoemaker, W. R., Jones, S. E., Muscarella, M. E., Behringer, M. G., Lehmkuhl, B. K., and Lennon, J. T. 2021. Microbial population dynamics and evolutionary outcomes under extreme energy-limitation. *bioRxiv*. Publisher: Cold Spring Harbor Laboratory _eprint: https://www.biorxiv.org/content/early/2021/01/26/2021.01.25.428163.full.pdf.

Stamatakis, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics (Oxford, England)*, 30(9): 1312–1313.

Zambrano, M. M. and Kolter, R. 1996. GASPing for Life in Stationary Phase. *Cell*, 86(2): 181–184.

Zambrano, M. M., Siegele, D. A., Almirón, M., Tormo, A., and Kolter, R. 1993. Microbial competition: Escherichia coli mutants that take over stationary phase cultures. *Science (New York, N.Y.)*, 259(5102): 1757–1760.

Zinser, E. R. and Kolter, R. 2004. Escherichia coli evolution during stationary phase. *Research in Microbiology*, 155(5): 328–336.