Experimental evolution of evolvability

Michael Barnett[1], Lena Zellar[1] and Paul B. Rainey[1,2]

[1]Department of Microbial Population Biology, Max Planck Institute for Evolutionary Biology, August-Thienemann Str. 2, 24306 Plön, Germany

[2]Laboratory of Biophysics and Evolution, CBI, ESPCI Paris, Université PSL, CNRS, 75005 Paris, France

Abstract

It has been proposed that capacity to generate variation can evolve by natural selection such that mutation becomes biased toward adaptive outcomes. However the idea that selection might drive the evolution of evolvability has remained largely theoretical. Here we detail the *de novo* evolution of evolvability in experimental populations of bacteria, making explicit the selective processes and molecular events. Key to experimental realisation was a lineage-level birth-death dynamic, where lineage reproductive success was dependent upon capacity to mutate between two target phenotypic states, each optima in a repeating cycle of environments. Fueled by variation in evolutionary potential associated with unique mutational paths, lineages capable of mediating rapid and reliable transitions between states through local mutational bias emerged. The mechanism is analogous to that underpinning highly mutable "contingency loci" in pathogenic bacteria. Drawing upon detailed knowledge of the evolutionary history of lineages, we identify key steps in construction of the locus, including changes to the genotype-phenotype map and input from mutations that elevated transcription and concomitantly, rates of localised frame-shifting. We also document ancillary advantages to rapid phenotype switching, with lineages possessing this ability being more likely to acquire secondary adaptive mutations. Our results provide a mechanistic account of a previously puzzling evolutionary phenomenon and clarify conditions for the adaptive evolution of evolvability.

Introduction

Generation of adaptive variation by mutation is critical for long-term evolutionary success. If such capacity is limited, then lineages, no matter how well-adapted to the present, will be susceptible to extinction on change of environment. Conversely, those lineages adept at generating adaptive variation stand to persist in the face of change with additional opportunity provided to expand and further diversify. It has been argued that selection, working on lineages that differ in capacity to generate variation, can promote the evolution of evolvability [13, 29, 60, 54, 18, 51]

Mechanisms that increase adaptive variation are well recognised. For example, sexual reproduction increases the rate at which beneficial mutations accumulate, thus enhancing the pace of adaptive evolution[39]. Lineages with elevated global mutation rates due to mutations disrupting DNA repair or replication machinery (global hyper-mutators) have been observed in experimental populations of microbes [53, 24, 1, 32, 57] and are also found in nature [44]. However, as most mutations are deleterious, global hyper-mutation comes at long-term cost. In clonal organisms, this can lead to gradual deterioration of the genome via Muller's ratchet [42]. A preferable strategy would be to increases mutation rate specifically at loci where variation is most likely to prove adaptive. Although this would appear to require foresight on the part of evolution, precisely such mechanisms of adaptive local hyper-mutation are found in the "contingency loci" of bacterial pathogens[40].

Contingency loci are characterised by the presence of mutation-prone sequences, such as short-sequence repeats, often located in promoter regions, or less commonly, in open reading frames, associated with genes encoding products that determine interactions with the external environment [40]. Being prone to replication slippage [33], such loci mutate at high frequency causing lineages of descendant cells to be phenotypically diverse. In the case of bacterial pathogens this diversity is thought to be instrumental in capacity of lineages to persist in the face of challenges presented by host immune responses [3].

The selective causes underpinning the evolution of contingency loci, and the evolution of evolvability more generally, are unclear and often controversial [54, 38, 5, 45, 46]. In part this stems from confusion over definition, but particularly from the fact that evolvability as adaption is difficult to explain without invoking selection at levels above that of individuals, which is typically considered weak [38, 63, 23]. However, mechanisms that improve capacity for evolution have been observed in simulations where evolving populations experience selection toward varying and recurring goals [10, 28]. Under such conditions, genotype-phenotype mappings evolve such that phenotypic states adaptive in the past become more readily generated by mutation in the future. Should future selective environments resemble past ones, derived types are biased toward generating adaptive change.

Here, recognising the importance of selection acting on lineages, we took advantage of an experimental bacterial system where repeated mutational transitions between two phenotypic states is readily achieved. We then subjected multiple lineages to a selective regime in which lineages better able to produce adaptive phenotypic variation replaced those less able. In one meta-population, following more than 40 sequential transitions between states, we observed the *de novo* evolution of local hypermutability, analogous to that known in pathogenic bacteria. Of crucial importance were a set of general potentiating mutations that bought selection to bear on a single regulatory gene (*plfu0185*), followed by a sequence of 24 successive mutations in *pflu0185*, which alternately activated and inactivated gene function. En route mutations in the promoter region caused transcription to increase, which correlated with an increase in the rate of frame-shifting, ultimately facilitating reliable phenotypic reversals via expansion and contraction of a specific heptanucleotide (HN) sequence. We demonstrate that the derived locus itself is evolutionarily responsive, with copies of the HNsequence increasing in a predictable manner in response to stringent selection. Finally, we report an unanticipated consequence of contingency behaviour, namely, that lineages endowed with capacity to mutate at high frequency are more likely to acquire secondary adaptive mutations that improve general ecological performance.

Main

The bacterium *Pseudomonas fluorescens* diversifies in spatially structured microcosms [48]. Two niche specialist mutants, the mat-forming cellulose-over-producing (CEL$^+$) and the mat-colonising non-cellulose producing (CEL$^-$) types, cycle as a consequence of negative frequency-dependent interactions [47]. Such interactions make possible an experiment in which competing lineages are rewarded for capacity to repeatedly transition between two phenotypic states by mutation. Lineages that fail to generate the target state are replaced by extant lineages (Figure 1).

Evolution of lineages

The selective regime shown in Figure 1 imposes constraints [6] that cause lineages to function as units of selection [35]. While the experiment began with a single genotype in each of the four meta-populations, mutation ensures that by the first phenotypic transition, lineages vary one to another. Failure of lineages to produce the next phenotype in the cyclical series results in extinction, with death providing opportunity for extant lineages to export their evolutionary success, including their evolutionary potential. Because each consecutive stage is founded by a single colony, offspring lineages resemble parental types. Selection thus works over two timescales: the doubling time of cells, and the doubling time of lineages. Over the long-term, successful lineages stand to be those composed of cells

that enhance lineage-level evolvability, with opportunity for cumulative refinement by selection.

Knowledge of the fate of lineages within meta-populations allows precise depiction of evolutionary dynamics (Figure 2). Evident are instances in three of the four meta-populations (A-C) in which descendants of single lineages went to fixation, forming unbroken series of successful transitions. Meta-population D suffered a mass extinction at transition 48. If lineages had not been treated as units of selection, that is, if in the face of extinction, reproduction of lineages had been disallowed, then all meta-populations would all have faced early extinction (Supplementary Figure 1). That selection acts on both viability and fecundity is crucial, in absence of the latter, there is minimal opportunity for exploration of genetic (and thus phenotypic) space, or for refinement of traits.

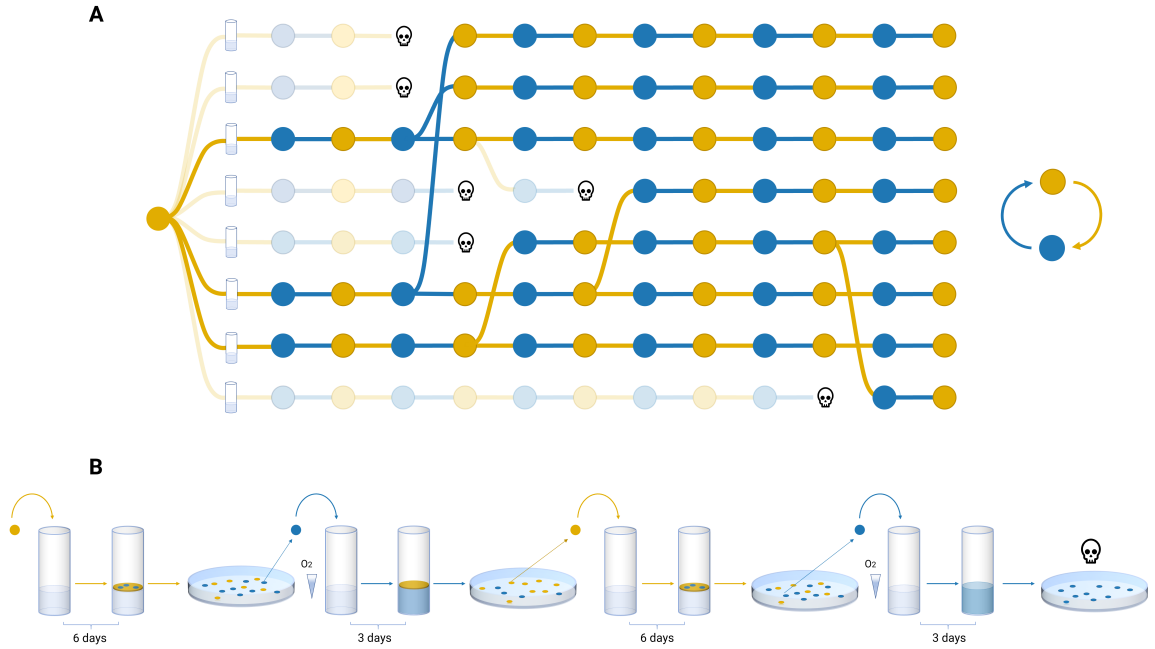

Figure 1: **Lineage selection**. A. Graphical representation of the experimental regime depicting eight competing lineages founded from a single common genotype. Lineages are required to transition repeatedly between CEL$^+$ (orange) and CEL$^-$ (blue) states by mutation. Lineages that fail to achieve the target phenotype are marked extinct, with extinction providing opportunity for extant lineages to export their evolutionary success. B. Schematic of protocol that begins by founding a microcosm with a single CEL$^+$ genotype. After six days in static culture, descendant cells are diluted and plated on agar with colony phenotype used to detect CEL$^-$ mutants. A single CEL$^-$ mutant is then used to found growth in a fresh microcosm, with a sample again examined by agar plate culture for the presence of CEL$^+$ types. A single CEL$^+$ type then founds the next bout of selection and the cycle repeats. The illustrated lineage goes extinct on the final transfer because it failed to generate the next (CEL$^+$) target phenotype. Key to the selection regime are frequency dependent interactions between CEL$^+$ and CEL$^-$: CEL$^+$ forms a mat at the surface and provides conditions that favour CEL$^-$ mutants; CEL$^-$ types deplete oxygen thus establishing conditions that favour the evolution of surface-colonising CEL$^+$ types. .

Whole genome sequence of derived lineages from meta-populations A-C were obtained at the end point of each series, but also periodically en route. In meta-populations A and C, lineages carrying loss-of-function (LoF) mutations in genes involved in mutational repair fixed (*mutS* and *uvrD*, respectively), with the point of origin marked in Figure 2. In contrast, no mutations affecting genes involved in DNA replication or repair were detected in the derived lineages of meta-population B, prompting continuation of the experiment.

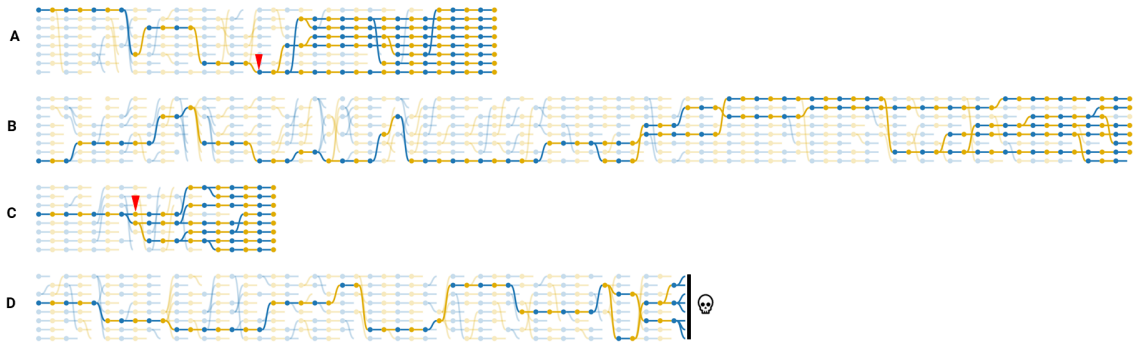

Figure 2: **Evolutionary dynamics of meta-populations**. Four CEL$^+$ genotypes were used to found each of four meta-populations (A-D) each composed of eight lineages. Orange nodes indicate the CEL$^+$ target phenotype and blue CEL$^-$. Lineage death occurred through failure to generate the target phenotype and provided extant lineages opportunity for birth. Solid lines depict the genealogy of final surviving lineages, with extinctions indicated by broken faded lines. Red arrows in A and C indicate the point of origin of mutator genotypes founding lineages that rapidly fixed (see text). .

Meta-population B ultimately cycled through 80 transitions (720 days within microcosms) with the experiment being terminated on evidence of gradual fixation of a single lineage. En route this meta-population experienced 120 lineage deaths and subsequent births. Fluctuation assays (based on resistance to rifampicin) of representative genotypes from derived lineages showed that global rates of mutation among lineages from meta-populations A and C were elevated 300-fold and 350-fold, respectively, however the mutation rate among derived lineages of meta-populations B and D did not differ to ancestral *P. fluorescens* SBW25 (Supplementary Figure 2).

Local hyper-mutability underpins lineage adaptation

To understand the mutational history and genetic bases of both successful and unsuccessful lineages whole-genome re-sequencing was used to identity and order of mutations underpinning transitions between CEL$^+$ and CEL$^-$ states. This revealed 548 mutational events that targeted 77 genes, 29 of which were hit multiple times. Despite the availability of diverse mutational paths a single regulatory gene, *pflu0185*, became the main mutational target, eventually accounting for 59% (306 of 520) of all phenotypic transitions (Figure 3). Moreover, the majority of these (186 of 300) were caused by

duplication and loss of a single heptanucleotide (HN)-sequence 'GGTGCCC' located near the 5' end of the gene (nucleotides 19-25). Mapping the occurrence of phenotypic transitions effected by changes in copy number of the HN-sequence on dynamics of lineages within meta population B (Figure 3C) showed that transitions involving gain and loss of the HN-sequence were associated with a decrease in the number of extinctions from 26% per transition to 10% (Figure 3A).

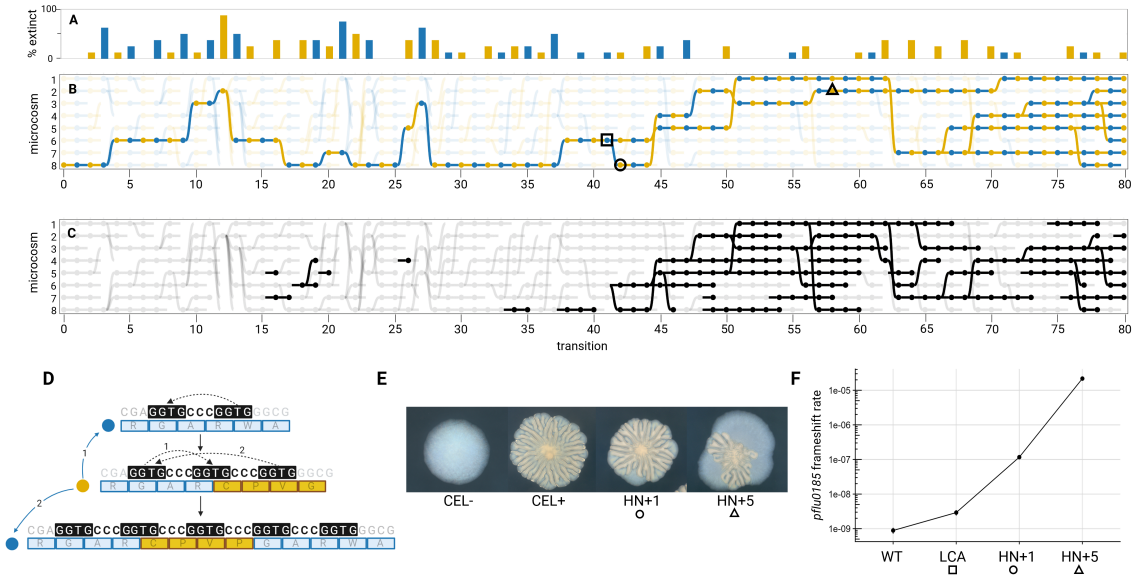

Figure 3: **Local hyper-mutability accounts for adaptive advantage of derived lineages in meta-population B**. A. Extinction events at CEL$^+$ (orange) and CEL$^-$ (blue) transitions. B. Lineage dynamics with focal points indicated by symbols around nodes (see text). X-axis labels are transitions (T) and Y-axes denotes microcosm (M). The node identified by the square box is thus T41 M6. C. Phenotypic transitions effected by changes in the number of copies of a HN-sequence in *pflu0185* . D. Proposed mechanism of slipped-strand mispairing. Top shows the DNA sequence of *pflu0185* surrounding and including the focal HN-sequence (GGTGCCC) and relevant downstream nucleotides (GGTG). The arrow indicates mispairing between GGTG repeats which in the forward direction leads to an additional copy of the HN-sequence (middle). Two copies (which have destroyed the reading frame) can return to a single copy by slippage in the reverse direction. Alternatively, mispairing in the two-copy state (middle), can, via reverse slippage, result in three copies of the HN-sequence (bottom), which also restores the correct reading frame. E. Image of CEL$^+$ and CEL$^-$ colonies on agar plates with derived lineages having either a single copy of the HN-sequence, or five copies – the latter showing evidence of sectoring indicative of an elevated mutation rate. F. Relationship between genotype and rate of frame-shift mutations in *pflu0185* determined by fluctuation assay. Symbols denote genotypes from (B). Error bars (largely contained within the symbols) are 95% CI.

Pflu0185 is a regulatory protein composed of two enzymatic domains: a diguanylate cyclase (DGC) — responsible for the synthesis of cyclic-di-GMP (c-di-GMP) — and a phosphodiesterase (PDE) — responsible for breakdown of c-di-GMP [49]. In *P. aeruginosa* functional analysis shows the protein to have PDE function, but no DGC activity [8], which is consistent with the genetic analysis reported here. The cellulose-synthase-encoding *wss* operon is allosterically activated by c-di-GMP, such that when levels are elevated the CEL$^+$ state is achieved, but when reduced, cells are CEL$^-$ [55]. With just

a single copy of the HN-sequence, PDE function is intact, c-di-GMP levels are low and cells achieve the CEL$^-$ state. On addition of an extra copy of the HN-sequence the reading frame is destroyed, PDE function is lost, c-di-GMP levels increase, and the CEL$^+$ state is realised.

Mechanistically, rapid changes in the number of copies of short sequence repeats, can be attributed to slipped-strand mispairing - a genomic instability that increases the likelihood of deleting or duplicating repeated units [34]. Factors initiating the first duplication are unclear, however, evidence presented below indicates contributions from mutations that enhance *pflu0185* transcription. But immediately relevant is the fact that the duplicated GGTGCCC sequence is flanked, downstream, by a further copy of GGTG and thus the focal sequence is GGTGcccGGTG (Figure 3D). Mis-pairing between the two copies of GGTG stand to catalyse a duplication event that is readily reversible by forward slippage. However, with two copies of the HN-sequence (three copies of GGCT), reverse slippage results in three additional copies of the HN-sequence and restores the reading frame. This is consistent with results of sequence analysis, which showed transitions to CEL$^+$ were underpinned by gain of one extra copy (but not two), whereas transitions in the reverse direction, while typically caused by loss of the additional copy, were sometimes corrected by addition of three copies.

Genotypes primed to transition between CEL$^+$ and CEL$^-$ via changes in the number of copies of the HN-sequence pass this enhanced evolutionary potential to offspring. Once in the CEL$^+$ state, with two copies of the HN-sequence, restoration of the correct reading frame was readily achieved. Indeed, 96% of lineages inheriting the duplication reactivated Pflu0185 in the next transition through correction of the reading-frame, and were ∼eight times more likely to survive the next transition compared to inheritance of any other regulatory mutation (95% CI [2.5, 26], P=0.0002, Fisher's exact test, two-tailed). Of lineages that achieved the two-copy (non-functional) state, 86 of 89 instances involved restoration of the *pflu0185* reading frame by contraction. In three instances the reading frame was restored by expansion of the tract by three additional copies. In these lineages, the HN-sequence was more likely to undergo further expansion at the next transition (5:3 contraction:expansion, P=0.0049, Fisher's exact test, two-tailed). As the number of repeats increased, the rate at which transitions occurred visibly increased (Figure 3E).

By T41 M6 (black square in Figure 3B) duplication of the HN-sequence became increasingly probable even in the face of numerous additional paths for generating LoF mutations — both within *pflu0185* and elsewhere in the genome. Indeed, alternative LoF mutations in *pflu0185* were occasionally observed late in evolution and, where irreparable, led to lineage extinction (Supplementary Figure S3).

The relationship between number of tandem repeats and rate at which phenotypic variants are generated [14] predicts that as copy number increases, so too will local mutation rate. While global measures of mutation rate reported above showed no increase, by translationally fusing *pflu0185* to an out-of-frame kanamycin resistance (*kanR*) gene, the rate of frame-shifting, specifically within *pflu0185*, could be assessed by fluctuation assay (Figure 3F). Surprisingly, a statistically significant three-fold increase in the rate of frame-shifting was evident in the genotype that marked the coalescent point for the lineage that went to fixation (T41 M6, black square in Figure 3B), despite the fact that this genotype had just a single copy of the HN-sequence. With two copies ((T42 M8), black circle in Figure 3A) the rate of frame-shifting increased a further 150-fold and with five copies (T58 M2, black triangle in Figure 3B) (the maximum observed in the selection experiment) the rate of frame-shifting was more than four orders of magnitude greater than the ancestral genotype (Figure 3F).

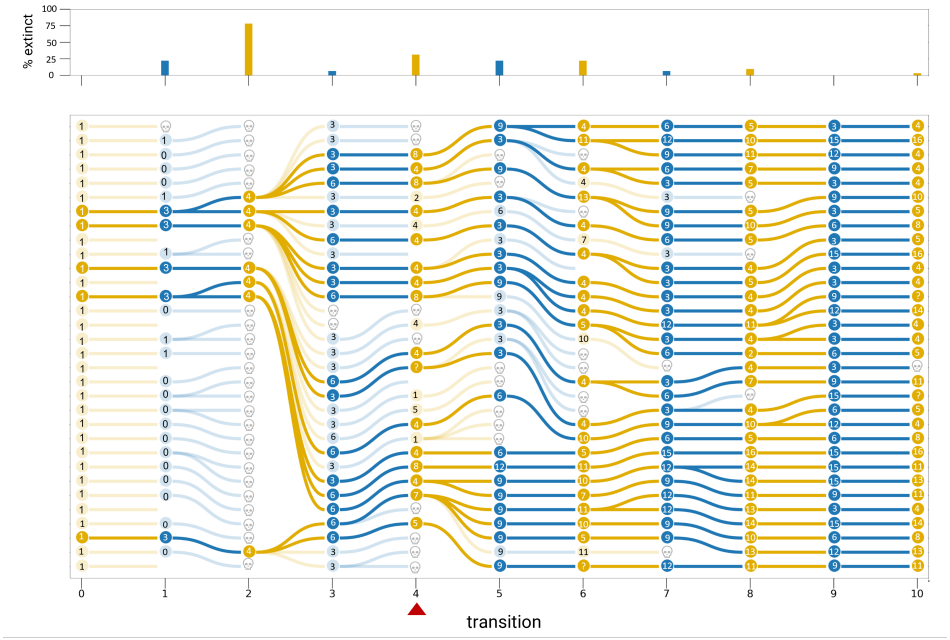

Figure 4: **Stringent selection drives expansion in hnrepeat copy number**. A total of 32 lineages were founded by genotype T42 M8 (black circle in Figure 3B) containing a single additional copy of the HN-sequence. Lineages were subject to the selection regime as shown in Figure 1, but population size was reduced 30-fold and lineages were required to achieve the target phenotype within 24 hours or face extinction. HN-repeat copy number is indicated within nodes. Colour code for phenotypic states are previously described. The red triangle indicates a point after which the stringency of selection was further increased by requiring successful lineages to have at least five colonies of the target phenotype present. The upper panel shows extinctions at each transition.

If the derived form of *pflu0185* has evolved to facility further evolution, then the locus itself should be evolvable. Accordingly, we took a single CEL$^+$ genotype carrying the initial HN-duplication (+1 GGTGCCC unit) (T42 M8 black circle in Figure 3B) and used this genotype to found 32 lineages that competed under a restrictive regime where population size was reduced by a factor of thirty and

persistence required re-establishment of the target state every 24 h (Figure 4).

The selective response, lineage-level dynamics, and changes in numbers of repeats over ten transitions are revealing. Most lineages managed — within 24 h — to achieve the CEL$^-$ state required for the first transition, but the majority did so by losing one copy of the HN-sequence. Without exception, those lineages that switched to CEL$^-$ by loss, were unable to transition to the CEL$^+$ state during the next 24 h period and thus went extinct. In sharp contrast, those lineages that achieved the CEL$^-$ state (at transition 1) by adding three additional copies, became progenitors of lineages that persisted throughout the course of the experiment. At the end point, lineages harboured on average 8 additional copies of the HN-sequence, with some acquiring as many as 16. A clear association is evident between the number of copies of the HN-sequence and likelihood of survival (Figure 4).

Navigating multiple paths and avoiding genetic dead-ends

Having observed the *de novo* evolution of localised adaptive hyper-mutability, we next sought understanding of mechanistic causes. Of interest were the genetic changes potentiating the first transient duplication events (T 16) (Figure 3C), but of particular focus were events between transitions 14 and 42 (Figure 3C) that marked a significant improvement in the likelihood that successive transitions were effected by changes in the number of copies of the HN-sequence.

Genome sequence data showed that the majority (89%) of phenotypic transitions were associated with single mutations, allowing genetic causes to be precisely identified. The switch between CEL$^-$ and CEL$^+$ states is asymmetrical with details of the genetic architecture of components affecting cellulose production being described in Figure 5A. Figures 5B-D provide a visual representation of principle changes for all lineages from meta-population B at three different transition stages. The full data set is available via an interactive website (micropop.evolbio.mpg.de/data/evolve) and Supplementary Video 1 allows sequential accumulation of mutations and genetic targets (up to T41) for all lineages of meta-population B to be view transition-by-transition. Figure 5b displays mutational history at transition T6 and shows early on, mutations in a range of genes, half of which are DGCs (green squares), and one of which, the DGC-encoding *awsR*, was targeted by mutation multiple times effecting repeated transitions between CEL$^-$ and CEL$^+$ (yellow self-loops circling *awsR*).

By transition T13 (Figure 5C), an increasing number of transitions were triggered by further compensatory mutations in *awsR*, however, utility of *awsR* ceased upon occurrence of a 54 base pair deletion at T9 M6, which fixed at transition 12, following an almost fatal extinction event. Additionally, the diversity of mutational routes further expanded with many, including mutations in the cellulose synthe-

sis operon, *wss*, proving dead-ends. Mutations that generate CEL$^-$ through destruction of the cellulose biosynthetic machinery accounted for 16% of all transitions, with these lineages being ∼10 times more likely to go extinct at the next transition compared to any other target (P<0.0001, Fisher's exact test, two-tailed). This stems from the fact that cellulose synthesis is essential for the CEL$^+$ phenotype, thus LoF mutations must be repaired at the next transition in order to avoid extinction.

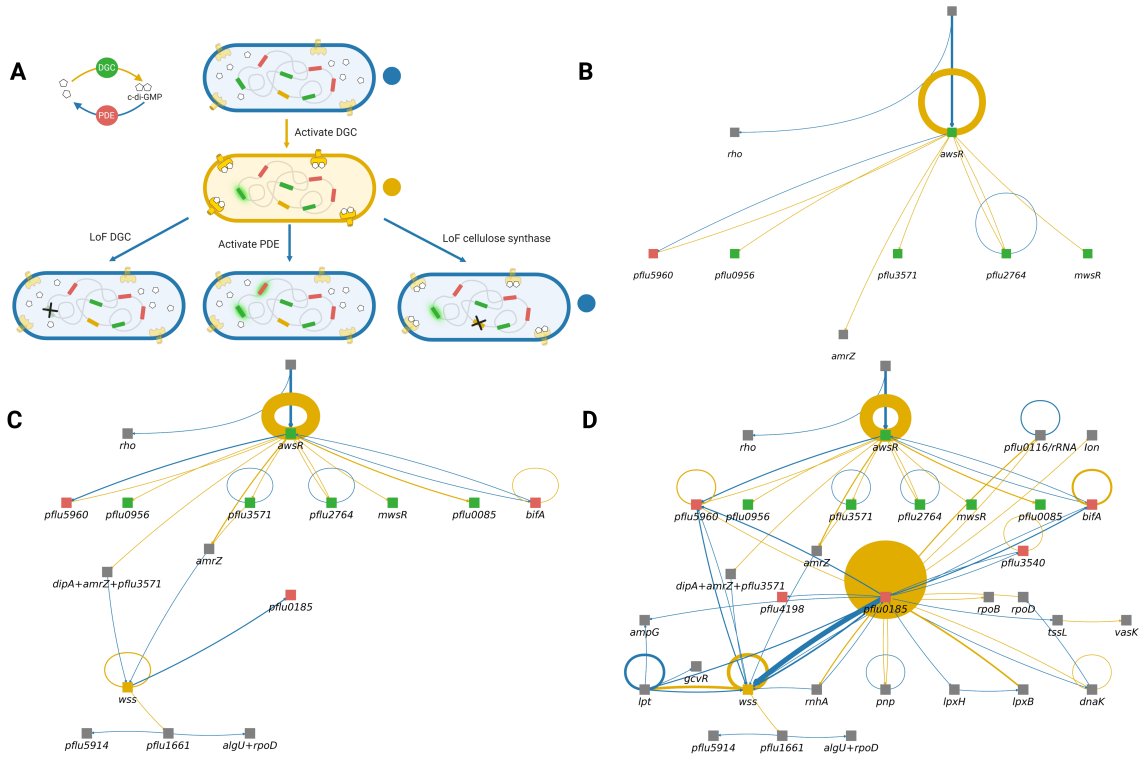

Figure 5: **Diversity of mutational paths generating variation in evolutionary potential**. A. Genetic architecture underpinning mutational transitions between CEL$^-$ and CEL$^+$ states. C-di-GMP is the critical signalling molecule and allosteric activator of cellulose synthesis. The molecule is synthesised from two molecules of guanosine-5'-triphosphate (GTP) (diamond symbols) by diguanylate cyclases (DGCs) and degraded by phosphodiesterases (PDEs). Red boxes in cells denote genes encoding PDEs, green boxes depict DGCs and the yellow box is the structural determinant of cellulose encoded by the 10 gene *wss* operon. The CEL$^+$ state is realised by mutations that elevate c-di-GMP levels. Once in the CEL$^+$ state, there exist three primary routes to CEL$^-$: mutational inactivation of DGCs, mutational activation of PDEs, or LoF mutations in cellulose synthesis. The SBW25 genome encodes 41 genes with roles in synthesis and degradation of c-di-GMP, with 394 of the 549 mutations identified within (or adjacent to) genes involved in modulation of c-di-GMP levels. B-D Snap shots of mutational trajectories at transition 6 (B), 13 (C) and 41 (D) for all lineages from meta-population B starting from the ancestral state (the CEL$^+$ genotoype used to found all eight lineages). The full sequence is available as a video (Supplementary Video 1). Lines indicate successive sequences of genetic targets (boxes). Line thickness indicates the frequency with which the mutational route was used. Blue lines indicate transitions from CEL$^+$ to CEL$^-$ and yellow lines from CEL$^-$ to CEL$^+$ as in Figure 5A. The yellow box depicts the cellulose synthesis encoding machinery. Self-loops, for example around *awsR* in C, and around *pflu0185* in D, indicate repeated transitions via successive compensatory mutations at these loci.

Genes involved in regulation of c-di-GMP provided multiple targets for successive transitions, with

LoF mutations often repaired by substitutions at residues elsewhere in the same regulatory gene (or through mutations in the 5' untranslated region that likely increased gene expression). In contrast, not a single instance of intragenic compensation was observed for *wss*, where repair came about almost exclusively through change to affected residues (either direct reversion or alternative substitutions) (Supplementary Table 1). This indicates a much tighter structure-function relationship for the *wss* locus and contributed to this target being a frequent end-point.

Also notable at transition T13 (Figure 5C) are mutations that target PDEs (red squares), with an unusual sequence of mutations bringing the PDE activity of *plu0185* into play for the first time. At T10 M3, in the lineage that went to fixation, three atypical mutations co-occurred. These targeted genes encoding a PDE (*dipA*), a DGC (*pflu3571*) and the global transcriptional regulator *amrZ*. The latter is conserved across *Pseudomonas* where it controls transcription of multiple DGCs [43]. We hypothesized this unusual triple-mutant path had caused the source of cyclic-di-GMP to become distributed across multiple DGC-encoding loci, such that LoF mutations at any single DGC, were no longer able to reduce c-di-GMP levels sufficient to generate the CEL$^-$ state. In support of this hypothesis, transposon mutagenesis was employed in an attempt to identify genes capable of suppressing the phenotype of CEL$^+$ at transition 10 (M3), but no DGC-encoding loci were identified. Assuming a Poisson distribution of transposon insertions across the SBW25 genome (6,700 genes), the probability of a gene remaining undetected given our screen of 54,000 mutants is $\sim$0.0003.

From transition 14 onward, successful lineages no longer generated CEL$^-$ via changes in DGCs, but did so instead by mutations that affected the activity of PDEs (Supplementary Figure 4). The sudden switch from DGCs as mutational targets to PDEs was surprising, as numerous other intact DGC-encoding genes capable of generating CEL$^+$ through single mutations are known [37]. The shift from DGCs as targets, to PDEs, marked a significant alteration in the range of mutational targets, which was instrumental in turning the focus of selection to PDEs and increasingly to *plfu0185*.

Figure 5D depicts mutational history of all lineages up to transition 41, the moment that transitions between CEL$^+$ and CEL$^-$ began to be caused by frequent changes in copy number of the HNrepeat (Figure 3C). Evident are a diversity of one-off mutational routes, the majority of which proved evolutionary dead-ends. Striking, however, is the fact that *pflu0185* became the target of multiple mutations that facilitated repeated switching between CEL$^+$ and CEL$^-$ states (thick self-loop circling *pflu0185*).

Concomitant increases in transcription and mutation rate at *pflu0185* potentiated repeat-associated hyper-mutation

Figure 6A shows a series of consecutive mutations in *pflu0185* that occurred in the focal lineage between T13 M6 and T41 M6 (Figure 3B). Each of these mutations was causally responsible for the switch between CEL$^+$ and CEL$^-$ via effects on the PDE activity of *pflu0185*. At the point of coalescence (T41, M6) *pflu0185* carried 24 mutations (including three direct reversals), either within, or upstream of the gene. Among this convoluted history of mutations, no single event stood out in terms of relevance to hyper-mutability caused by changes in copy number of the HNsequence. However, given a three-fold increase in mutation rate at T41 M6 (compared to ancestral SBW25) (Figure 3F), combined with the presence of multiple mutations in the upstream region, attention turned to the effect of selected mutations on the rate of frame-shifting within *pflu0185*. This was achieved by fluctuation assay using the translational *pflu0185-kanR* gene fusion.

On introduction of the construct into genotypes from different points along the mutational series, but with particular focus on those mutations occurring within the upstream region, checks were firstly made to determine whether the fusion allowed expression of kanamycin resistance to a level sufficient to enable conduct of fluctuation assays. The results, shown in Figure 6B, revealed a series of surprises. While the MIC of ancestral SBW25 to kanamycin was 8 µg/ml, the first mutation in the promoter region (-53 A→G), which generated the CEL$^+$ phenotype by reducing PDE expression, led to a four-fold reduction in MIC (a level insufficient to perform a fluctuation assay). The next mutation (-197 C→T), also in the promoter region, caused CEL$^+$ to switch to CEL$^-$ by restoring PDE activity via changes to expression that returned the MIC (and thus transcription) to that of the ancestral type. The following transition, from CEL$^-$ to CEL$^+$, occurred via a mutation in the PDE domain (V579M), which resulted in reduction in activity sufficient to allow c-di-GMP levels to rise to a level that triggered the CEL$^+$ state. The next transition (to CEL$^-$) occurred, once again, via an upstream mutation (-167 C→T), which caused a further ten-fold increase in MIC (and thus transcription) sufficient to compensate for the previous reduction in PDE activity. Following the -167 C→T mutation were a series of changes in the open reading frame of *pflu0185* , punctuated by four further promoter mutations, with the last of these achieving a 30-fold increase in MIC over the ancestral level.

We next asked whether sense could be made of the series of upstream mutations. The changes, documented in Figure 6C, and placed in context of surrounding nucleotide sequence, deliver a promoter sequence typical of those found in *Pseudomonas* [16, 58]. Such mutational changes are consistent with the measured elevation in transcription being sufficient to deliver increased resistance to kanamycin.

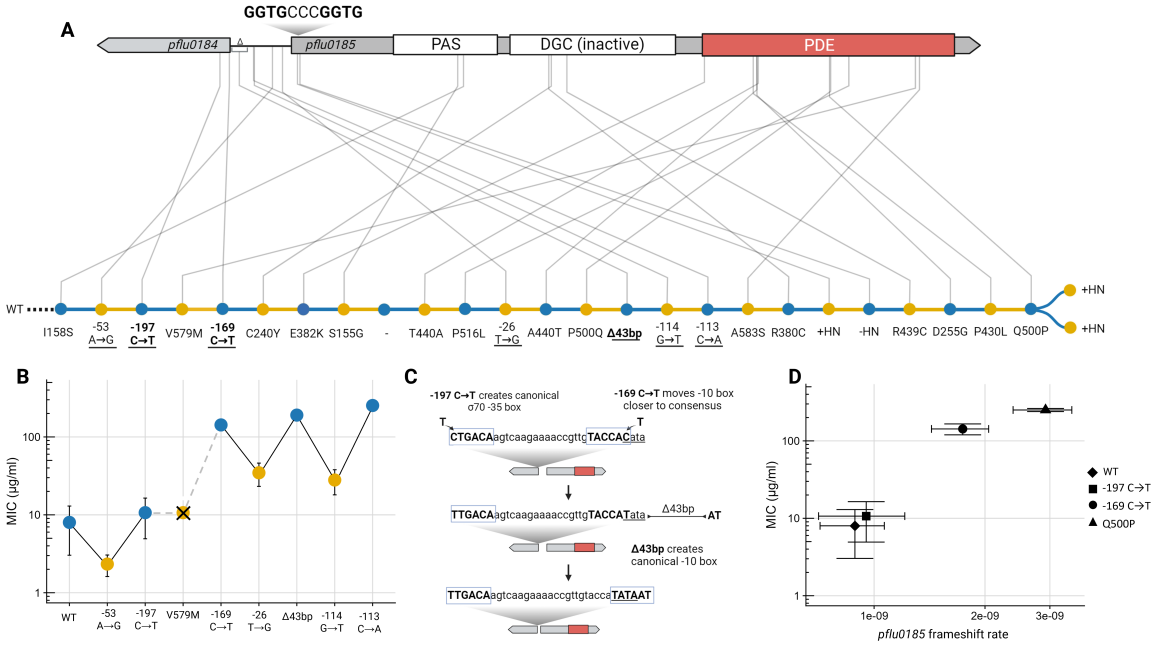

Figure 6: **Mutational history of potentiating mutations in *pflu0185* .**A. Genetic structure of *pflu0185* and sequence of successive mutations modulating changes in phenotypic state. The top portion of A shows the *pflu0185* open reading frame with location of the HN-sequence, PAS signalling domain, inactive DGC domain and, in red, the PDE domain. Also shown is the upstream gene and intergenic space where multiple mutations arose that affected transcription. B. Effects of mutations on transcription of *pflu0185* and CEL$^+$ / CEL$^-$ status determined through assay of MIC in the presence of kanamycin (putative promoter mutations are underlined in A). Mutation V579M is included to demonstrate its role in reducing catalytic activity of the PDE, necessitating further increases in *pflu0185* transcription (see text). C. Sequence changes in upstream regions (in bold type) placed in context of surrounding sequence that contribute to formation of a typical *Pseudomonas* promoter. D. Relationship between transcription (as determined by MIC assay) and rate of frame-shift mutations in *pflu0185* . Note that the MIC of Q500P is equivalent to the last upstream mutation (-133 C→A). Error bars are 95% CIs

.

In terms of the possibility that these changes might have potentiated proclivity of the HN-sequence to expand in copy number, the relationship between MIC and rate of frameshift mutations in *pflu0185* was determined by fluctuation assay. The results show that the -169 C→T mutation produced a highly significant two-fold increase in the rate of frameshfting compared to ancestral (wildtype) *P. fluorescens* SBW25. This was further increased by mutation Q500P (T41 M6) that occurred just just prior to the point where expansion and contraction of the HN-sequence became increasingly likely.

Local hyper-mutation of *pflu0185* -facilitated additional adaptive evolution

A final observation concerns transitions between CEL$^+$ and CEL$^-$ that were associated with more than one mutation. Those transitions effected by changes in HN-repeat number were more than three times as likely to be associated with additional (secondary) mutations. Of 91 such transitions, 19 harboured

secondary mutations (21%). In contrast, of 167 transitions effected by mutations not involving the HN-sequence, only 10 carried secondary mutations (6%) (P=0.0007, Fisher's exact test, two-tailed).

While secondary mutations that hitchhike with transition-causing mutations are to be expected (Supplementary Table 2), mutations in loci encoding genes connected to chemotaxis and motility (Supplementary Table 3), have conceivable ecological relevance. Moreover, two loci, the aerotaxis receptor *aer* (*pflu4551)* and a methyl-accepting chemotaxis protein (MCP), were observed to occur repeatedly (six and four times, respectively) across multiple lineages during transitions mediated by the HN-sequence (Figure 7A). Such striking parallelism suggests that these mutations were themselves adaptive. Figure 7B elaborates the benefit presumed to accrue to lineages derived from genotypes that switch phenotype early, and reliably.

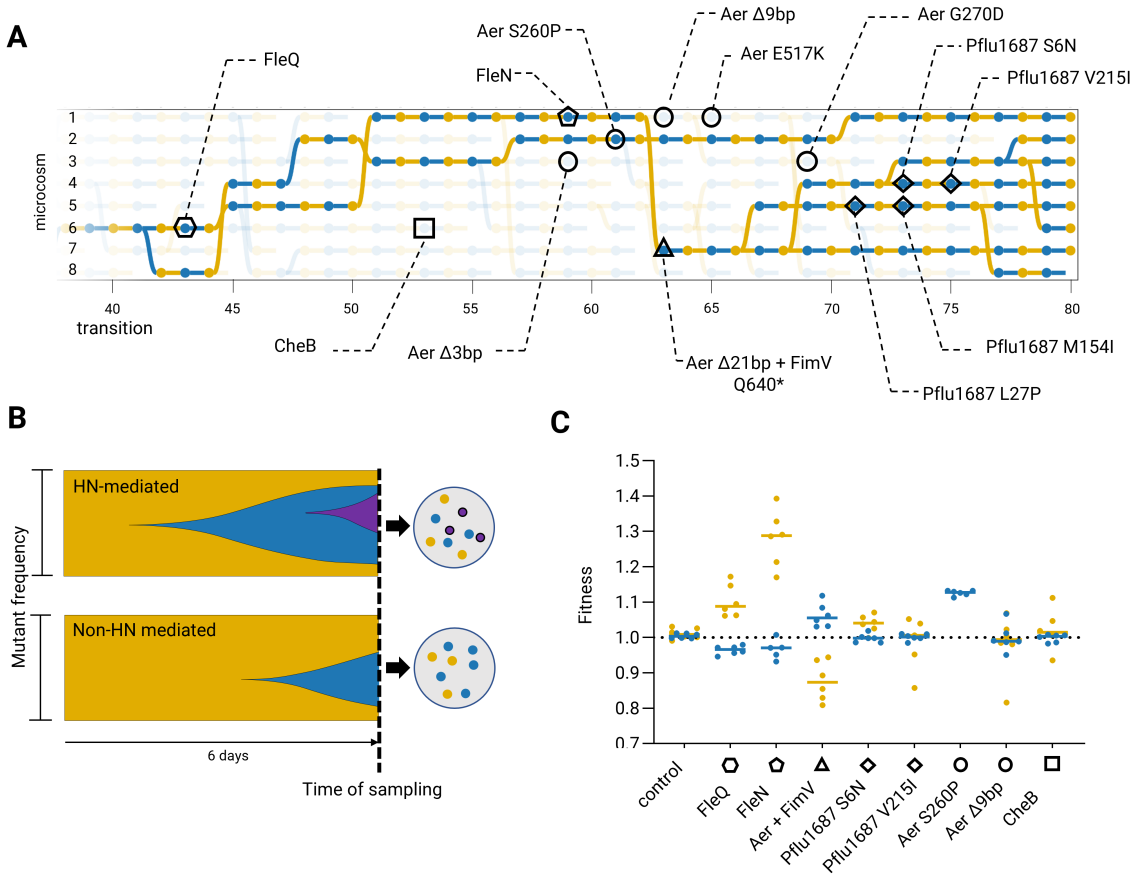

Figure 7: **Secondary mutations and fitness effects**. A. Genealogy of successful lineages from meta-population B between transition 40 and 80 on which secondary mutations of predicted ecological significance are indicated. B. Muller plots showing predicted consequences of presence (top) and absence (bottom) of local hyper-mutation: early and reliable switching provides enhanced opportunity for secondary mutations (purple) to reach high frequency (top). C. Fitness consequences of mutations determined by competitive assay against immediate ancestral types in CEL$^+$ (orange) and CEL$^-$ (blue) backgrounds in static (CEL$^+$ assays) and shaken (CEL$^-$ assays) culture.

To test the adaptive hypothesis, the fitness effects of a sample of mutations were determined relative to immediate ancestral types, both in CEL$^+$ and CEL$^-$ backgrounds. With emphasis on the lineage that descends from T42 M6, the first mutation examined (in T43 M6) was in FleQ (R371C), a master regulator of motility and polysaccharide production (including cellulose) [26, 21]. Relative to the immediate ancestral types the *fleQ* mutation increased fitness ∼10% in CEL$^+$, but marginally decreased fitness in CEL$^-$. The next mutation, in FleN (L42Q) (T58 M1), a gene that also regulates flagella through interactions with FleQ [11], increased fitness of CEL$^+$ by a further ∼30%, with with minimal effects in CEL$^-$. Relative to the fitness of T42, M6, the two mutations combined contributed a ∼40% fitness gain in CEL$^+$.

The genotype at T62 M7 carries two additional mutations (the second only instance in the experiment where three mutations occurred at the same transition), one a 21 base pair deletion in Aer and the other a stop codon within FimV (Q640*). These two mutations reduced fitness of CEL$^+$ by ∼10%, while increasing fitness of CEL$^-$ by ∼

Fitness effects of secondary mutations in other lineages were also examined, including changes in the chemotaxis protein CheB (T92P at T53 M6) and further mutations in Aer (S260P at T60 M2 and Δ9bp at T62 M1). Fitness effects due to mutations in CheB (T92P) and Aer (Δ9bp) were undetected in both CEL$^+$ and CEL$^-$, but the Aer (S260P) mutation increased fitness in CEL$^-$ by ∼10%. Together the range of effects shown in Figure 7C are consistent with the hypothesis that at least some secondary mutations co-occurring with changes in HN-repeat number are adaptive.

Discussion

Evolution by natural selection is a blind process, but living systems often appear endowed with evolutionary foresight [13]. Mechanistically, this is conceivable. Certain configurations of gene regulatory networks[61, 10, 59], developmental systems [29, 17] and mutational processes [40, 25, 22, 7] have apparent adaptive utility in future environments. Ability to realise such advantages requires not only memory of past history, but also ability to regenerate previously achieved phenotypic states. While regulation of gene expression allows environment-dependent modulation of physiology, critical to success in diverse and unpredictable future environments, is generation of phenotypic states that exceed the breadth of those achievable by stereotypical modes of gene regulation. Here we have detailed how prior evolutionary history can be incorporated into the genetic architecture of a single cell such that successful past phenotypic states – achievable solely by mutation – can be repeatedly realised in future environments.

Central to our claims is an experiment in which lineages evolved as units of selection, where those lineages better able to generate, by mutation, adaptive phenotypic variation, replaced those that were less proficient (Figure 1). Of the four original meta-populations (Figure 2), one went extinct, two evolved global hyper-mutators, but in the fourth, a single lineage emerged with capacity to switch rapidly between phenotypic states via expansion and contraction of a short nucleotide repeat in a manner precisely analogous to that of contingency loci in pathogenic bacteria [40]. Moreover, when multiple lineages of a single genotype carrying the mutational history of 42 previous transitions, but carrying just a single additional copy of the HN-repeat (T42 M8, Figure 3C), were propagated under a regime that required phenotypic switching on a daily basis, lineages with up to 16 copies of the repeat rapidly emerged (Figure 4). This demonstrates not only capacity of the HN-repeat within *pflu0185* to generate adaptive phenotypic variation, but shows that the locus, in its derived state, is itself, evolvable [46].

Selection works most forcefully on Darwinian individuals and is typically weak at higher levels of organisation [38, 63, 54]. Here the death and birth of lineages established a feedback between lineage-level properties and the evolutionary dynamics of cells. While selection within microcosms favours fast-growing cells, successful lineages are those with the capacity to generate phenotypic variants suited to the future environment. Moreover, selection over the doubling time of lineages serves to align cell and lineage fitness, precisely as occurs during, for example, the evolutionary transition from cells to multicellular life [6].

Data presented in Figure 4 (and available online) provides an unprecedented view of evolutionary change. In standard selection experiments details of mutational history are often missing, for example through unobserved extinction events, or difficult to unravel because of absence of connection between mutation and phenotypic effects. Here, the entire mutational history of a single meta-population has been laid bare. Moreover, given knowledge of the genetic architecture underpinning CEL$^+$ and CEL$^-$ states, it has been possible to causally connect hundreds of mutations to their phenotypic consequences.

Examination of the patterns of mutational history shows how the effects of selection accumulate: new progress is constructed on top of old. But also evident are the roles of chance (and serendipity), particularly with regard to refinement of mutational targets. Notable, for example, was the move from DGCs to PDEs, caused by mutations that bought about distributed production of c-di-GMP, and which in turn arose in response to near fatal collapse of meta-population B caused by elimination of *awsR* as a viable mutational path (Figure 5C).

With selection increasingly focusing on PDEs, *pflu0185* was targeted by mutations multiple times,

prior to this locus acquiring hyper-mutator status. That a single gene can withstand being repeatedly broken and fixed to the extent shown in Figure 6 is remarkable. While the series of mutations was instrumental in the increasing the likelihood that the HNsequence within *pflu0185* , none stood out as obvious candidates. However, noting that switches to CEL$^+$, due to mutations in the open reading frame of *pflu0185* , were often compensated for by mutations in the promoter region, we predicted that mutations in the upstream sequence effected a reduction in c-di-GMP by elevating transcription sufficient to compensate for previously reduced catalytic activity. This indeed proved to be true with repeated interplay between mutations affecting gene function and transcriptional activity eventually elevating transcription ∼15-fold (Figure 6B). Moreover, elevated transcription was associated with significant increases in the rate of local frame-shift mutations. It is this, namely the connection between transcription and frame-shifting, that we suggest was instrumental in enhancing the likelihood that the HNsequence duplicated.

Mechanistic connections between our observed increases in transcription and mutational frame-shifting are unclear, however, transcription is known to be mutagenic, with highly expressed genes being more error prone [12]. Transcription-associated mutagenesis [27] and transcription-coupled repair [20] could both plausibly play roles. While transcription-associated mutagenesis can have pathological effects on genome integrity, it has been harnessed during evolution to drive localized and rapid genetic change as evident in the vertebrate immune system, where transcription is required for somatic hypermutation [56]. Also relevant is that elevated transcription exposes single stranded DNA [62], which is a requirement for slipped-strand mispairing [34]. In this context the 5'-location of the HNsequence is likely a significant factor.

Observation of additional (secondary) adaptive mutations associated with local hyper-mutation of *pflu0185* led us to reason that increased propensity to switch enhances the possibility that secondary adaptive mutations, which would otherwise be stalled by the waiting time required to achieve the primary transition-mediating mutation, have opportunity to increase in frequency (Figure 7B). Indeed, data presented in Figure 7 supports this hypothesis showing how evolution can be sped up but without the threat of Muller's ratchet that arises from global hyper mutation [52, 9]: ability to generate adaptive phenotypic variation provides a 'head-start' to adapt to other aspects of the local environment with feedback affecting further lineage success.

Finally, we turn to broader implications of the study and particularly the extent to which findings here might have general relevance. The selective regime employed was highly contrived with selection on lineages being strictly enforced. It is unlikely that such restrictive conditions are encountered in nature, with the plausible exception of those presented to bacterial pathogens and commensals

that persist in environments where the host immune response is a constant and pervasive factor [40]. Pathogenic bacteria evolving under such conditions are likely to experience strong lineage-level selection with repeated transition through restrictive bottlenecks [41]. As we have shown here, precisely these conditions serve to promote the evolution of evolvability. Indeed, that the genomes of obligate pathogens and commensals are replete with contingency loci bears testimony to this fact.

But we further suggest that selection on lineages is likely more important than generally appreciated. Competing lineages need only differ in evolutionary potential, for example in how adaptive phenotypes are instantiated by mutation. Provided there is sufficient time for phenotypic consequences to manifest, then lineages differing in capacity to generate adaptive variation stand to compete [64]. This is likely to be particularly relevant in situations where transmission / migration involves passage through bottlenecks that require variation to be generated afresh[30, 36, 4, 19]. Moreover, provided that challenges to lineage survival repeat across time, then adaptations that enhance evolvability, shaped by lineage selection for the benefit of lineages, can be expected.

Materials and Methods

Experimental evolution

Both CEL$^+$ and CEL$^-$ were propagated 25 ml glass microcosms (containing 6 ml of King's Medium B (KB)) at 28°C, without shaking. The four meta-populations were founded by single genotypes derived from a previous, but conceptually similar, experiment [24]. Founding genotypes carried on average nine unique mutations, primarily in characterized cyclic-di-GMP regulators.

For the CEL$^+$ to CEL$^-$ phase, microcosms were incubated for six days, vortexed, and 50 μl of a 3.35 $\times$ 10^5-fold dilution spread on a KB agar plate. After 48 hours growth each plate was then examined for the presence of CEL$^-$ colonies, with a single colony used to inoculate a fresh microcosm and initiate a three day growth period. 100 μl of an 8.0 $\times$ 10^6-fold dilution of each microcosm was plated and screened for CEL$^+$ types. The process was iterated. A random number generator was used to determine the extant lineage that would replace extinct lineages. When an entire meta-population failed to generate the target state twice over, propagation was ceased (as in meta-population D).

The births and deaths of lineages were recorded, and the resulting lineage dynamics visualized using Colgen. Overnight cultures of each selected CEL$^-$ and CEL$^+$ colony were stored at -80°C in 33% glycerol saline. The 'stringent selection' experiment followed the same procedure as above but with cells evolved in 96-well plates containing 200μl of KB, and a reduced period of 24 hours given before dilution and screening.

Sequencing and detection of mutations

Visualization and knowledge of genealogy was used to inform both whole-genome sequencing and targeted Sanger sequencing to detect and order mutational events. For meta-population B, 482 clones representing successive time-points of the experiment were sequenced, allowing precise ordering of mutations underpinning the majority of the 520 transitions. Remaining transitions and associated mutations were then inferred by parsimony or confirmed by Sanger sequencing. In one instance, to resolve an ambiguity, a mutation (-197 C$\rightarrow$T *pflu1085*) was re-created to confirm its phenotypic effect.

Genomic DNA was extracted from overnight cultures using the DNeasy DNA Blood & Tissue Kit (Qiagen) according to the manufacturer's spin column protocol. Library preparation and whole-genome sequencing to a depth of
($\sim$50x coverage was conducted at either the Max Planck Institute for Evolutionary Biology (Plön, Germany) or Eurofins Genomics (Constance, Germany) using Illumina NextSeq, MiSeq, or NovaSeq. The resulting reads were assembled against the reference *P. fluorescens* SBW25 genome (GenBank

accession:AM181176.4) using tje clonal option of Breseq with default parameters [15]. Mutations were identified using breseq's HTML output and in the case of predicted novel junctions manually inspected using Geneious Prime. Marginal predictions were confirmed by Sanger sequencing. PCR products for Sanger sequencing were purified with EXOSAP and sequencing conducted by the Max Planck Insitute for Evolutionary biology, prepared using the BigDye Terminator v3.1 cycle sequencing kit followed by the BigDye XTerminator purification kit (ThermoFisher) according to manufacturer's instructions.

Strain construction

Isothermal Gibson Assembly reactions with the NEBuilder HiFi DNA Assembly master mix from New England BioLabs, were used for all molecular cloning procedures. Constructs integrated into the SBW25 genome were flanked by two 400-900 bp fragments with homology to SBW25 to enable recombination and assembled into the transfer plasmid pUISacB (accession). Primer overhangs were generated using the online NEBuilder tool (nebuilder.neb.com) and are available on request. All PCRs were performed using master mixes of high-fidelity polymerases (Phusion or Q5) and purified with the QIAquick PCR purification kit (Qiagen). pUISacB was extracted from overnight cultures using the QIAprep Spin Miniprep Kit (Qiagen), linearized with SpeI, amplified by PCR and digested with DpnI to remove remaining plasmid template. Assembled constructs were transformed into chemically competent *E. coli* Top 10 cells and integrated into SBW25 via two-step allelic exchange as previously described [2], but with final mutant selection using sucrose counter-selection on tryptone yeast extract and sucrose (TYS10) plates (10 g/L tryptone, 5 g/L yeast extract, 10% (W/V) filtered sucrose, 15 g/L agar).

For the *pflu0185* 'frame-shift detector' used to measure local frame-shift rate, a fusion was constructed with *kanR* via the linker sequence 'GCTAGCAGCGCAAGCGCAAGCGCA' (ASSASASA), such that the latter gene was 1bp out of frame with respect to *pflu0185* (an in-frame version of the same *kanR* fusion was also constructed). Three synonymous mutations engineered into the 5' end of *kanR* were required to remove aberrant start codons that existed in the out-of-frame sequence (which otherwise re-started the correct reading-frame). Strains expressing fluorescent proteins for use in fitness assays were constructed via standard conjugation of recipient SBW25 cells and *E. coli* donor containing pMRE-Tn7-145 (scarlet) or pMRE-Tn*7*-152 (GFP) with media containing 0.1% w/v arabinose to induce Tn*7* transposase genes [50].

Transposon mutagenesis

Transposon mutagensis using Tn*5* was conducted according to the previously described method [21]. Briefly, this involved a tri-parental conjugation between the recipient CEL$^+$ and an *E. coli* donor

containing the plasmid pCM639 with the IS-Ω-kan/hah transposon along with *E. coli* containing the helper plasmid pRK2013. Suppressor mutants were identified by screening for CEL$^-$ colonies. The site of insertion was determined by amplifying the transposon-chromosome junction through arbitrarily-primed PCR with Sanger sequencing of the resulting product.

Minimum inhibitory concentrations

Minimum inhibitory concentration (MIC), used as a proxy for transcription, was determined by plating ($\sim 10^8$ cells of the relevant genotypes engineered with an in-frame *pflu0185*:*kanR* fusion onto KB agar and applying a kanamycin MIC test strip (Liofilchem). After 24 hours growth the MIC was determined by identifying where the margin of growth intercepted the strip in three independent replicates. In cases where genotypes exhibited high levels of resistance extending beyond the maximum measurable range of the test strips (256 μg/ml), the assay was instead conducted on agar containing either 100 or 200 μg/ml of kanamycin and the concentrations summed.

Fluctuation assays

Colony counts on selective and non-selective plates were used to estimate mutation rates using the rSalvador R package (v1.9), which generates a maximum likelihood estimate of mutation rate under the Luria-Delbruck model while accounting for variation in plating efficiency.[65] The likelihood ratio method was used to calculate confidence intervals and the statistical significance of differences between mutation rate estimates. Global mutation rates were assessed using rifampicin (100μg/ml) and kanamycin (at concentrations of 10-50μg/ml depending on *pflu0185* expression) used to determine rates of frame-shifting within *pflu0185* . Twelve independent cultures were used to determine spontaneous mutants; 6 were used to determine population size.

Fitness assays

Fitness was measured by 1:1 competition in either shaken (for CEL$^-$) or static (for CEL$^+$) microcosms. Competition took place between strains fluorescently marked strains with either mScarlet-I with sGFP2, that were otherwise isogenic but for the secondary mutation of interest. Starting populations were derived from colonies inoculated into KB microcosms and grown 24 hours at 28°C in a shaking incubator. Competitors were then mixed 1:1 and this mix diluted in filtered PBS (0.22μM filter) with the initial ratio of scarlet to GFP obtained using flow cytometry with the MACSQuant VYB. 6 μl of the 1:1 mix was also used to initiate competition, which ran for either 24 (shaken) or 48 (static) hours. Following competitive growth, each microcosm was vortexed vigorously and in the case of static CEL$^+$ microcosms sonicated to break up cell aggregates. Cultures were again diluted in filtered PBS and the cell counts and final ratio determined. Fitness was then calculated as the ratio of Malthusian parameters [31].

Acknowledgements

The authors acknowledge generous core support from the Max Planck Society.

Author contributions

MB and PBR: experimental concept and design. MB: experimental work and analysis with technical assistance by LZ. MB and PBR wrote the paper.

References

[1] Massimo Amicone, Vítor Borges, Maria João Alves, Joana Isidro, Líbia Zé-Zé, Sílvia Duarte, Luís Vieira, Raquel Guiomar, João Paulo Gomes, and Isabel Gordo. Mutation rate of sars-cov-2 and emergence of mutators during experimental evolution. *Evolution, medicine, and public health*, 10(1):142–155, 2022.

[2] Eleni Bantinaki, Rees Kassen, Christopher G Knight, Zena Robinson, Andrew J Spiers, and Paul B Rainey. Adaptive divergence in experimental populations of pseudomonas fluorescens. III. mutational origins of wrinkly spreader diversity. *Genetics*, 176(1):441–453, May 2007.

[3] Christopher D Bayliss, Dawn Field, E Richard Moxon, et al. The simple sequence contingency loci of haemophilus influenzae and neisseria meningitidis. *The Journal of clinical investigation*,

107(6):657–666, 2001.

[4] Hubertus JE Beaumont, Jenna Gallie, Christian Kost, Gayle C Ferguson, and Paul B Rainey. Experimental evolution of bet hedging. *Nature*, 462(7269):90–93, 2009.

[5] G Bell. The evolution of evolution. *Heredity*, 94(1):1–2, 2005.

[6] Andrew J Black, Pierrick Bourrat, and Paul B Rainey. Ecological scaffolding and the evolution of individuality. *Nat Ecol Evol*, 4(3):426–436, March 2020.

[7] Malgorzata Bzymek and Susan T Lovett. Instability of repetitive dna sequences: the role of replication in multiple mechanisms. *Proceedings of the National Academy of Sciences*, 98(15):8319–8325, 2001.

[8] Yu-Ming Cai, Kai-Wei Yu, Ji-Hong Liu, Zhao Cai, Zun-Hao Zhou, Yang Liu, Tian-Fu Wang, and Liang Yang. The c-di-gmp phosphodiesterase pipa (pa0285) regulates autoaggregation and pf4 bacteriophage production in pseudomonas aeruginosa pao1. *Applied and Environmental Microbiology*, 88(12):e00039–22, 2022.

[9] Alejandro Couce, Larissa Viraphong Caudwell, Christoph Feinauer, Thomas Hindré, Jean-Paul Feugeas, Martin Weigt, Richard E Lenski, Dominique Schneider, and Olivier Tenaillon. Mutator genomes decay, despite sustained fitness gains, in a long-term experiment with bacteria. *Proc. Natl. Acad. Sci. U. S. A.*, 114(43):E9026–E9035, October 2017.

[10] Anton Crombach and Paulien Hogeweg. Evolution of evolvability in gene regulatory networks. *PLoS computational biology*, 4(7):e1000112, 2008.

[11] Nandini Dasgupta and Reuben Ramphal. Interaction of the antiactivator flen with the transcriptional activator fleq regulates flagellar number in pseudomonas aeruginosa. *Journal of bacteriology*, 183(22):6636–6644, 2001.

[12] Abhijit Datta and Sue Jinks-Robertson. Association of increased spontaneous mutation rates with high levels of transcription in yeast. *Science*, 268(5217):1616–1619, 1995.

[13] Richard Dawkins. The evolution of evolvability. In *Artificial life*, pages 201–220. Routledge, 1989.

[14] X De Bolle, C D Bayliss, D Field, T van de Ven, N J Saunders, D W Hood, and E R Moxon. The length of a tetranucleotide repeat tract in haemophilus influenzae determines the phase variation rate of a gene with homology to type III DNA methyltransferases. *Mol. Microbiol.*, 35(1):211–222, January 2000.

[15] Daniel E Deatherage and Jeffrey E Barrick. Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. *Engineering and analyzing multicellular systems: methods and protocols*, pages 165–188, 2014.

[16] Patricia Domínguez-Cuevas and Silvia Marqués. Compiling sigma-70-dependent promoters. *Virulence and gene regulation*, pages 319–343, 2004.

[17] Jeremy Draghi and Günter P Wagner. Evolution of evolvability in a developmental model. *Evolution*, 62(2):301–315, 2008.

[18] Ilan Eshel. Clone selection and the evolution of modifying features. *Theoretical Population Biology*, 4(2):196–208, 1973.

[19] Jenna Gallie, Eric Libby, Frederic Bertels, Philippe Remigi, Christian B Jendresen, Gayle C Ferguson, Nicolas Desprat, Marieke F Buffing, Uwe Sauer, Hubertus JE Beaumont, et al. Bistability in a metabolic network underpins the de novo evolution of colony switching in pseudomonas fluorescens. *PLoS Biology*, 13(3):e1002109, 2015.

[20] Ann Ganesan, Graciela Spivak, and Philip C Hanawalt. Transcription-coupled dna repair in prokaryotes. *Progress in molecular biology and translational science*, 110:25–40, 2012.

[21] Stephen R Giddens, Robert W Jackson, Christina D Moon, Michael A Jacobs, Xue-Xian Zhang, Stefanie M Gehrig, and Paul B Rainey. Mutational activation of niche-specific genes provides insight into regulatory networks and bacterial function in a complex environment. *Proceedings of the National Academy of Sciences*, 104(46):18247–18252, 2007.

[22] Christopher J Graves, Vera ID Ros, Brian Stevenson, Paul D Sniegowski, and Dustin Brisson. Natural selection promotes antigenic evolvability. *PLoS pathogens*, 9(11):e1003766, 2013.

[23] Christopher J Graves and Daniel M Weinreich. Variability in fitness effects can preclude selection of the fittest. *Annual review of ecology, evolution, and systematics*, 48:399–417, 2017.

[24] Katrin Hammerschmidt, Caroline J Rose, Benjamin Kerr, and Paul B Rainey. Life cycles, fitness decoupling and the evolution of multicellularity. *Nature*, 515(7525):75–79, November 2014.

[25] Ruth Hershberg and Dmitri A Petrov. Evidence that mutation is universally biased towards AT in bacteria. *PLoS Genet.*, 6(9):e1001115, September 2010.

[26] Jason W Hickman and Caroline S Harwood. Identification of fleq from pseudomonas aeruginosa as ac-di-gmp-responsive transcription factor. *Molecular microbiology*, 69(2):376–389, 2008.

[27] Sue Jinks-Robertson and Ashok S Bhagwat. Transcription-associated mutagenesis. *Annual review of genetics*, 48:341–359, 2014.

[28] Nadav Kashtan, Elad Noor, and Uri Alon. Varying environments can speed up evolution. *Proc. Natl. Acad. Sci. U. S. A.*, 104(34):13711–13716, August 2007.

[29] Marc Kirschner and John Gerhart. Evolvability. *Proceedings of the National Academy of Sciences*, 95(15):8420–8427, 1998.

[30] Alex K Lancaster and Joanna Masel. The evolution of reversible switches in the presence of irreversible mimics. *Evolution*, 63(9):2350–2362, 2009.

[31] Richard E Lenski, Michael R Rose, Suzanne C Simpson, and Scott C Tadler. Long-term experimental evolution in escherichia coli. i. adaptation and divergence during 2,000 generations. *The American Naturalist*, 138(6):1315–1341, 1991.

[32] Mathilde Lescat, Adrien Launay, Mohamed Ghalayini, Mélanie Magnan, Jérémy Glodt, Coralie Pintard, Sara Dion, Erick Denamur, and Olivier Tenaillon. Using long-term experimental evolution to uncover the patterns and determinants of molecular evolution of an escherichia coli natural isolate in the streptomycin-treated mouse gut. *Molecular ecology*, 26(7):1802–1817, 2017.

[33] G Levinson and G A Gutman. Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Mol. Biol. Evol.*, 4(3):203–221, May 1987.

[34] Gene Levinson and George A Gutman. Slipped-strand mispairing: a major mechanism for dna sequence evolution. *Molecular biology and evolution*, 4(3):203–221, 1987.

[35] Richard C Lewontin. The units of selection. *Annual review of ecology and systematics*, 1(1):1–18, 1970.

[36] Eric Libby and Paul B Rainey. Exclusion rules, bottlenecks and the evolution of stochastic phenotype switching. *Proceedings of the Royal Society B: Biological Sciences*, 278(1724):3574–3583, 2011.

[37] Peter A Lind, Andrew D Farr, and Paul B Rainey. Experimental evolution reveals hidden diversity in evolutionary pathways. *elife*, 4:e07074, 2015.

[38] Michael Lynch. The frailty of adaptive hypotheses for the origins of organismal complexity. *Proceedings of the National Academy of Sciences*, 104(suppl_1):8597–8604, 2007.

[39] Michael J McDonald, Daniel P Rice, and Michael M Desai. Sex speeds adaptation by altering the dynamics of molecular evolution. *Nature*, 531(7593):233–236, 2016.

[40] E Richard Moxon, Paul B Rainey, Martin A Nowak, and Richard E Lenski. Adaptive evolution of highly mutable loci in pathogenic bacteria. *Current biology*, 4(1):24–33, 1994.

[41] Richard Moxon and Edo Kussell. The impact of bottlenecks on microbial survival, adaptation, and phenotypic switching in host-pathogen interactions. *Evolution*, 71(12):2803–2816, December 2017.

[42] Hermann Joseph Muller. The relation of recombination to mutational advance. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 1(1):2–9, 1964.

[43] Candela Muriel, Eva Arrebola, Miguel Redondo-Nieto, Francisco Martínez-Granero, Blanca Jalvo, Sebastian Pfeilmeier, Esther Blanco-Romero, Irene Baena, Jacob G Malone, Rafael Rivilla, and Marta Martín. AmrZ is a major determinant of c-di-GMP levels in pseudomonas fluorescens F113. *Sci. Rep.*, 8(1):1979, January 2018.

[44] Antonio Oliver, Rafael Cantón, Pilar Campo, Fernando Baquero, and Jesús Blázquez. High frequency of hypermutable pseudomonas aeruginosa in cystic fibrosis lung infection. *Science*, 288(5469):1251–1253, 2000.

[45] Joshua L Payne and Andreas Wagner. The causes of evolvability and their evolution. *Nat. Rev. Genet.*, 20(1):24–38, January 2019.

[46] Massimo Pigliucci. Is evolvability evolvable? *Nature Reviews Genetics*, 9(1):75–82, 2008.

[47] Paul B Rainey and Katrina Rainey. Evolution of cooperation and conflict in experimental bacterial populations. *Nature*, 425(6953):72–74, 2003.

[48] Paul B Rainey and Michael Travisano. Adaptive radiation in a heterogeneous environment. *Nature*, 394(6688):69–72, 1998.

[49] Ute Römling, Lian-Ying Cao, and Feng-Wu Bai. Evolution of cyclic di-GMP signalling on a short and long term time scale. *Microbiology*, 169(6), June 2023.

[50] Rudolf O Schlechter, Hyunwoo Jun, Michał Bernach, Simisola Oso, Erica Boyd, Dian A Muñoz-Lintz, Renwick CJ Dobson, Daniela M Remus, and Mitja NP Remus-Emsermann. Chromatic bacteria–a broad host-range plasmid and chromosomal insertion toolbox for fluorescent protein expression in bacteria. *Frontiers in Microbiology*, 9:3052, 2018.

[51] John Maynard Smith and John Maynard-Smith. *The evolution of sex*, volume 4. Cambridge University Press Cambridge, 1978.

[52] P D Sniegowski, P J Gerrish, T Johnson, and A Shaver. The evolution of mutation rates: separating causes from consequences. *Bioessays*, 22(12):1057–1066, December 2000.

[53] Paul D Sniegowski, Philip J Gerrish, and Richard E Lenski. Evolution of high mutation rates in experimental populations of e. coli. *Nature*, 387(6634):703–705, 1997.

[54] Paul D Sniegowski and Helen A Murphy. Evolvability. *Current Biology*, 16(19):R831–R834, 2006.

[55] Andrew J Spiers, Sophie G Kahn, John Bohannon, Michael Travisano, and Paul B Rainey. Adaptive divergence in experimental populations of pseudomonas fluorescens. i. genetic and phenotypic bases of wrinkly spreader fitness. *Genetics*, 161(1):33–46, May 2002.

[56] Ursula Storb, Hong Ming Shen, Nancy Michael, and Nayun Kim. Somatic hypermutation of immunoglobulin and non–immunoglobulin genes. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 356(1405):13–19, 2001.

[57] François Taddei, Miroslav Radman, John Maynard-Smith, Bruno Toupance, Pierre-Henri Gouyon, and Bernard Godelle. Role of mutator alleles in adaptive evolution. *Nature*, 387(6634):700–702, 1997.

[58] Christopher A Vakulskas, Keith M Brady, and Timothy L Yahr. Mechanism of transcriptional activation by pseudomonas aeruginosa exsa. *Journal of bacteriology*, 191(21):6654–6664, 2009.

[59] Berta Verd, Nicholas AM Monk, and Johannes Jaeger. Modularity, criticality, and evolvability of a developmental gene regulatory network. *Elife*, 8:e42832, 2019.

[60] Nathaniel Virgo, Eran Agmon, and Chrisantha Fernando. Lineage selection leads to evolvability at large population sizes. In *Artificial Life Conference Proceedings*, pages 420–427. MIT Press One Rogers Street, Cambridge, MA 02142-1209, USA journals-info . . . , 2017.

[61] Günter P Wagner and Lee Altenberg. Perspective: complex adaptations and the evolution of evolvability. *Evolution*, 50(3):967–976, 1996.

[62] Kenneth D Westover, David A Bushnell, and Roger D Kornberg. Structural basis of transcription: separation of rna from dna by rna polymerase ii. *Science*, 303(5660):1014–1016, 2004.

[63] George Christopher Williams. *Adaptation and natural selection: A critique of some current evolutionary thought*, volume 61. Princeton university press, 1966.

[64] Robert J Woods, Jeffrey E Barrick, Tim F Cooper, Utpala Shrestha, Mark R Kauth, and Richard E Lenski. Second-order selection for evolvability in a large escherichia coli population. *Science*, 331(6023):1433–1436, 2011.

[65] Qi Zheng. rsalvador: An R package for the fluctuation experiment. *G3*, 7(12):3849–3856, December 2017.