Nov 25, 2019

Antagonism in Evolutionary Opportunities Results in Non-Monotonic Evolution Across an Environmental Gradient

Megan G. Behringer[§*], Wei-Chin Ho[§*], Samuel F. Miller, John C. Meraz, Gwyneth F. Boyer, Michael Lynch

Center for Mechanisms of Evolution, The Biodesign Institute, Arizona State University, Tempe, AZ 85287.

[*]These authors contributed equally to this work.

[§]Please direct correspondence to:

Megan G. Behringer: megan.behringer@asu.edu
Wei-Chin Ho: weichinh@asu.edu

Evolutionary outcomes are dictated by both deterministic (e.g., natural selection) and stochastic (e.g., genetic drift or environmental uncertainty) forces[1-8]. The repeated finding of similar phenotypes or genotypes across multiple natural or experimental populations in analogous environments suggest that deterministic forces are strong enough for reinforcing some predictability[8-12]. However, the extent to which the phenotypes discovered in two extreme environments can be interpolated to an intermediate environment remains largely unknown, as increased stochastic forces at intermediate environments cannot be ruled out. Experimental evolution can help uncover the constraints of evolutionary predictability and reveal whether there are non-monotonic patterns of deterministic or stochastic forces across environmental gradients. Here, we perform experimental evolution of *Escherichia coli* under three nutritional transfer periods (TP): every day, every 10 days, and every 100 days, representing different levels of starvation. After 900 days of experimental evolution, intermediately starved cells had evolved a distinct lower growth rate and higher mutation rate, neither of which were found in the other extreme treatments. Because the intermediately starved populations exhibit significantly high molecular parallelism, their distinct phenotypes are likely due to non-monotonic deterministic forces instead of stronger stochastic forces. Our results demonstrate novel complexities associated with evolutionary predictability across environmental gradients.

Natural microbial populations, including facultative pathogens with clinical or agricultural importance, frequently encounter environments with scarce resources[13,14]. As an opportunistic pathogen, *E. coli*'s habitat is broad, ranging from soil and wastewater to the lower gut, often oscillating between feast and famine[15-17]. Further, some *E. coli* are capable of surviving during periods of prolonged starvation, accumulating mutations in the genes *rpoS*[18] and *lrp*[19] that confer

24 a phenotype known as growth advantage in stationary phase (GASP)[20-22]. Other phenotypes
25 observed in *E. coli* during resource limitation involve increased mutation rates[23,24], as populations
26 experiencing repeated exposure to resource limitation have exhibited the evolution of trade-offs
27 between maximum growth rate and sustainable growth[25]. Although these studies and others report
28 genetic and phenotypic responses to resource limitation, there has been a lack of replication across
29 parallel treatments or varying timescales. As such, conclusions that the phenotypic response to
30 starvation, or any selective pressure, scales progressively is premature. To address this issue, we
31 surveyed the evolutionary response of populations along a selective gradient to determine if there
32 are any constraints on the genotypes and phenotypes associated with long-term starvation.

33 To establish the evolution experiment, we used ancestral lines with two different genetic
34 backgrounds (GB): a wild-type strain (WT) and a WT-derived strain with impaired methyl-
35 directed mismatch repair (MMR-) wherein an engineered deletion of *mutL* yields a ~150× increase
36 in the single nucleotide mutation rate[26]. With three TPs (one day, 10 days, and 100-days), each
37 TP/GB combination was replicated in eight parallel experiments, resulting in 48 experimental
38 populations (8×3×2), allowing us to examine the parallelism within each treatment combination.
39 After 900 days of experimental evolution, we quantified the growth behaviours. Notably, 10-day
40 populations take significantly longer to reach their maximum growth rate ($T_{\max}$) when assayed in
41 96-well plates than ancestral, 1- and 100-day populations (**Fig. 1a**, WT: t_{anc} = 14.91, P_{anc} = 4.5 x
42 10^{-10}; $t_{1\text{-}day}$ = 11.65, $P_{1\text{-}day}$ = 4.6 x 10^{-10}; $t_{100\text{-}day}$ = 1.99, $P_{100\text{-}day}$ = 0.057; MMR-: t_{anc}= 7.24, P_{anc} = 2.5
43 x 10^{-5}; $t_{1\text{-}day}$ = 7.81, $P_{1\text{-}day}$ = 8.4 x 10^{-6}; $t_{100\text{-day}}$ = 1.87, $P_{100\text{-}day}$ = 0.073; pairwise Welch's t-test, two-
44 sided, adjusted by FDR). In addition, 10-day populations present significantly smaller colonies on
45 LB agar than ancestral, 1- and 100-day populations (**Fig. 1b-e**, WT: t_{anc} = 44.99, P_{anc}< 1 x10^{-16};
46 $t_{1\text{-}day}$ = 5.56, $P_{1\text{-}day}$ = 3.3 x 10^{-4}; $t_{100\text{-}day}$ = 4.85, $P_{100\text{-}day}$ = 5.6 x 10^{-4}; MMR-: t_{anc} = 16.10, P_{anc} = 5.4

47 x 10^{-10}; $t_{1\text{-}day} = 6.16$, $P_{1\text{-}day} = 4.5 \times 10^{-5}$; $t_{100\text{-}day} = 2.22$, $P_{100\text{-}day} = 0.039$; pairwise Welch's t-test, two-
48 sided, adjusted by FDR). This severe reduction in growth rate in the 10-day populations illustrates
49 just one of the non-monotonically evolving phenotypes observed in response to repeated resource
50 limitation.

51 To further study the underlying molecular mechanisms for phenotypic evolution and
52 determine if the rate of genomic evolution in each of the TP/GB conditions follows the same
53 discordant pattern as growth rate, we performed high-throughput metagenomic sequencing every
54 100 days (mean coverage > 100×, **Supplementary Table 1.1**) and estimated genomic divergence
55 by summing the derived allele frequencies (DAFs) for all single nucleotide polymorphisms (SNPs)
56 at each time point. The rate of genomic evolution is then defined as the slope in the linear
57 regression of summed DAFs against the number of generations. We found that in both the WT and
58 MMR- backgrounds, the rates in 10-day populations are the highest followed by the rates in 100-
59 day populations (**Fig. 2a**). This increased rate of genomic evolution is still observed when other
60 measurements for genomic divergence are used (**Extended Data Fig. 1**). Nonetheless, when
61 comparing the probability of fixation in nonsynonymous, intergenic, and synonymous mutations,
62 we found no evidence suggesting that stronger positive selection drives the elevated rate of genome
63 evolution in 10-day populations (**Extended Data Fig. 2**). Therefore, below we will examine two
64 other candidate factors: mutation rates and subpopulation structure.

65 Previous studies have illustrated that *E. coli* can experience increased mutation rates during
66 stationary phase[23]. Thus, we performed fluctuation tests[27] to determine how resource limitation
67 had affected the mutation rate and if more resource-depleted conditions had selected for newly
68 arising mutators or anti-mutators. We observed significant effects of TP in both genetic
69 backgrounds (WT: $P = 1.4 \times 10^{-5}$; MMR-: $P = 0.014$; Kruskal-Wallis test, **Fig. 2b**). Ten-day clones

70 surprisingly evolved significantly higher mutation rates than daily clones (WT: P = 8.9 x 10^{-6};
71 MMR-: P = 5.6 x 10^{-3}, Mann-Whitney U test) but 100-day clones did not (WT: P = 0.48; MMR-:
72 P = 0.068, Mann-Whitney U test), once more depicting a non-monotonic pattern in phenotypic
73 evolution as the intermediate, 10-day environment produced the most extreme mutation rate.

74 Another candidate factor that may increase evolutionary rates is subpopulation structure[28].
75 As it has previously been shown that our culture environment supports subpopulation structure in
76 daily-transferred populations[29], we wanted to determine how TP affected genetic diversity. We
77 used a clade-aware hidden Markov chain (caHMM) which assumes the coexistence of two clades
78 (major and minor) to infer which mutations belong to the basal clade (i.e., mutations that sweep
79 through the entire population before the establishment of subpopulations) and which belong to
80 either the major or minor clade[30] (**Fig. 3a**). caHMM revealed that TP (P = 6.1 x 10^{-6}, Scheirer-
81 Ray-Hare test) affects the stability of subpopulation structure (measured by the duration that two
82 clades are observed to coexist; **Fig. 3b**) and that subpopulation structure in 100-day populations is
83 significantly less stable than in daily populations (P = 3.1 x 10^{-5}, Mann-Whitney U test) and 10-
84 day populations (P = 4.2 x 10^{-5}, Mann-Whitney U test). Consistently, the proportion of fixed
85 mutations inferred to belong to the basal clade in a population is negatively correlated with the
86 stability of subpopulation structure (**Extended Data Fig. 3**) and is greatest in 100-day populations
87 (**Extended Data Fig. 4**). Plotting the temporal distribution of DAFs reveals that mutations with
88 intermediate DAFs (0.3-0.7) are less abundant in 100-day populations (**Extended Data Fig. 5**),
89 consistent with a lower tendency to form subpopulation structure. This lower tendency is unlikely
90 due to insufficient evolutionary time. For example, by 300 days, daily-transferred lines have
91 experienced enough genomic evolution to promote subpopulation structure in 13 of 16 populations.
92 However, by 500 days (calculated based on **Fig. 2a**), the same amount of genomic evolution only

93 results in 2 of 16, 100-day populations exhibiting subpopulation structure. Therefore, TP
94 contributes to the increased rate of genomic evolution, but the extremely long nutrient-replenishing
95 time experienced by 100-day populations reduces the likelihood of establishing stable
96 subpopulation structure.

97 We previously observed that biofilm formation was one of the major phenotypes emerging
98 in our culture conditions and was generally indicative of the evolution of subpopulation structure[29].
99 Here, we observed that while daily and 10-day populations produce visible biofilms at the surface-
100 air interface, 100-day populations do not (**Fig. 3c**). If biofilm formation is essential to the
101 establishment of stable subpopulation structure in our culture tube environment, then the absence
102 of biofilm formation in 100-day populations is consistent with the inability of the caHMM to detect
103 the stable coexistence of multiple clades within 100-day populations. Further, the thickness of the
104 biofilms appears greater in 10-day populations, revealing another non-monotonically evolving
105 phenotype.

106 Lastly, to seek evidence for positive selection and distinguish which mutations are likely
107 adaptive across the nutrient-limitation gradient from those promoted to fixation due to genetic
108 draft[31], we focused on the nonsynonymous mutations and calculated for each gene a *G*-score
109 indicating the degree to which parallel mutations are overrepresented[32].A higher sum of all G-
110 scores indicates that observed mutations tend to be found in a smaller subset of genes within a
111 TP/GB combination. Compared to the simulated null distribution, the observed sum of all *G*-scores
112 is higher in all TP/GB combinations (**Fig. 4a-b; Supplementary Table 1.2**), revealing a
113 significant amount of parallel mutation and suggesting that positive selection is shaping these
114 populations. Interestingly, the statistical significance of *G*-scores for 10-day and 100-day
115 populations is much greater than for daily populations, suggesting that the pool of adaptive

116 mutations for 10-day and 100-day populations is more limited, possibly due to their environmental
117 harshness. A similar analysis also reveals significant overrepresentation of genes affected by
118 parallel indels or IS-element insertions (**Fig. 4c; Supplementary Table 1.3**) highlighting
119 parallelism in likely loss-of-function mutations. In addition to genic-level parallelism, we
120 identified a set of 115 nonsynonymous mutations that arose in at least two experimental
121 populations in a TP/GB combination (**Extended Data Fig. 6; Supplementary Table 1.4**). Such
122 nucleotide-level parallelism again suggests positive selection in action.

123 By determining the functional implications of genes under positive selection, we can begin
124 to understand the cellular processes that must be altered in order to sustain repeated starvation. GO
125 analysis of the list of genes overrepresented for nonsynonymous SNPs revealed that populations
126 subjected to extended resource limitation (10- and 100-day) are enriched for mutations in genes
127 related to global gene expression (regulation, termination, and fidelity of transcription and
128 translation) as well as resource import and metabolism (**Supplementary Table 1.5-7**). Of the
129 genes associated with gene expression and metabolism, we notably observe mutations in genes
130 that regulate or are regulated by the stringent response including *crp* and *sspA*. We also observed
131 mutations in essential genes involved in gene expression: a nonsynonymous mutation in the
132 transcription terminator *rho* that arose in four 100-day populations; and two mutations in the
133 translation elongation and ribosome recycling factor, *fusA*, that arose across five 10-day
134 populations (**Supplementary Table 1.4**). We additionally focused on candidate mutations with
135 the potential to influence the observed changes in mutation rates (GO:0006281 - DNA Repair or
136 GO:0006260 - DNA replication) and found more parallel mutations in 10-day populations than in
137 1- or 100-day populations (**Supplementary Table 1.8**). This finding is consistent with the non-
138 monotonic pattern in mutation-rate evolution. Further studies are needed to understand how

139 alteration of genes responsible for basic molecular processes, such as replication, transcription,
140 and translation benefit populations experiencing repeated starvation.

141 Several implications can be drawn from the reported results. First, if the evolution of
142 cooperation is key to the emergence of population substructure[29], our findings suggest cooperation
143 is harder to evolve in extremely resource-limited environments, possibly due to a combination of
144 low mutation rates and limited pool of beneficial mutations. This echoes the findings that resource-
145 deficient environments promote detachment of biofilm formers[33] or metastasis of cancer cells[34].

146 Second, genes involved in global gene expression and resource import/metabolism are
147 targets for adaptive mutations in starved populations. However, despite our 100-day populations
148 spending the vast majority of their time in long-term stationary phase, we strikingly observed a
149 lack of mutations affecting genes commonly associated with the growth advantage in stationary
150 phase (GASP) phenotype , such as *rpoS*[18,35] and *lrp*[19]. Further investigation is needed to determine
151 if the absence of these mutations is due to strain specific/culture environment differences or
152 because these mutations are transient, with benefits that are no longer realized after repeated
153 exposure to extreme starvation.

154 Third, while elevated mutation rates are often found associated with stressful
155 environments[36] and are observed in 10-day populations, the extremely long nutrient-replenishment
156 cycles of the 100-day populations lowered the tendency towards elevated mutation rates.
157 Population-genetic models have suggested that higher mutation rates can be tolerated in
158 populations experiencing constant environmental fluctuations[37-39]. Our experimental results of 10-
159 day populations suggest that, when environmental fluctuations are neither too rapid nor too
160 infrequent, mutator alleles are best tolerated and able to fix. Consistent with a reduced tolerance
161 for increased mutation rates, in both 1-day and 100-day populations, it is common to observe a

162 reduction in mutation rate across individuals with MMR- backgrounds (**Fig. 2**), although the
163 genetic source of this reduction is not apparent. Clearly, these changes in the evolved mutation
164 rate are determined by a multiplicity of factors requiring more detailed study.

165 Finally, we observed that the responses of many molecular and phenotypic traits are non-
166 monotonic across an environmental gradient of nutrient limitation. One possible reason for the
167 deceleration of trait evolution in 100-day populations may be related to an economic theory known
168 as 'the maximization principle' (**Fig. 4d**). Although populations may experience an increase in
169 evolutionary opportunity with more extreme levels of resource limitation (e.g., increases in
170 beneficial mutation rates, distance from fitness optima), they also experience increases in
171 evolutionary constraints (e.g., increases in deleterious mutation rates, and decreases in effective
172 population sizes, resource availability, and accessible mutational trajectories). If both the total
173 evolutionary opportunity and evolutionary constraint increase such that their curves recede and
174 then approach, then the greatest net evolutionary opportunity will exist in an intermediate
175 environment. Consequently, the largest magnitude of adaptation will not exist at the most extreme
176 environments, leading to non-monotonic trait evolution. These results raise questions about the
177 predictability of evolution and advocate a need for more investigation into molecular and
178 phenotypic evolution along environmental gradients.
179
180
181
182
183
184

185 **References**

186 1 Lassig, M., Mustonen, V. & Walczak, A. M. Predicting evolution. *Nat Ecol Evol* **1**, 77

187 2 Jain, K. & Krug, J. Deterministic and stochastic regimes of asexual evolution on rugged
188 fitness landscapes. *Genetics* **175**, 1275-1288

189 3 Rouzine, I. M., Rodrigo, A. & Coffin, J. M. Transition between stochastic evolution and
190 deterministic evolution in the presence of selection: General theory and application to
191 virology. *Microbiol Mol Biol R* **65**, 151-185

192 4 Kimura, M. Stochastic processes and distribution of gene frequencies under natural
193 selection. *Cold Spring Harb Sym* **20**, 33-53

194 5 Fisher, R. A. On the dominance ratio. *Bull Math Biol* **52**, 297-318

195 6 Nosil, P. *et al.* Natural selection and the predictability of evolution in Timema stick insects.
196 *Science* **359**, 765-770

197 7 Chaturvedi, S. *et al.* The predictability of genomic changes underlying a recent host shift
198 in Melissa blue butterflies. *Mol Ecol* **27**, 2651-2666

199 8 Blount, Z. D., Lenski, R. E. & Losos, J. B. Contingency and determinism in evolution:
200 Replaying life's tape. *Science* **362**, eaam5979

201 9 Rosenblum, E. B., Parent, C. E. & Brandt, E. E. The molecular basis of phenotypic
202 convergence. *Annu. Rev. Ecol. Evol. Syst* **45**, 203-226

203 10 Parker, J. *et al.* Genome-wide signatures of convergent evolution in echolocating mammals.
204 *Nature* **502**, 228-231

205 11 Stern, D. L. The genetic causes of convergent evolution. *Nat Rev Genet* **14**, 751-764

206 12 Losos, J. B. Convergence, adaptation, and constraint. *Evolution* **65**, 1827-1840

207 13 Barton, I. S., Fuqua, C. & Platt, T. G. Ecological and evolutionary dynamics of a model
208 facultative pathogen: *Agrobacterium* and crown gall disease of plants. *Environ Microbiol*
209 **20**, 16-29
210 14 Brown, S. P., Cornforth, D. M. & Mideo, N. Evolution of virulence in opportunistic
211 pathogens: generalism, plasticity, and control. *Trends Microbiol* **20**, 336-342
212 15 Johnson, J. R. & Russo, T. A. Extraintestinal pathogenic *Escherichia coli*: "The other bad
213 *E. coli*". *J Lab Clin Med* **139**, 155-162
214 16 Smith, J. L., Fratamico, P. M. & Gunther, N. W. Extraintestinal pathogenic *Escherichia*
215 *coli*. *Foodborne Pathog Dis* **4**, 134-163
216 17 Jang, J. *et al.* Environmental *Escherichia coli*: ecology and public health implications-a
217 review. *J Appl Microbiol* **123**, 570-581
218 18 Farrell, M. J. & Finkel, S. E. The growth advantage in stationary-phase phenotype
219 conferred by *rpoS* mutations is dependent on the pH and nutrient environment. *J Bacteriol*
220 **185**, 7044-7052
221 19 Zinser, E. R. & Kolter, R. Prolonged stationary-phase incubation selects for *lrp* mutations
222 in *Escherichia coli* K-12. *J Bacteriol* **182**, 4361-4365
223 20 Zambrano, M. M. & Kolter, R. GASPing for life in stationary phase. *Cell* **86**, 181-184
224 21 Finkel, S. E. & Kolter, R. Evolution of microbial diversity during prolonged starvation. *P*
225 *Natl Acad Sci USA* **96**, 4023-4027
226 22 Finkel, S. E. Long-term survival during stationary phase: evolution and the GASP
227 phenotype. *Nat Rev Microbiol* **4**, 113-120
228 23 Sniegowski, P. Evolution: bacterial mutation in stationary phase. *Curr Biol* **14**, R245-R246

229 24 Loewe, L., Textor, V. & Scherer, S. High deleterious genomic mutation rate in stationary
230 phase of *Escherichia coli*. *Science* **302**, 1558-1560
231 25 Ying, B. W. *et al.* Evolutionary consequence of a trade-off between growth and
232 maintenance along with ribosomal damages. *PLoS One* **10**, e0135639
233 26 Lee, H., Popodi, E., Tang, H. X. & Foster, P. L. Rate and molecular spectrum of
234 spontaneous mutations in the bacterium *Escherichia coli* as determined by whole-genome
235 sequencing. *Proc Natl Acad. Sci. USA* **109**, E2774-E2783
236 27 Foster, P. L. Methods for determining spontaneous mutation rates. *Methods Enzymol* **409**,
237 195-213
238 28 Nei, M. & Takahata, N. Effective population-size, genetic diversity, and coalescence time
239 in subdivided populations. *J Mol Evol* **37**, 240-244
240 29 Behringer, M. G. *et al. Escherichia coli* cultures maintain stable subpopulation structure
241 during long-term evolution. *Proc Natl Acad Sci USA* **115**, E4642-E4650
242 30 Good, B. H., McDonald, M. J., Barrick, J. E., Lenski, R. E. & Desai, M. M. The dynamics
243 of molecular evolution over 60,000 generations. *Nature* **551**, 45-50
244 31 Gillespie, J. H. Genetic drift in an infinite population. The pseudohitchhiking model.
245 *Genetics* **155**, 909-919
246 32 Tenaillon, O. *et al.* Tempo and mode of genome evolution in a 50,000-generation
247 experiment. *Nature* **536**, 165-170
248 33 Hunt, S. M., Werner, E. M., Huang, B. C., Hamilton, M. A. & Stewart, P. S. Hypothesis
249 for the role of nutrient starvation in biofilm detachment. *Appl Environ Microb* **70**, 7418-
250 7425

251 34 Garcia-Jimenez, C. & Goding, C. R. Starvation and pseudo-starvation as drivers of cancer
252 metastasis through translation reprogramming. *Cell Metab* **29**, 254-267
253 35 Lange, R. & Hengge-Aronis, R. The nlpD gene is located in an operon with rpoS on the
254 Escherichia coli chromosome and encodes a novel lipoprotein with a potential function in
255 cell wall formation. *Mol Microbiol* **13**, 733-743
256 36 Williams, A. B. & Foster, P. L. Stress-induced mutagenesis. *EcoSal Plus* **5**
257 37 Lukacisinova, M., Novak, S. & Paixao, T. Stress-induced mutagenesis: Stress diversity
258 facilitates the persistence of mutator genes. *PLoS Comput Biol* **13**, e1005609
259 38 Ishii, K., Matsuda, H., Iwasa, Y. & Sasaki, A. Evolutionarily stable mutation rate in a
260 periodically changing environment. *Genetics* **121**, 163-174
261 39 Tanaka, M. M., Bergstrom, C. T. & Levin, B. R. The evolution of mutator genes in bacterial
262 populations: the roles of environmental change and timing. *Genetics* **164**, 843-854
263

264 **Figure Legends**

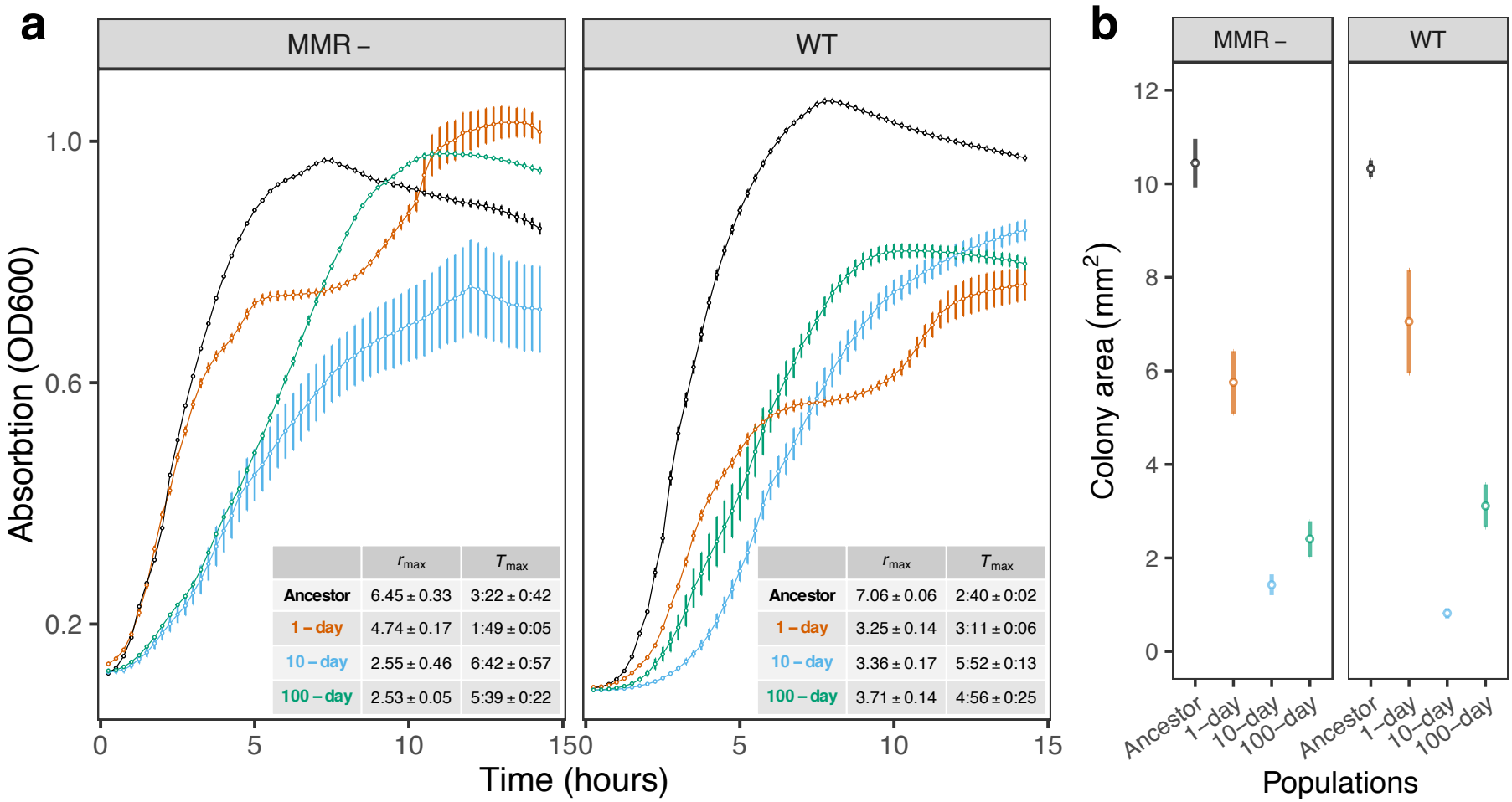
265 **Figure 1 Phenotypic divergence of evolved lines after 900 days. (a)** Growth curves of ancestor
266 (black) and evolved 1- (orange), 10- (blue), and 100-day (green) populations cultured in LB
267 broth at 37 °C. Circles denote average absorbance at 600 nm measured every 15 minutes for 14
268 hours. Error bars represent ± s.e.m.. Nested table displays the averages ± s.e.m. for growth
269 parameters, r_{max} and T_{max}. T_{max} is presented in units of time (hours:minutes). **(b)** Average colony
270 sizes of evolved populations and ancestors after 24 h of growth on LB agar. Error bars represent
271 ± s.e.m. Photos of **(c)** 1-day (Population 108), **(d)** 10-day (Population 408), and **(e)** 100-day
272 (Population 508) populations co-cultured with the WT ancestor on TA agar and incubated at
273 37°C for 24 h to show variation in colony size. Colonies from evolved populations appear red
274 while WT ancestor colonies appear beige. Inset in **d** provides a magnified look at ancestor and
275 10-day colonies. The black arrow points to a 10-day population colony, highlighting the evolved
276 extreme reduction in colony size.

277

278 **Figure 2 Evolutionary dynamics of evolved populations. (a)** Each panel shows the mean rate
279 of genomic evolution among the parallel experimental populations for each transfer period (1,
280 10, or 100 days) in MMR- (mismatch repair defective) or WT (wild-type) genetic backgrounds.
281 At each time point, the circle and the bar represent the mean and the standard error. From left to
282 right represent the sequencing profiles sampled at days 0, 90, 200, 300, 400, 500, 600, 700, 800,
283 and 900. Note that the data for populations with 100-day transfer periods are missing at day 500.
284 The line represents the linear regression against the number of generations. The estimated slope
285 (*b*) and corresponding standard error are also shown in each panel. **(b)** Each panel shows
286 mutation rates of populations after 900 days of experimental evolution in different combinations

287 of transfer period (1, 10, or 100-days) and genetic background (MMR- for mismatch repair
288 defective or WT for wild-type). In each combination, four evolved lines were tested, and two or
289 four clones per evolved line were isolated and measured. The open circle and error bar represent
290 the mean and the 95% confidence interval for each clone. The grey dashed line represents the
291 mutation-rate measurement of the corresponding ancestor.
292
293 **Figure 3 Ecological dynamics in experimental evolution.** Muller plots of major (yellow) and
294 minor (pink) clade abundance for (**a**) Population 101 (1-day, WT) and Population 503 (100-day,
295 WT). Black arrows denote the maximum duration that two clades co-exist. (**b**)The distribution
296 of durations of coexistence of major and minor clades as determined by the caHMM among the
297 parallel experimental populations for each transfer period/genetic background combination. **(c)**
298 Presence/absence of biofilm at the surface area of evolved cultures after 24 h. Surface biofilm
299 can be observed in daily and 10-day populations but is absent in 100-day populations and the
300 ancestor.
301
302 **Figure 4 Evolutionary parallelism as evidence for positive selection. (a)** The arrow and
303 vertical dashed line show the observed sum of *G*-scores, representing the extent of parallel
304 mutation for a particular TP/GB combination. The histogram shows the distribution of 20,000
305 simulated sums of *G*-scores. The significance of the observed sums can be evaluated by *z*-scores
306 ($z > 1.65$ for one-tailed $P < 0.05$). **(b)** List of genes with at least one nonsynonymous mutation
307 that are likely under positive selection. Significance levels (simulated *P*-values with Bonferroni
308 correction) are shown by the different non-black colours of tiles. Genes with no such hits in a
309 particular TP/GB combination are shown by black tiles. **(c)** List of genes significantly

310 overrepresented for structural mutations that are likely under positive selection. Yellow tiles
311 highlight genes observed in at least two populations and yielding simulated *P*-values < 0.05 after
312 Bonferroni correction. **(d)** A hypothetical model describing the non-monotonic evolution of traits
313 observed across environmental gradients. For example, if the environmental gradient represents
314 environmental changes that range from 'more similar' to 'more different' from the ancestral
315 environment (x-axis, black), then factors such as increased beneficial mutation rates due to stress
316 induced mutagenesis or phenotypic distance from fitness optima may contribute to evolutionary
317 opportunity (left y-axis, blue). However, this evolutionary opportunity may be bounded by
318 factors including increased deleterious mutation rates due to stress induced mutagenesis, lower
319 effective population sizes due to more severe bottlenecks and higher lethality, or deep fitness
320 valleys across the adaptive landscape, contributing to evolutionary constraints (right y-axis, red).
321 If along the environmental gradient, evolutionary opportunities increase in a concave-down or
322 linear fashion while evolutionary constraints increase in a concave-up or linear fashion and both
323 are non-parallel, then the maximum net-evolutionary opportunity (dotted lines, purple) may be
324 greatest in intermediate environments resulting in non-monotonic evolution.



	r_{max}	T_{max}
Ancestor	6.45 ± 0.33	3:22 ± 0:42
1 – day	4.74 ± 0.17	1:49 ± 0:05
10 – day	2.55 ± 0.46	6:42 ± 0:57
100 – day	2.53 ± 0.05	5:39 ± 0:22

	r_{max}	T_{max}
Ancestor	7.06 ± 0.06	2:40 ± 0:02
1 – day	3.25 ± 0.14	3:11 ± 0:06
10 – day	3.36 ± 0.17	5:52 ± 0:13
100 – day	3.71 ± 0.14	4:56 ± 0:25

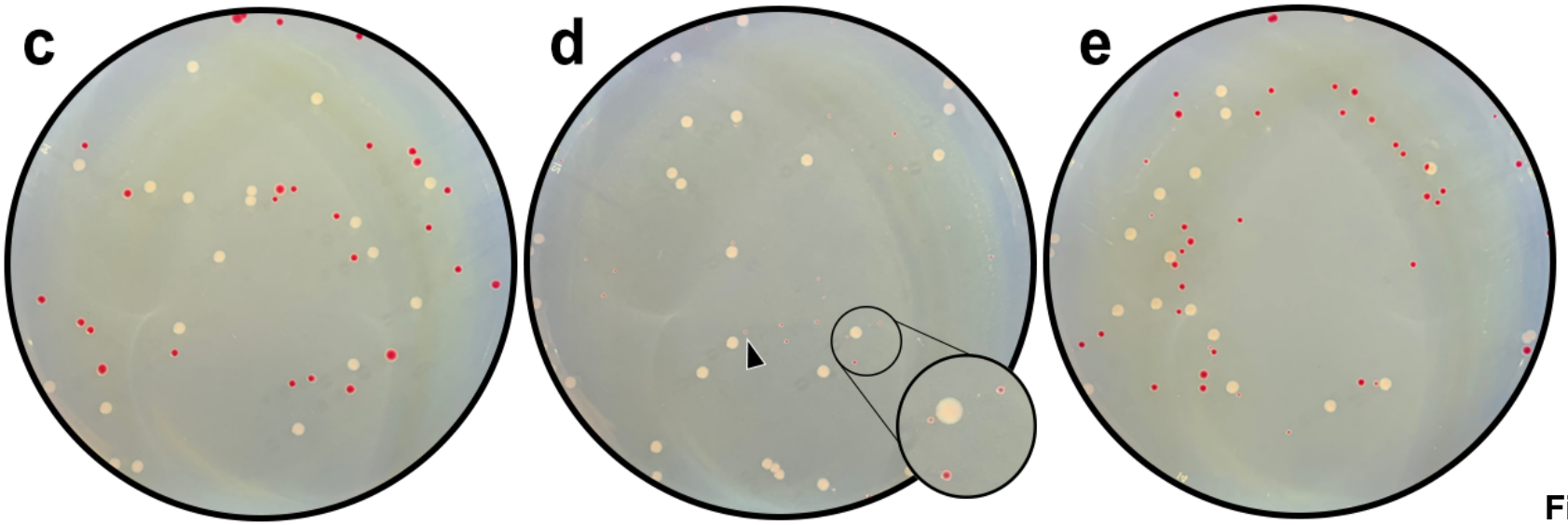


Fig.1

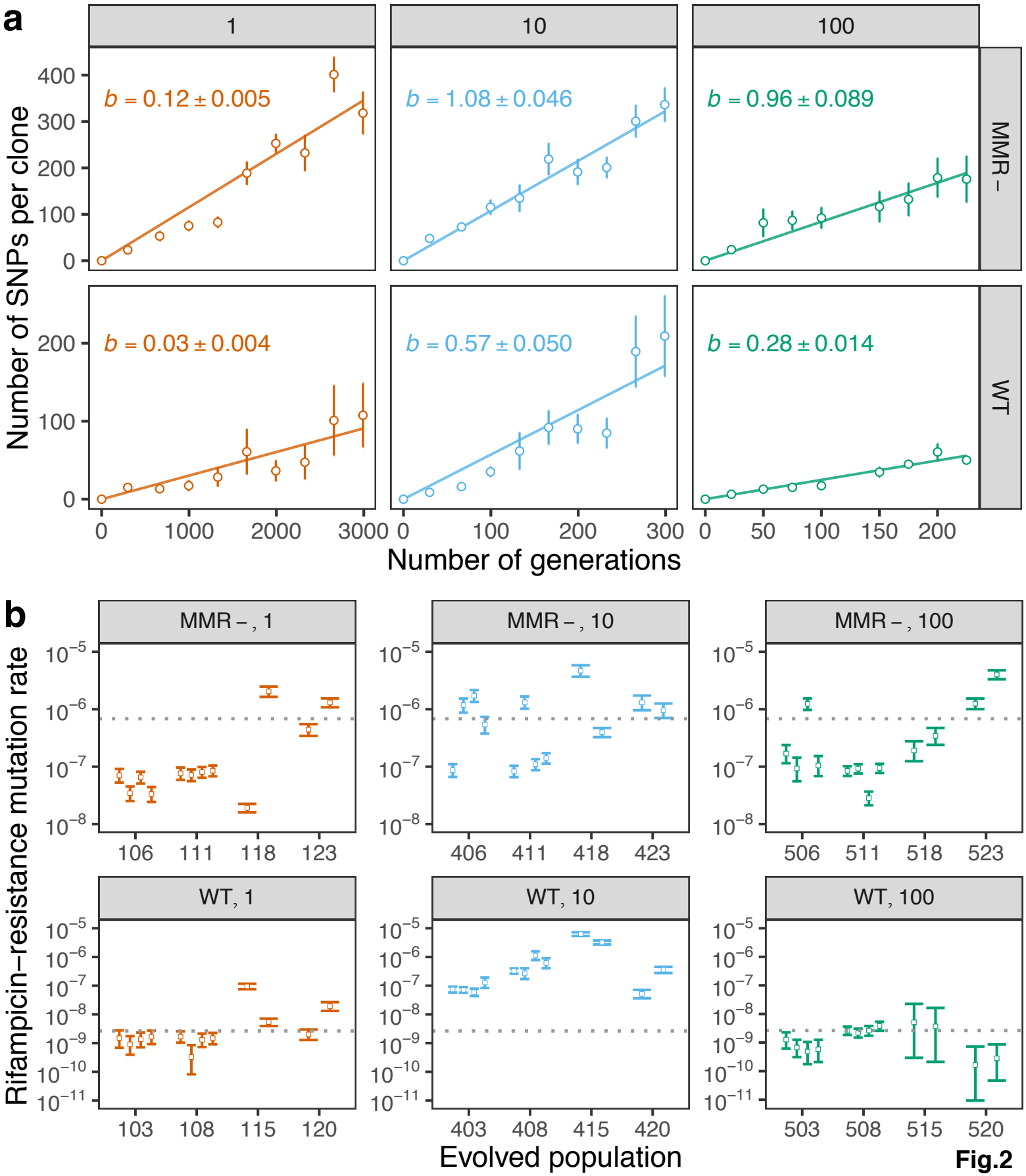

a
1
10
100
MMR –
WT
b = 0.12 ± 0.005
b = 1.08 ± 0.046
b = 0.96 ± 0.089
b = 0.03 ± 0.004
b = 0.57 ± 0.050
b = 0.28 ± 0.014
Number of SNPs per clone
Number of generations
b
MMR –, 1
MMR –, 10
MMR –, 100
WT, 1
WT, 10
WT, 100
Rifampicin–resistance mutation rate
Evolved population

Fig.2

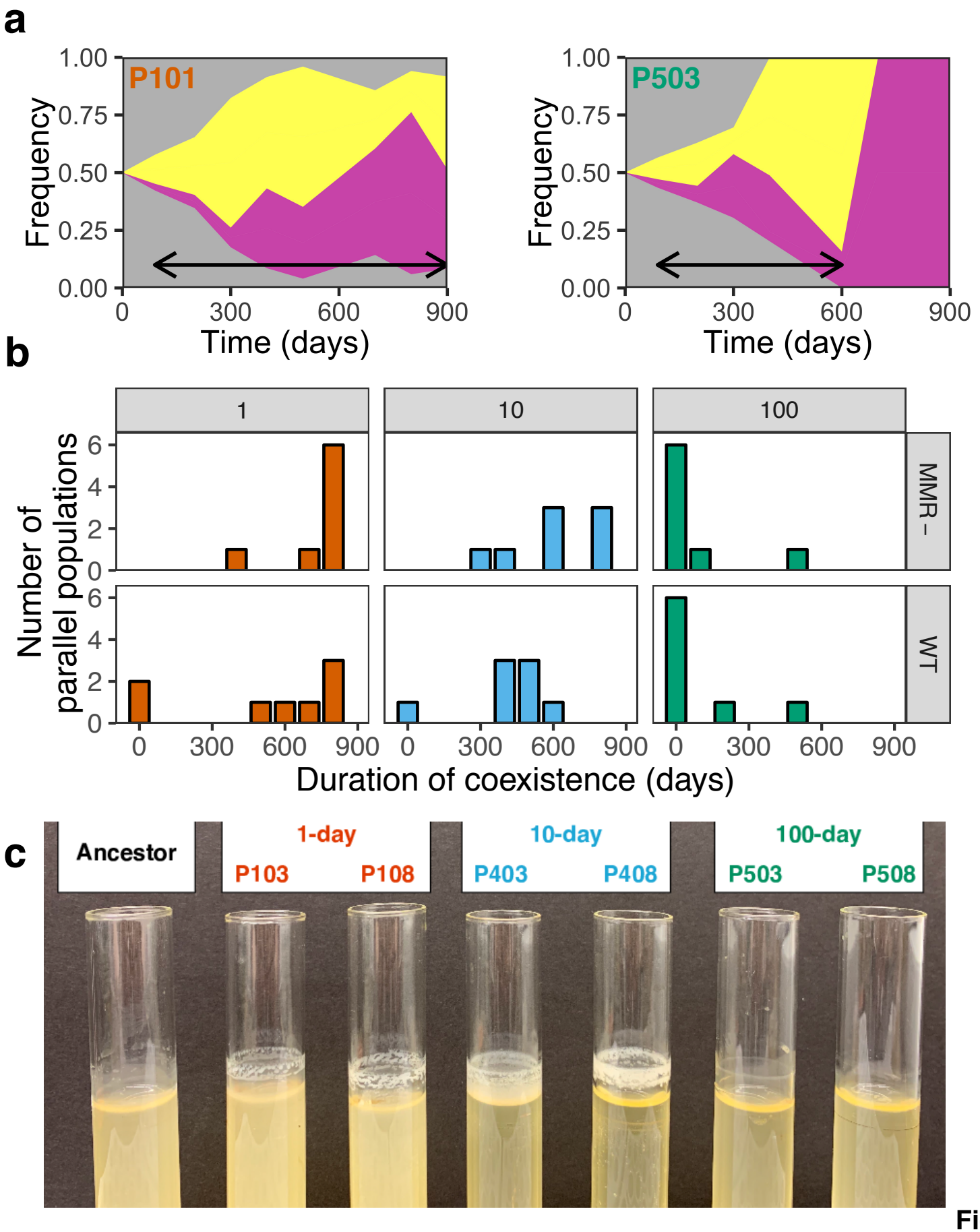

a
P101
P503
Frequency
1.00
0.75
0.50
0.25
0.00
0
300
600
900
Time (days)
b
1
10
100
MMR−
WT
Number of parallel populations
0
2
4
6
Duration of coexistence (days)
c
Ancestor
1-day
P103
P108
10-day
P403
P408
100-day
P503
P508

Fig.3

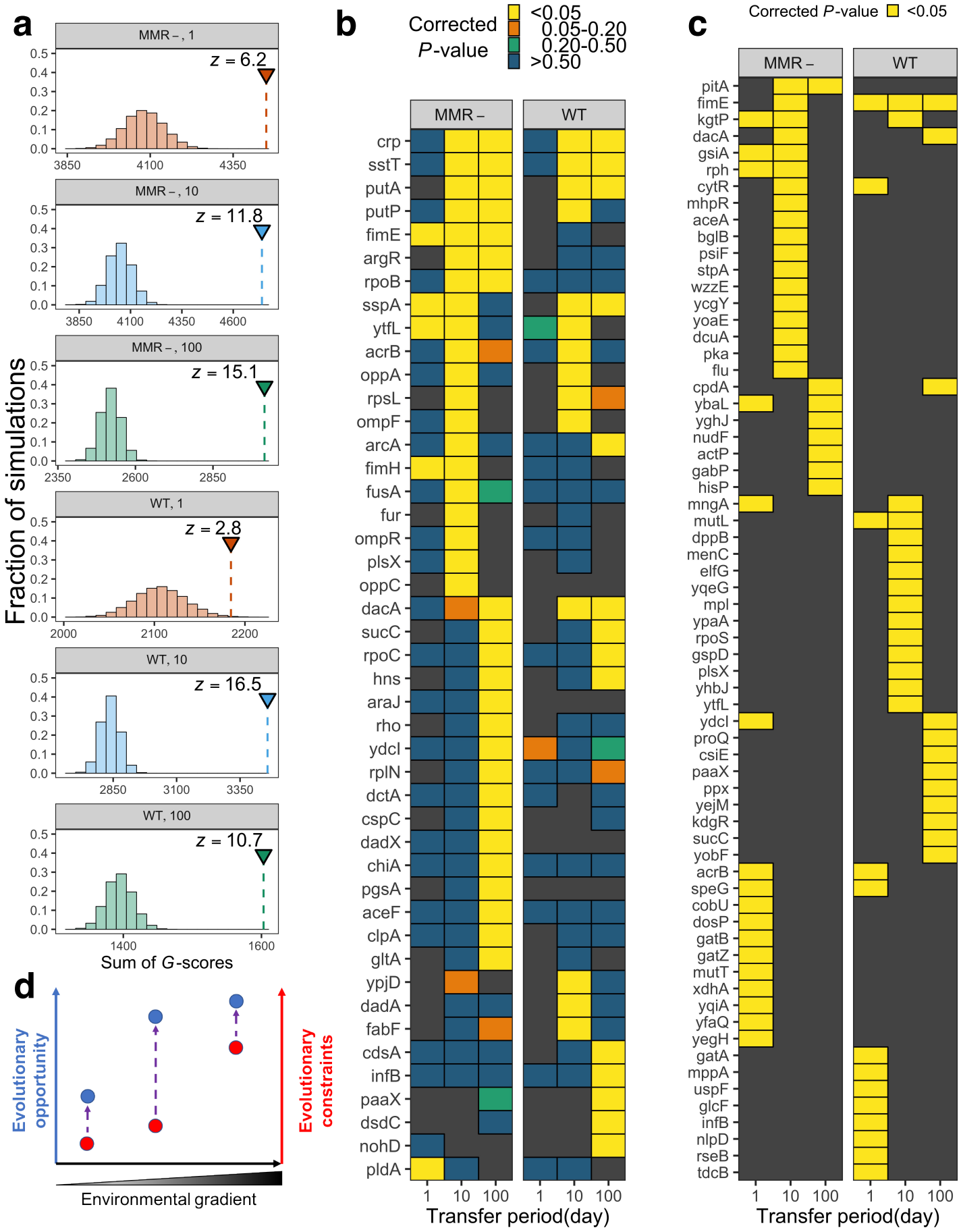

a
MMR – , 1
z = 6.2
MMR – , 10
z = 11.8
MMR – , 100
z = 15.1
WT, 1
z = 2.8
WT, 10
z = 16.5
WT, 100
z = 10.7
Fraction of simulations
Sum of G-scores
b
Corrected P-value
<0.05
0.05-0.20
0.20-0.50
>0.50
MMR –
WT
crp
sstT
putA
putP
fimE
argR
rpoB
sspA
ytfL
acrB
oppA
rpsL
ompF
arcA
fimH
fusA
fur
ompR
plsX
oppC
dacA
sucC
rpoC
hns
araJ
rho
ydcI
rplN
dctA
cspC
dadX
chiA
pgsA
aceF
clpA
gltA
ypjD
dadA
fabF
cdsA
infB
paaX
dsdC
nohD
pldA
Transfer period(day)
c
Corrected P-value <0.05
MMR –
WT
pitA
fimE
kgtP
dacA
gsiA
rph
cytR
mhpR
aceA
bglB
psiF
stpA
wzzE
ycgY
yoaE
dcuA
pka
flu
cpdA
ybaL
yghJ
nudF
actP
gabP
hisP
mngA
mutL
dppB
menC
elfG
yqeG
mpl
ypaA
rpoS
gspD
plsX
yhbJ
ytfL
ydcI
proQ
csiE
paaX
ppx
yejM
kdgR
sucC
yobF
acrB
speG
cobU
dosP
gatB
gatZ
mutT
xdhA
yqiA
yfaQ
yegH
gatA
mppA
uspF
glcF
infB
nlpD
rseB
tdcB
Transfer period(day)
d
Evolutionary opportunity
Evolutionary constraints
Environmental gradient

Fig.4

325 **Materials and Methods**

326 **Source of starting cell lines**

327 Experimental populations originated from one of four ancestral strains. Populations
328 denoted as WT are descendants of PMF2, a prototrophic derivative of *E. coli* K-12 str. MG1655,
329 while populations denoted as MMR- are descendants of PMF5 (PMF2, Δ*mutL*) provided by the
330 Foster Lab. To introduce a neutral marker that would allow for the periodic screening of
331 contamination between populations, lambda red recombineering was used to introduce a 3513 bp
332 deletion removing the *araBAD* operon from one-half of both the WT and MMR-starting strains.
333 This deletion was confirmed by the appearance of red colonies after plating on TA agar (1%
334 Arabinose, 1%Tryptone, 0.5% NaCl, 0.1% Yeast Extract, 0.005% TTC (Sigma T8877)) and next-
335 generation sequencing.

336

337 **Experimental evolution**

338 Long-term cultures were established by inoculating 10 mL of LB-Miller broth (BD Difco)
339 with a single-isolated progenitor colony cultivated overnight at 37 °C on LB agar plates. Resulting
340 cultures were maintained in 16- × 100 mm glass culture tubes containing 10 mL LB broth shaking
341 at 175 rpm at 37°C. Selection conditions involved subjecting cultures to varying degrees of
342 repeated starvation in which 1 mL of thoroughly-vortexed culture was transferred into 10 mL of
343 fresh LB broth every day, every 10-days, or every 100-days. To maintain an historical record of
344 the populations and to revive populations in the case of contamination or extinction, 1 mL of
345 culture was collected pretransfer every 30 days from daily and 10-day populations, frozen with
346 liquid nitrogen, and stored at -80°C. In order to not disrupt any potential ecological structure in
347 100-day populations, copies of these cultures were only collected pretransfer every 100-days. In

348 addition, 100-day populations were maintained in triplicate (1 main culture and 2 backup cultures)
349 with only the main culture being used to seed new triplicate cultures every 100 days. In the case
350 of a whole population extinction, one backup culture would replace the main culture and the
351 experiment would continue. This practice was only necessary during the first 200-days as past this
352 time point, no additional whole-population extinction events occurred. Populations were screened
353 for cross-contamination on days 90, 200, 300, 400, 500, 600, 700, 800, and 900 by streaking on
354 MacConkey Agar (BD Difco) supplemented with 0.4% Arabinose.
355
356 **Evaluation of growth phenotypes**
357 Growth curves were measured using an Epoch2 Microplate Spectrophotometer (Biolog).
358 To prepare cultures for growth rate measurement, 900-day populations and day-0 progenitor
359 strains were revived frozen -80°C stock by transferring a scrape of ice to 16- x 100 mm culture
360 tubes containing 10 mL of LB broth. Cultures were incubated overnight at 37°C and shaking at
361 175 rpm. Following overnight culture, 1.5 µL of each revived population or progenitor were
362 transferred into a well in a 96-well flat bottom culture plate (Corning/Falcon, 353072) containing
363 150 µL of LB broth. Evolved populations were measured as an average of eight replicates and
364 progenitor strains were measured as an average of 16 replicates, with an additional eight wells
365 dedicated to negative controls consisting solely of LB broth. After inoculation, lids were placed
366 on 96-well plates and were sealed with parafilm before placement in the spectrophotometer.
367 Throughout the duration of growth curve measurement, cultures were maintained at 37°C, with
368 double orbital shaking at 548 cpm. Growth measurements were taken every 15m over the course
369 of 16h at an OD of 600 nm. Maximum growth rate ($r_{\max}$) was calculated for all measurable
370 replicates (WT: n_{ans} = 32, $n_{1\text{-}day}$ = 16, $n_{10\text{-}day}$ = 16, $n_{100\text{-}day}$ =16; MMR: n_{ans} = 29, $n_{1\text{-}day}$ = 16, $n_{10\text{-}day}$

371 = 12, $n_{100\text{-}day}$ =16) using the Gen5 software (Biotek) as the maximum slope between two timepoints
372 and time to r_{max} (T_{max}) was recorded as the midpoint between r_{max} determining timepoints. For
373 daily lines experiencing diauxic growth, calculations of r_{max} and T_{max} were restricted to the first
374 logistic growth phase.

375 To measure colony size, revived 900-day populations and day-0 progenitor strains were
376 grown overnight at 37°C and shaking at 175 rpm in 16- x 100 mm culture tubes containing 10 mL
377 of LB broth. Following overnight incubation, serial dilutions of 100 µL of culture in 900 µL of
378 phosphate buffered saline were made down to dilutions of 10^{-5}, before plating 100 µL of diluted
379 culture on to TA agar (1.6% agar, 1% arabinose, 1% tryptone, 0.5% NaCl, 0.1% yeast extract,
380 0.0005% TTC) and incubating for 24h at 37°C. TA plates were then photographed with a 20 mm
381 scale in frame using an iPhone XS. Photos were then analysed with ImageJ (NIH,
382 https://imagej.nih.gov/ij/), by first setting the pixel to distance ratio using the in-frame scale before
383 flatting the photo to 8-bit black and white, setting the threshold to highlight the colonies, and using
384 the 'measure particles' tool. Measurements for six arbitrary colonies per analysed population were
385 recorded.

386 Presence and absence of biofilm at the surface-air interface was determined by reviving
387 900-day populations and day-0 progenitor strains in 16- x 100 mm culture tubes containing 10 mL
388 of LB broth overnight at 37°C and shaking at 175 rpm. Following incubation, culture tubes were
389 visually examined for the presence of biofilm attached to the culture tube at the surface-air
390 interface.

391

392

393

394 **DNA isolation and high-throughput sequencing**

395 High resolution population tracking was conducted by collecting 1 mL of culture at day 90, 200,
396 300, 400, 500, 600, 700, 800, and 900 of experimental evolution. DNA was extracted using the
397 DNeasy UltraClean Microbial Kit (Qiagen 12224; formerly MO BIO UltraClean Microbial DNA
398 Kit) and submitted to either the Hubbard Center for Genomic Analysis at the University of New
399 Hampshire, the Center for Genomics and Bioinformatics at Indiana University, or the CLAS
400 Genomics Facility at Arizona State University for library preparation and sequencing. Libraries
401 were generated with the Nextera DNA Library Preparation Kit (Illumina, FC-121-1030) following
402 an augmented protocol for optimization of reagent use[40] before being pooled and sequenced as
403 paired-end reads to a target depth of 100x on an Illumina HiSeq 2500 (UNH) or an Illumina
404 NextSeq 500 (Indiana; ASU).

405

406 **Sequencing analysis**

407 Sequencing analysis was performed on the Mason and Carbonate high-performance computing
408 clusters at Indiana University. Sequencing reads were quality controlled using Cutadapt v.1.9.1[41]
409 to remove residual adapters and trim low quality sequences. QCed sequences were then mapped
410 to the *Escherichia coli* K-12 substr. MG1655 reference genome (NC_000913.3) and all mutations
411 and their frequencies were identified using Breseq v.0.30.2 with the predict-polymorphisms
412 parameter setting[42]. All mutations that were identified to previously exist in ancestral lines were
413 discarded from the following analysis. Further, our analysis only included the samples which
414 passed the following additional four quality checks: 1) we required mean sequencing depths > 10;
415 2) any WT sample identified to contain the 1,830 bp deletion in *mutL* from the PMF5 progenitor
416 strain was discarded; 3) regions lacking sequencing coverage (i.e. depth = 0) must be smaller than

417 5% of the genome; and 4) the sequencing result should reflect the correct genetic background in
418 terms of *ara* markers, including a nonsynonymous SNP at position 66528, an intergenic SNP at
419 position 70289, and a multiple base substitution mutation (SUB) at position 66533. For an *ara+*
420 line, we required either of two SNPs showing DAF < 0.2. For an *ara-* line, we required either of
421 two SNPs showing DAF > 0.8 or the SUB is detected. In the end, 395 genomic profiles passed QC
422 and were included in the following analysis (**Supplementary Table 1.1**).

423 In order to remove the mutations that originated from the starting clone before experimental
424 evolution, we discarded any mutations with a DAF = 100% at one time point for at least 11
425 experimental populations with the same genetic background from the analysis. In addition, as the
426 highly repetitive sequences in *rsx* genes are known to cause errors in SNP calling[43], they were also
427 removed from the analysis.

428

429 **Guaranteed generations**

430 The calculation of guaranteed number of generations is based on the following two observations.
431 First, pre-transfer population density in the culture tube was measured at day 200, 300, 400, 500,
432 600, 700, 800, 900, by serial dilution in phosphate-buffered saline (PBS) before plating on LB agar
433 for CFU counts and revealed that the population density does not significantly change throughout
434 experimental evolution (**Extended Data Fig. 7**). This suggests the carrying capacity of
435 experimental populations is recovered within a transfer period after 1:10 dilution and the
436 populations must have experienced at least 3.3 ($\log_2 10$) generations between transfer periods.
437 Therefore, the first estimate of guaranteed generations (g_1) per day equals 3.3, 0.33, and 0.033
438 when the transfer period is one day, 10 days, and 100 days, respectively. Second, we found that
439 between transfer periods novel mutations arise and are able to be fixed within one sequencing

interval. Therefore, given the carrying capacity K, the population should at minimum experience $\log_2 K$ cell divisions in 100 days. Using estimated $K = 1.21 \times 10^{10}$, 2.13×10^{8}, and 3.71×10^{7} (**Extended Data Fig. 7**), the second estimate of guaranteed generations (g_2) per day equals 0.30, 0.28, and 0.25 when transfer period is one day, 10 days, and 100 days, respectively. The final number of guaranteed generations (g) for each transfer period is then determined as the larger value between the corresponding g_1 and g_2.

Given that K was only estimated pre-transfer, when the feeding cycle is long it is possible that populations have reached to a density K^* higher than K and decreased back to K within one feeding cycle. If this is true, the first estimated guaranteed generations (g_1') per day should be $[\log_2(K^*/K) + 3.3]/\mathrm{TP}$, where TP is the transfer period. While g_1' is likely to be larger than g, the observed fold-changes in the genomic evolutionary rates between in longer transfer periods cannot be fully explained by this possibility. For example, if the 9.4-fold change in 10-day lines compared to daily lines were entirely due to this possible underestimation, g_1' needs to be 9.4-fold higher than g_1, which requires an implausible $K^* = 4.7 \times 10^{16}$. Similarly, if the 8.8-fold changes in 100-day lines compared to daily lines were entirely due to this possible underestimation, g_1' needs to be 8.8-fold higher than g, which also requires an implausible $K^* = 7.2 \times 10^{64}$. Therefore, even if we had tracked the population density throughout the entire experimental evolution, it is likely that we still could not find substantial effects of changing population densities on our observations.

Rate of genomic evolution

We quantified the level of genomic divergence for each experimental population at each time point by summing all DAFs of detected mutations. Then we calculated the mean genomic divergence across all eligible experimental populations in each TP/GB combination

463 (**Supplementary Table 1.1**). When estimating the rate of genomic evolution, the linear regression
464 was performed by the function "lm" in R with formula "mean genomic divergence ~ guaranteed
465 generations + 0" (linear model), which enforces the y-intercept as 0.

466 To explore the possibility of nonlinear relationship between the genomic divergence and
467 time, we also performed regression using the formula "mean genomic divergence ~ guaranteed
468 generations + square root of guaranteed generations + 0" (nonlinear model), which was previously
469 proposed to catch the trend of diminishing returns[32]. However, the results of nested ANOVA test
470 suggest no strong evidence supporting that the nonlinear model is substantially better than the
471 linear model (**Extended Data Fig. 8**).

472

473 **Fluctuation test and mutation rate estimation**

474 To quantify the genetic mutation rate of our evolved populations and their ancestors, we performed
475 fluctuation tests as described[22] on clones isolated from each evolved population at 900 days.
476 Briefly, fluctuation tests measure the rate of resistance to the antimicrobial rifampicin which is
477 conferred by mutations to *rpoB*. For each TP/GB combination, the four *ara*- populations were
478 assayed. For the first two populations with smaller population numbers, four independent clones
479 were picked; for the last two populations with larger population numbers; two independent clones
480 were picked. For the WT or MMR- ancestor, five independent clones were used. For each clone,
481 40 replicate experiments were performed. Number of mutants as determined by CFU counts/mL
482 were converted to an estimated mutation rate by the function "newton.LD" function in the R
483 package "rSalvador"[44] for each clone.

484

485

486 **Clade-aware hidden Markov chain analysis (caHMM) and longest coexistence time**

487 To test coexistence time in each experimental population, we performed clade-aware
488 hidden Markov model (caHMM) using a modified version of previously released code[30]. Our
489 modifications included: 1) modifying the code in "calculate_clade_hmm_wrapper.py," in order to
490 continue the iteration of experimental populations, even when the major clade frequency of all
491 time points for an experimental population cannot be estimated; 2) removing the original cut-off
492 for minimum coexistence time, as that parameter is too long for our experiments; and 3) increasing
493 the iteration number for hidden Markov chain from 5 to 50. When performing caHMM, we
494 constructed an annotated time course of mutations using the original format. If a mutation is a
495 structural mutation (DEL, INS, and MOB), the sequencing depth is determined by the mean
496 sequencing depth across the mapped adjacent nucleotide ("coverage_plus" or "coverage_minus")
497 of associated junction candidates (JCs) in breseq annotated file. The sequencing depth of time 0
498 was set as 100.

499 Using the major clade frequency (f_M) and minor clade frequency (f_m) estimated by caHMM**,**
500 we calculated the longest coexistence time as follows: if there is no time point showing $0.2 < f_M$
501 <0.8 and $0.2 < f_m < 0.8$, its longest length of coexistence is defined as zero; otherwise, we found
502 the longest time intervals within which every time point shows $0.01 < f_M < 0.99$ or $0.01 < f_m < 0.99$.

503 After calculating the longest length of coexistence for each population, to test whether TP
504 and GB affects the longest length of coexistence, we performed a Scheirer-Ray-Hare test,
505 (nonparametric version of two-way ANOVA) using the function "scheirerRayHare" in the R
506 package "rcompanion"[45]. The sample size for each TP/GB combination is eight. The succeeding
507 *post hoc* analysis is done by Mann-Whitney *U* test with the function "wilcox.test" in R.

508

509 **Identification of candidate mutations for positive selection**

510 For the populations where caHMM does not finish, we instead performed well-mixed
511 hidden Markov chain (wmHMM) using the code "calculate_well_mixed_hmm_wrapper.py,"
512 which also infers the fixation status of a mutation under a one-clade model. The single clade in
513 wmHMM is defined as the basal clade. Candidate mutations for adaptation are then defined as
514 mutations that satisfy either of the following two criteria: (1) the mutations recognized as fixed
515 mutations in any of the basal, major, or minor clade in the output file from caHMM; (2) the
516 mutations that show DAF > 0.5 in at least two time points.

517

518 **Calculation of *G*-scores and GO term enrichment**

519 To test whether the set of candidate nonsynonymous mutations in a TP/GB combination
520 were influenced by positive selection, we calculated *G*-score for each gene[32]. A larger *G*-score
521 means more overrepresentation for a gene. To calculate *G*-scores, we first counted the observed
522 number of fixed nonsynonymous mutations in gene *i* per TP/GB combination (O_i), and the
523 expected number for gene *i* (E_i) was calculated by $O_{tot}(L_i/L_{tot})$, where $O_{tot} = \Sigma_i O_i$, L_i is the number
524 of nonsynonymous sites for gene *i,* and $L_{tot} = \Sigma_i L_i$. The *G*-score for gene *i* (G_i) was then calculated
525 by $2O_i\ln(O_i/E_i)$ or defined as zero when $O_i = 0$. Note that $2O_i\ln(O_i/E_i)$ could be smaller than zero
526 when $0 < O_i < E_i$. This can lead to a paradox that G_i is smaller when a gene was hit by a few times
527 compared to the situation that the same gene was not hit at all. Therefore, we also defined $G_i = 0$
528 when $2O_i\ln(O_i/E_i) < 0$ in the following analysis.

529 Because the null expectation of *G*-scores varies with total number of fixed nonsynonymous
530 mutations (**Extended Data Fig. 9**), we performed 20,000 simulations in each of which O_{tot} hits are
531 randomly distributed among all L_{tot} sites across all genes in the genome. To evaluate the

532 significance of the sum of G-scores (**Fig. 3a**), we calculated the corresponding z score by (the
533 observed sum - mean of simulated sums) / (standard deviation of simulated sums). To evaluate the
534 significance of G-score for gene i, we directly compared the G_i to the distribution of 20,000
535 simulated G_i. The P-value was determined by the proportion of simulated G_i larger or equal to the
536 observed G_i. The multiple test correction was performed by multiply each gene's P-value by the
537 number of genes with at least one hit by the set of fixed nonsynonymous mutations (Bonferroni
538 correction). Only the genes with Bonferroni corrected P-value < 0.05 are called significant. Using
539 this set of significant genes for each TP/GB combination, we performed enrichment tests of gene
540 ontology biological processes (BP) terms using the function "enrichGO", the organismal database
541 org.EcK12.eg.db, and a q-value cut-off = 0.05 in R package "DOSE"[46].
542

543 **Overrepresentation of the genes affected by structural mutations**

544 To test whether the set of genes that were affected by structural mutations (indels and IS-
545 element insertions) in a TP/GB combination were influenced by positive selection, we also
546 calculated a G-score for each gene[32]. To calculate G-scores, we first counted the observed number
547 of populations with any structural mutations in gene i per TP/GB combination (O_i), and the
548 expected number for gene i (E_i) was calculated by $O_{tot}(L_i/L_{tot})$, where $O_{tot} = \Sigma_i O_i$, L_i is the gene
549 length for gene i, and $L_{tot} = \Sigma_i L_i$. The G-score for gene i (G_i) was then calculated by $2O_i\ln(O_i/E_i)$,
550 following the methods described in the above section.

551 We also performed 20,000 simulations and determined the Bonferroni corrected P-value
552 for each gene i following the methods described in the above section. As a result, we found all the
553 genes with $O_i \geq 2$ show Bonferroni corrected P-value < 0.05.
554

555 **Data Availability**

556 All resulting sequences have been deposited to NCBI's short read archive as BioProject:
557 PRJNA532905. All non-sequencing derived data needed to make figures 1A, 1B, and 2B can be
558 found on GitHub (https://github.com/LynchLab/ECEE_Starvation).

559

560 **Code Availability**

561 All custom scripts and R code utilized to analyse data are deposited on GitHub
562 (https://github.com/LynchLab/ECEE_Starvation).

563

564 **Methods-Only References**

565 40 Baym, M. *et al.* Inexpensive multiplexed library preparation for megabase-sized genomes.
566 *PLoS One* **10**, e0128036

567 41 Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads.
568 *EMBnet.journal* **17**, 10-12

569 42 Deatherage, D. E. & Barrick, J. E. Identification of mutations in laboratory-evolved
570 microbes from next-generation sequencing data using breseq. *Methods Mol Biol* **1151**, 165-188

571 43 McCloskey, D. *et al.* Evolution of gene knockout strains of E. coli reveal regulatory
572 architectures governed by metabolism. *Nature communications* **9**, 3796

573 44 Zheng, Q. rSalvador: An R package for the fluctuation experiment. *G3 (Bethesda)t* **7**, 3849-
574 3856

575 45 Mangiafico, S. S. *Summary and analysis of extension program evaluation in R*. (2016).

576 46 Yu, G. C., Wang, L. G., Yan, G. R. & He, Q. Y. DOSE: an R/Bioconductor package for
577 disease ontology semantic and enrichment analysis. *Bioinformatics* **31**, 608-609

578 **Acknowledgements**

579 We thank B. Choi, T. Doak, D.A. Drummond, P. Foster, W. Guo, J. T. Lennon, H. Long, J.
580 McKinlay, S. Snyder, E. Thorley, A. Urquidez, S. Vidrios, and S. Walls for their assistance and
581 helpful comments. High Performance Computing Resources were provided and maintained by the
582 National Center for Genome Analysis Support at Indiana University. This work was supported by
583 Army Research Office Grants ARO65308-LS-MU and W911NF-14-1-0411 and National
584 Institutes of Health Grant F32GM123703 and R35GM122566.

585

586 **Author Contributions**

587 M.G.B., W-C.H., and M.L. designed research; M.G.B., W-C.H., S.F.M., G.F.B., J.C.M.,
588 performed experiments; M.G.B, W-C.H. and G.F.B. conducted bioinformatic analyses; M.G.B.,
589 and W-C.H., and S.F.M, analysed data; and M.G.B., W-C.H., and M.L. wrote the paper.

590

591 **Ethics declarations**

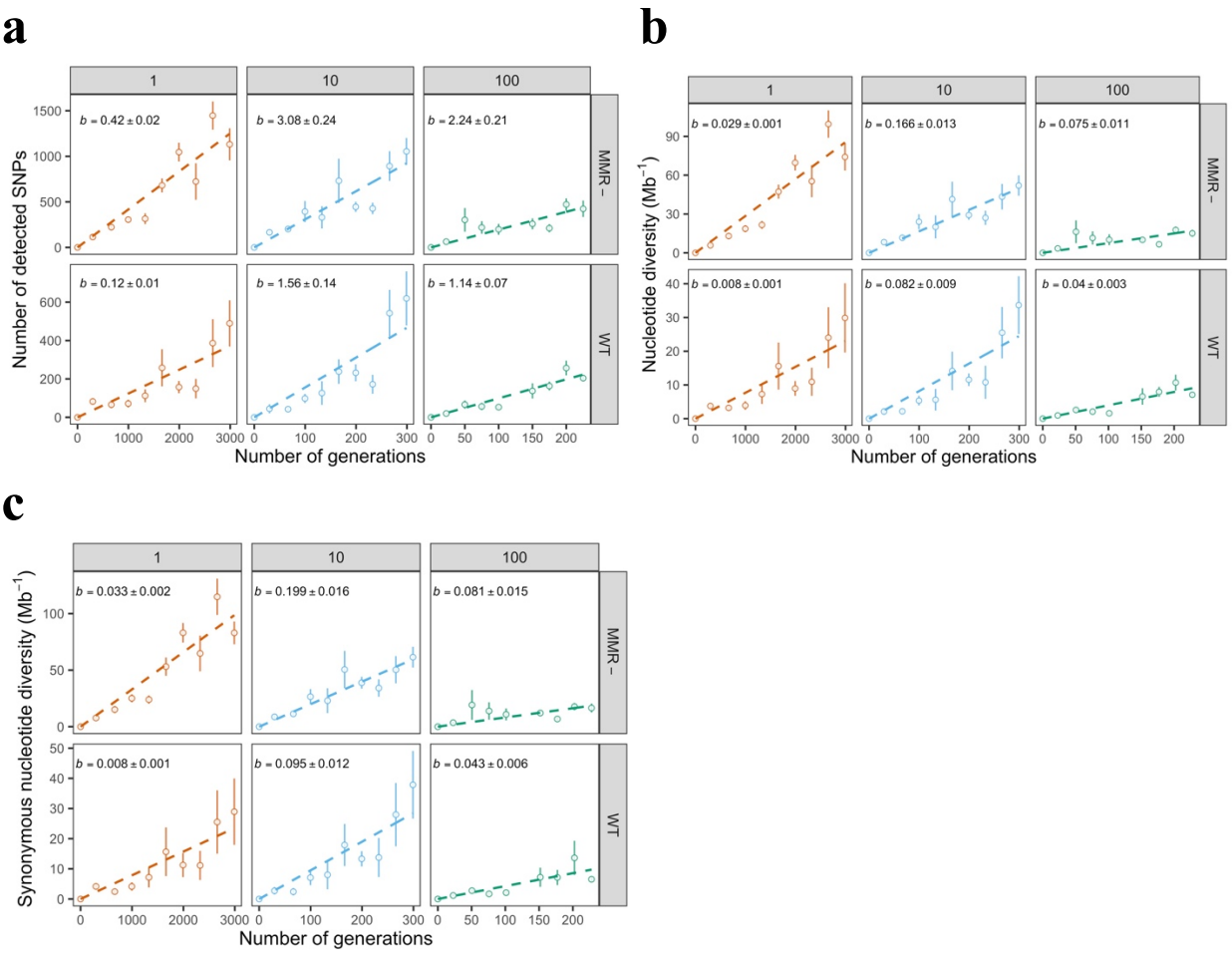
592 The authors declare no competing financial interests

593

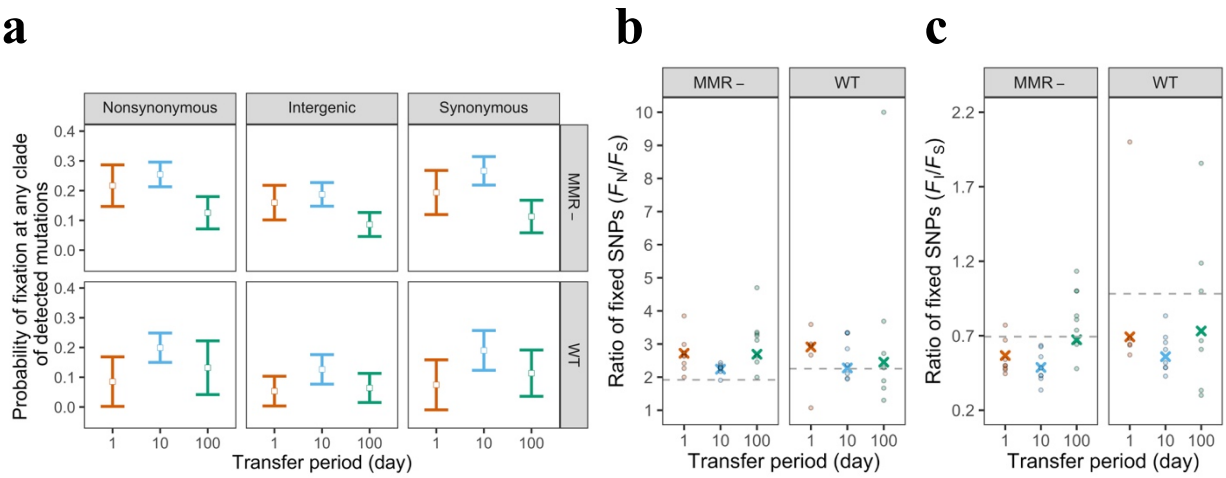
594 **Additional Information**

595 **Supplemental Information data** is available for this paper.

596 Correspondence and requests for materials should be addressed to Megan Behringer
597 (megan.behringer@asu.edu) and Wei-Chin Ho (weichinh@asu.edu)

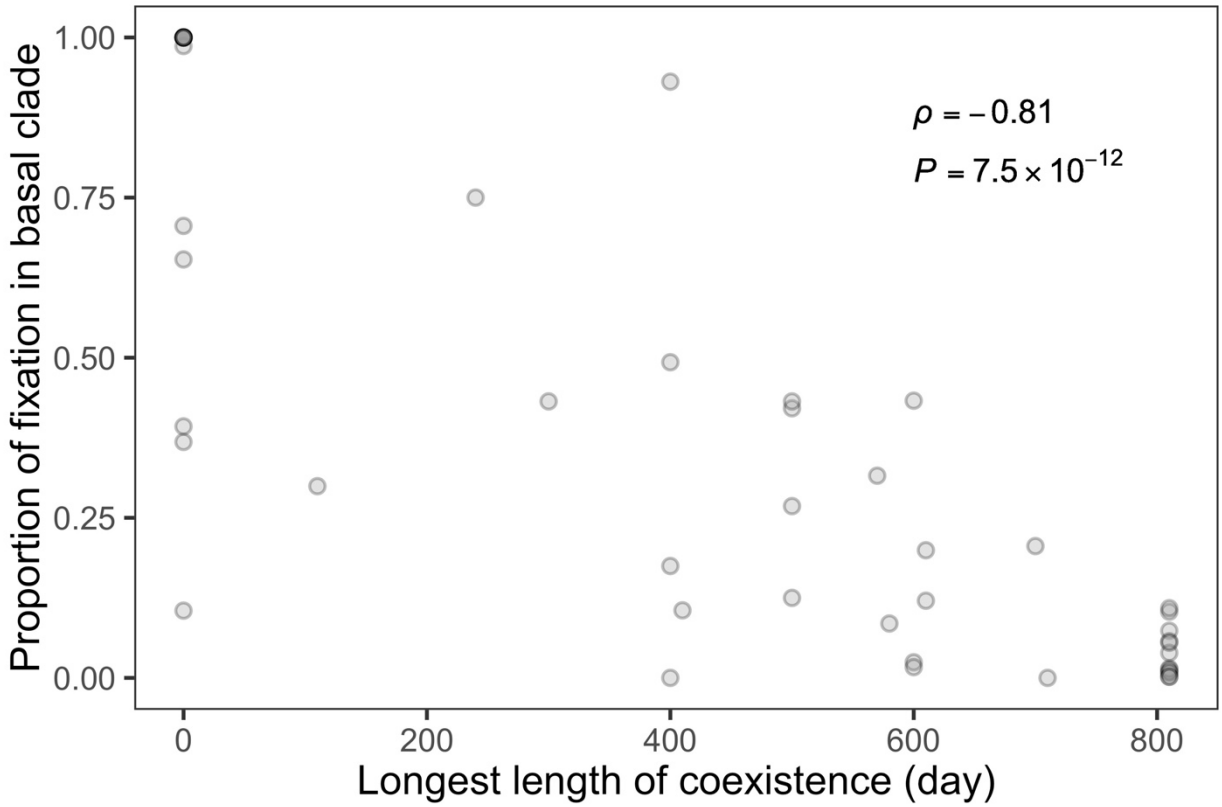


Extended Data Figure 1. Temporal dynamics in various measurements of genomic diversity. Genomic diversity is measured by (**a**) number of SNPs, (**b**) nucleotide diversity, and (**c**) synonymous nucleotide diversity, respectively. The y-axis shows mean results among the parallel populations in a combination of one transfer period (1 day, 10 days, or 100 days) and one genetic background (MMR- for mismatch repair defected or WT for wild-type). From left to right represent the sequencing profiles sampled at day 90, 200, 300, 400, 500, 600, 700, 800, and 900 (the data of populations with transfer periods = 100days at day 500 are missing). At each time point, the circle and the bar represent the mean +/- standard error of mean. The dashed lines represent the linear regression lines of number of SNPs against guaranteed generations, and the estimated slopes (*b*) and corresponding standard errors are also shown.

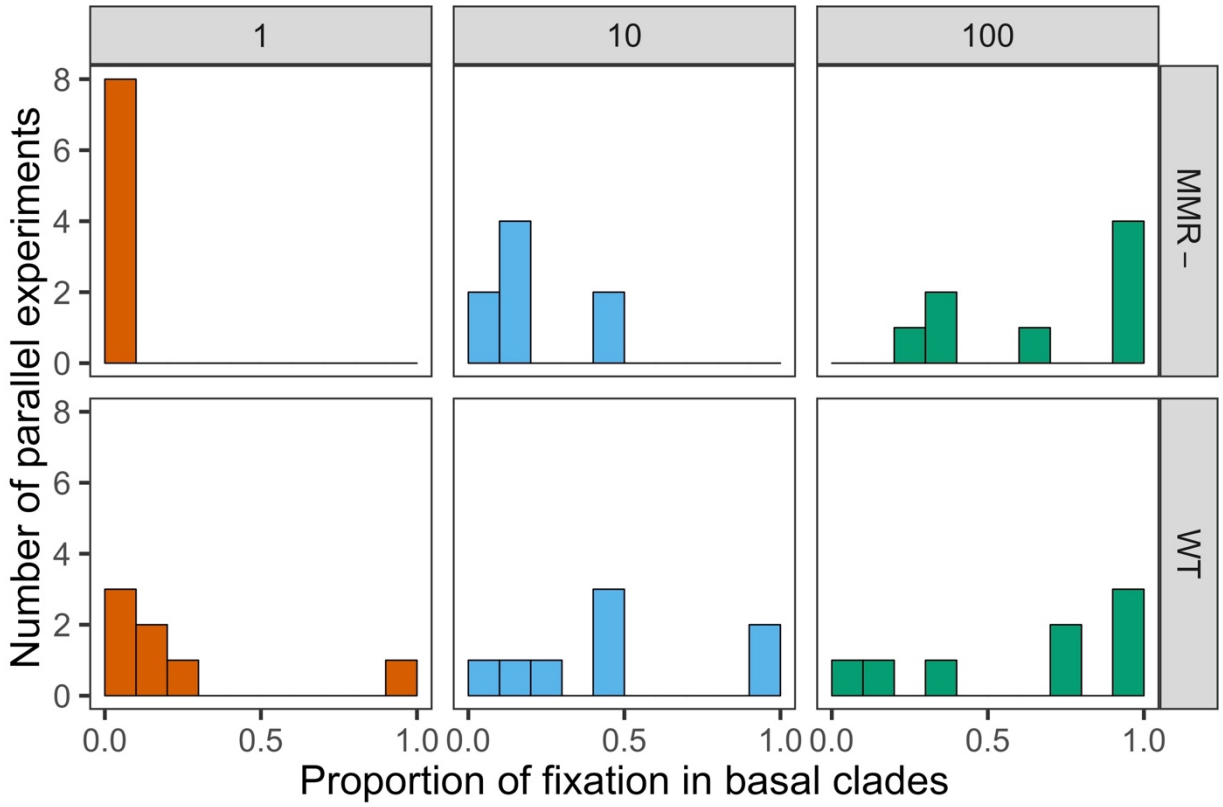


Extended Data Figure 2. No evidence for stronger positive selection in any transfer period. **(a)** Similar probabilities of synonymous, nonsynonymous, and intergenic mutations fixed in any clade were found in populations with larger transfer periods. The dot and error bar represent mean probability and its 95% confidence interval (estimated by 1.96 times of standard errors) across parallel experimental populations. The fixed probably here is conditioning on the detectability of mutations. **(b-c)** Populations with longer transfer periods do not show significant higher fixation rate of nonsynonymous mutations and intergenic mutations compared to the fixation rate of synonymous mutations. If positive selection is strong, one would expect to see that mutations with a greater potential to change protein functionality or levels (e.g. nonsynonymous mutations and intergenic mutations) are more likely to be fixed within a clade than mutations with lesser potential (e.g. synonymous mutations). Therefore, higher ratios of the number of nonsynonymous mutations fixed in any clade (F_N) to the number of synonymous mutations fixed in any clade (F_S), or higher ratios of the number of integenic mutations fixed in any clade (F_I) to F_S, is expected. **(b)** A cross is calculated by the summed F_N and F_S of all parallel experimental populations with a specific transfer period and genetic background. A dot is calculated by the F_N and F_S in each parallel experimental population. The dots with $F_S = 0$ are

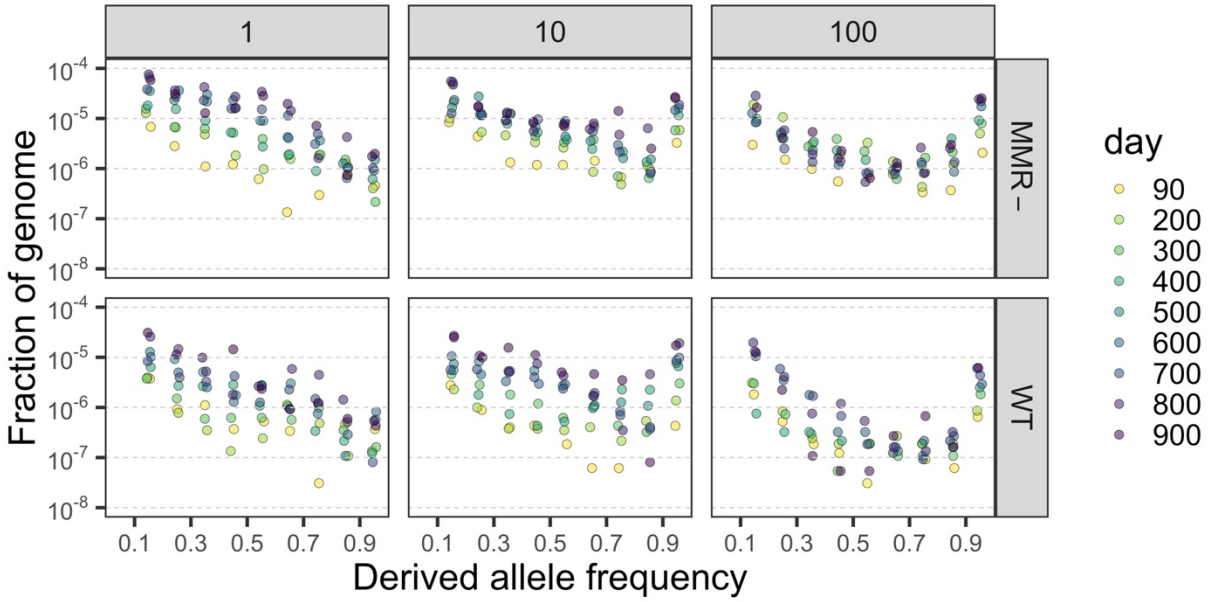
632 not shown. The dashed lines show the ratio of the number of nonsynonymous mutations to the
633 number of synonymous mutations detected in mutation accumulation experiments. **(c)** A cross is
634 calculated by the summed F_I and F_S of all parallel experimental populations with a specific
635 transfer period and genetic background. A dot is calculated by the F_I and F_S in each parallel
636 experimental population. The dots with $F_S = 0$ are not shown. The dashed lines show the ratio of
637 the number of intergenic mutations to the number of synonymous mutations detected in mutation
638 accumulation experiments.



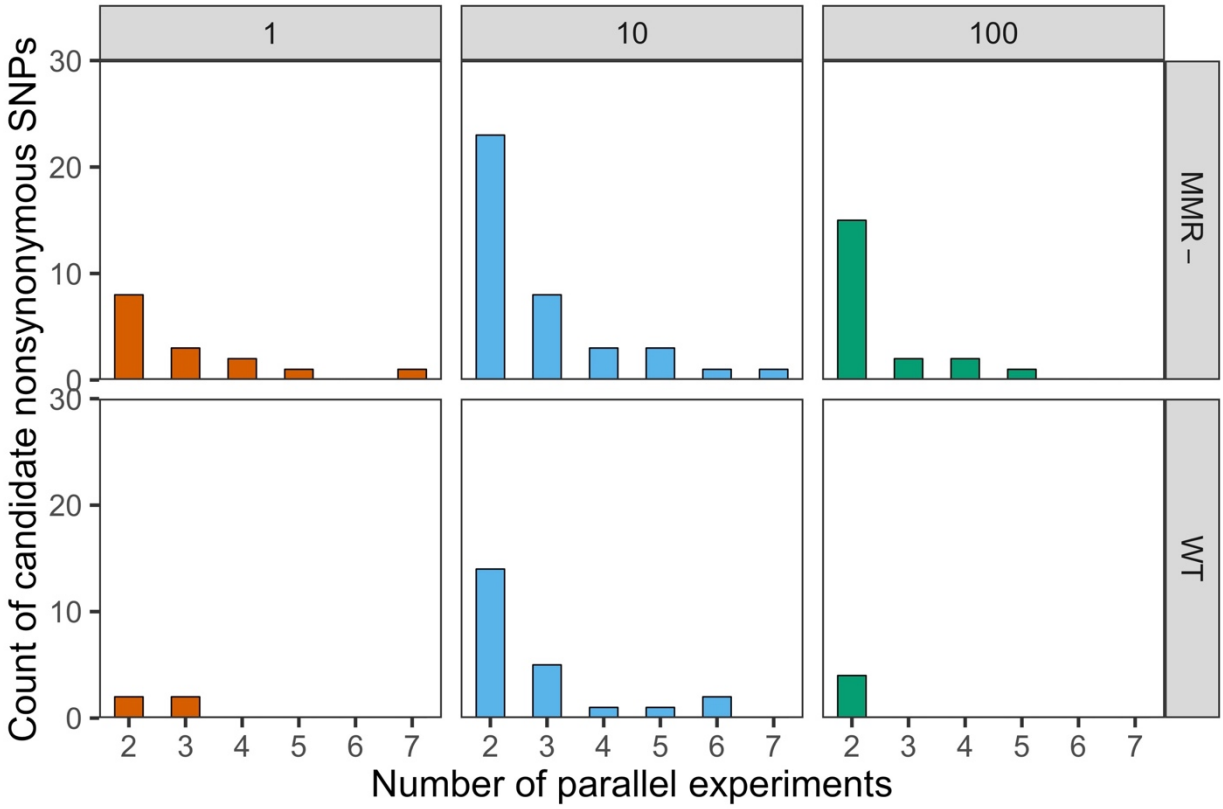
Extended Data Figure 3. Negative correlation between the proportion of fixation occurred in basal clade and the longest length of coexistence across all experimental populations. Each dot is an experimental population. ρ : Spearman's rank correlation coefficient. P: p-value of Spearman's rank correlation.



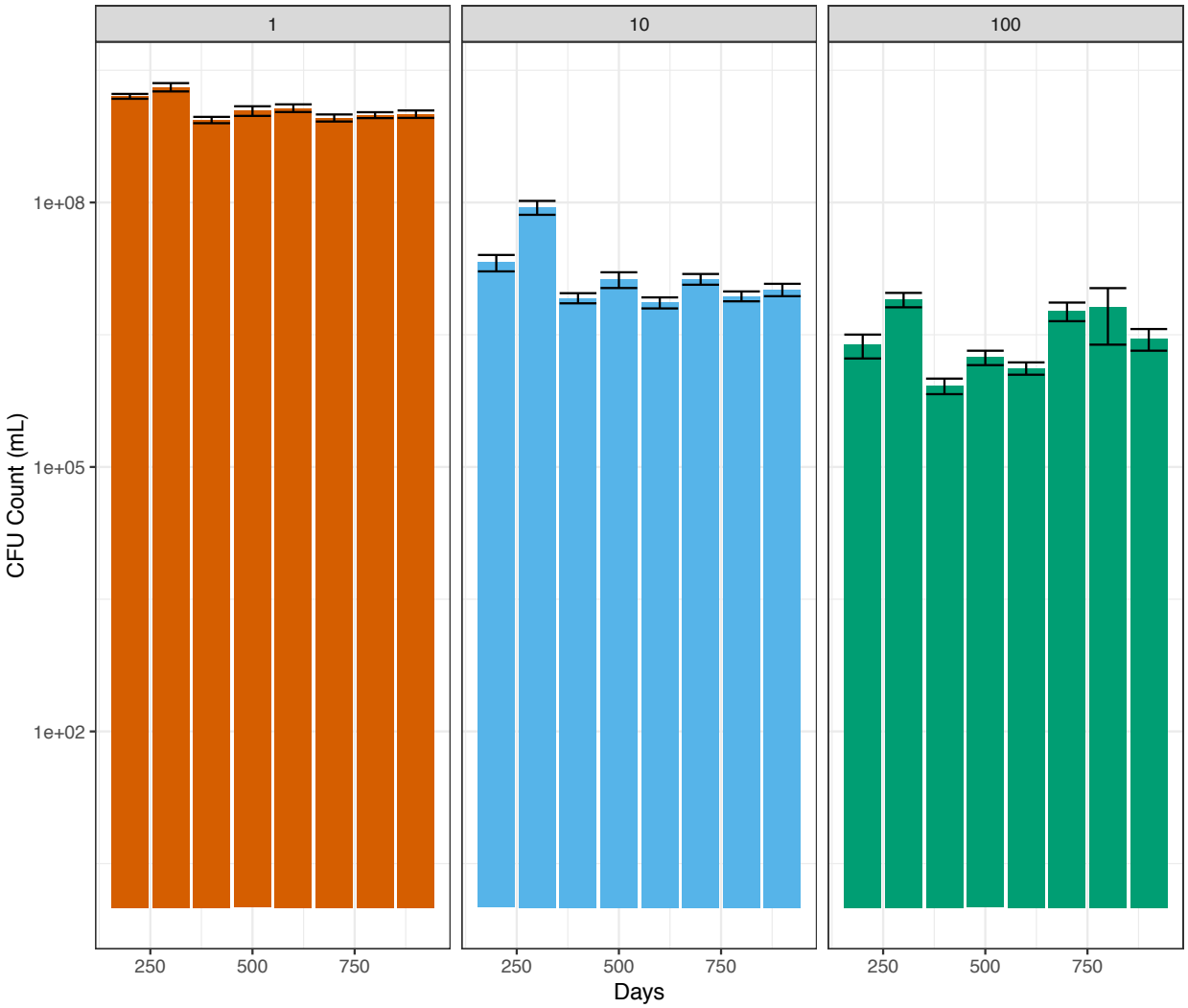
Extended Data Figure 4. The distribution of proportion of fixation occurred in basal clades for each transfer period/genetic background combination. We found TP significantly affects the proportion of fixation occurred in basal clades ($P = 2.5 \times 10^{-4}$, Scheirer-Ray-Hare test). The difference is significant between daily and 100-day populations ($P = 5.5 \times 10^{-5}$, Mann-Whitney *U* test), 10-day and 100-day populations ($P = 0.016$, Mann-Whitney *U* test), and daily and 10-day populations ($P = 1.6 \times 10^{-3}$, Mann-Whitney *U* test).



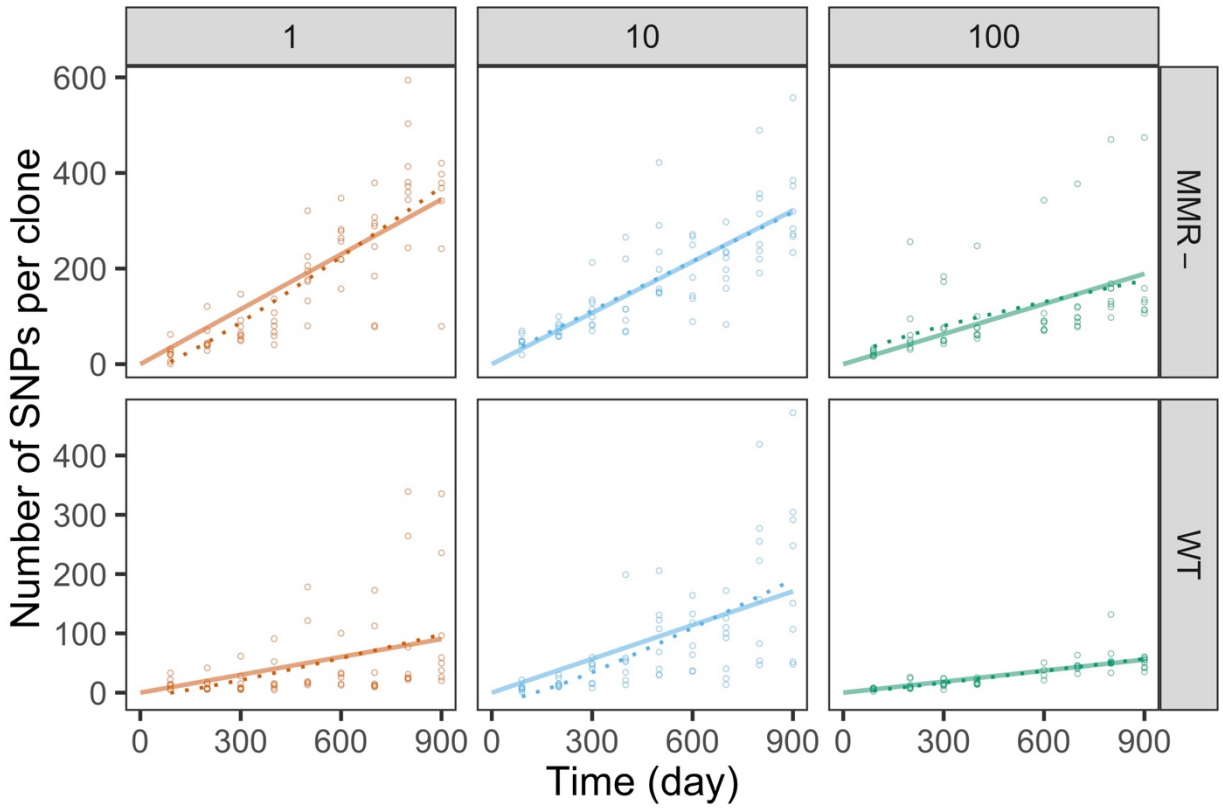
Extended Data Figure 5. Temporal distribution of mean derive allele frequencies (DAF) in evolved populations. Each dot shows a mean fraction of genome showing a range of DAF with bin size 0.1 in different time point of experimental evolution. The different time points of experimental evolution are shown by colors.



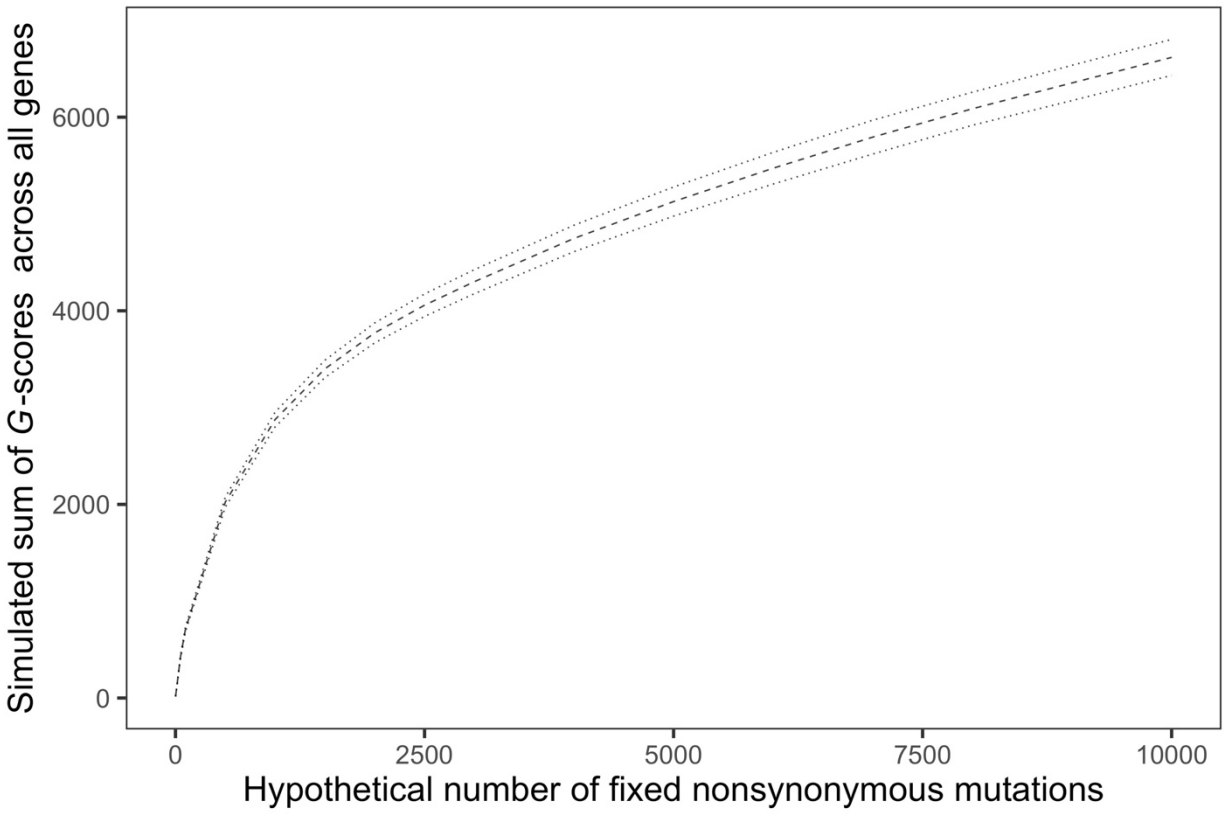
Extended Data Figure 6. Distribution of number of parallel experiments where the same fixed nonsynonymous SNPs are identified. All six combinations of one transfer periods (1 day, 10 days, or 100days) and in one genetic background (MMR- for mismatch repair defected or WT for wild-type) are shown. Because of a very low probability that any one nucleotide is mutated at least twice ($< 10^{-4}$ by simulation with a random sample of 10^4 mutations), our observed nucleotide-level parallelism again suggests positive selection in action.



674
675 **Extended Data Figure 7. Density measurement of experimental populations during**
676 **experimental evolution.** Pre-transfer population density of 1- (orange), 10- (blue) and 100-day
677 (green) populations as determined by CFU counts. Bars represent mean population density of
678 WT and MMR- populations after 200, 300, 400, 500, 600, 700, 800, and 900 days of evolution
679 and error bars represent ±s.e.m.
680



Extended Data Figure 8. Linear and nonlinear regression between total frequency of SNPs and days. Each panel shows the results in one of transfer periods (TP: 1, 10, or 100 days) and in MMR- (mismatch repair defected) or WT (wild-type) genetic background. At each time point, a circle represents the number of SNPs per clone in a parallel population. Sequencing profiles were sampled at days 90, 200, 300, 400, 500, 600, 700, 800, and 900. Note that the data of populations with 100-day transfer periods are missing at day 500. The solid line represents the result of linear regression ($y \sim 0 + x$); The dotted line represents the result of nonlinear regression ($y \sim 0 + x + x$^0.5). Model comparison were performed by nested ANOVA: $P = 0.028$, 0.69, and 0.23 for TP = 1, 10, and 100, respectively in MMR- background; $P = 0.40$, 0.063, and 0.40 for TP = 1, 10, and 100, respectively in WT background. The nested ANOVA test was performed by the function "anova" in R.



693
694
695 **Extended Data Figure 9. Null expectation of *G*-scores vary with total number of mutations**
696 **considered.** The dashed line and dotted lines summarize the mean and 95% range across 1000
697 simulations of hypothetical number of fixed nonsynonymous mutations = 1, 50, 100, 500, 1000,
698 1500, 2000, 2500, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000.