

1 Evolutionary dynamics of *de novo* mutations and mutant lineages arising in a simple, constant
2 environment

3

4 Margie Kinnersley^{1,¶}, Katja Schwartz^{3,¶}, Jacob Boswell², Dong-Dong Yang², Gavin Sherlock^{3,*}
5 Frank Rosenzweig^{1,2,*}

6

7 ¹Division of Biological Sciences, The University of Montana, Missoula, MT 59812

8 ²School of Biological Sciences, Georgia Institute of Technology, Atlanta, GA 30332

9 ³Department of Genetics, Stanford University School of Medicine, Stanford, CA 94305-5120

10

11

12 ¶These authors contributed equally to this work.

13

14 *Co-corresponding authors: frank.rosenzweig@biology.gatech.edu and gsherloc@stanford.edu

15

16 **Keywords:** *E. coli*, adaptive evolution, chemostat, metagenomics, population sequencing, clone
17 sequencing, clonal interference, parallelism

18

19 **Running Title:** Parallelism and clonal interference in evolving bacterial populations

20

21 **Abstract**

22 A large, asexual population founded by a single clone evolves into a population teeming with
23 many, whether or not its environment is structured, and whether or not resource levels are
24 constant or fluctuating. The maintenance of genetic complexity in such populations has been
25 attributed to balancing selection, or to either clonal interference or clonal reinforcement, arising
26 from antagonistic or synergistic interactions, respectively. To distinguish among these
27 possibilities, to identify targets of selection and establish when and how often they are hit, as
28 well as to gain insight into how *de novo* mutations interact, we carried out 300-500 generation
29 glucose-limited chemostat experiments founded by an *E. coli* mutator. To discover all *de novo*
30 mutations reaching $\geq 1\%$ frequency, we performed whole-genome, whole-population sequencing
31 at $\sim 1000X$ -coverage every 50 generations. To establish linkage relationships among these
32 mutations and depict the dynamics of evolving lineages we sequenced the genomes of 96 clones
33 from each population when allelic diversity was greatest. Operon-specific mutations that enhance
34 glucose uptake arose to high frequency first, followed by global regulatory mutations. Late-
35 arising mutations were related to energy conservation as well as to mitigating pleiotropic effects
36 wrought by earlier regulatory changes. We discovered extensive polymorphism at relatively few
37 loci, with identical mutations arising independently in different lineages, both between and
38 within replicate populations. Out of more than 3,000 SNPs detected in nearly 1,800 genes or
39 intergenic regions, only 17 reached a frequency $\geq 98\%$, indicating that the evolutionary
40 dynamics of adaptive lineages was dominated by clonal interference. Finally, our data show that
41 even when mutational input is increased by an ancestral defect in DNA repair, the spectrum of
42 beneficial mutations that reach high frequency in a simple, constant resource-limited

- 43 environment is narrow, resulting in extreme parallelism where many adaptive mutations arise but
44 few ever go to fixation.

45 **Author Summary**

46 Microbial evolution experiments open a window on the tempo and dynamics of evolutionary
47 change in asexual populations. High-throughput sequencing can be used to catalog *de novo*
48 mutations, determine in which lineages they arise, and assess allelic interactions by tracking the
49 fate of those lineages. This *adaptive genetics* approach makes it possible to discover whether
50 clonal interactions are antagonistic or synergistic, and complements genetic screens of induced
51 deleterious/loss-of-function mutants. We carried out glucose-limited chemostat experiments
52 founded by an *E. coli* mutator and performed whole-genome, whole-population sequencing on
53 300-500 generation evolutions, cataloging 3,346 *de novo* mutations that reached $\geq 1\%$ frequency.
54 Mutations enhancing glucose uptake rose to high frequency first, followed by global regulatory
55 changes that modulate growth rate and limiting resource assimilation, then by mutations that
56 favor energy conservation or mitigate pleiotropic effects of earlier regulatory changes. We
57 discovered that a few loci were highly polymorphic, with identical mutations arising
58 independently in different lineages, both between and within replicate populations. Thus, when
59 mutational input is increased by an ancestral defect in DNA repair, the spectrum of beneficial
60 mutations that arises under constant resource-limitation is narrow, resulting in extreme
61 parallelism where many adaptive mutations arise but few ever become fixed.

62

63

64 **Introduction**

65 Evolution experiments using microbes have enlarged our understanding of the tempo and
66 dynamics of evolutionary change, as well as how selection, drift and historical contingency
67 influence evolutionary trajectories. Combined with high throughput sequencing, experimental
68 microbial evolution (EME) can now be used to identify substantial numbers of *de novo*
69 beneficial mutations in laboratory populations, to determine in which lineages they arise and the
70 fate of those lineages, and to evaluate the sign and strength of possible epistatic interactions [1-
71 3]. This approach, adaptive genetics, based on analyzing cohorts of spontaneous beneficial
72 mutations to determine how their frequencies fluctuate over time, constitutes a mode of inquiry
73 that complements traditional genetic screening of induced deleterious/loss-of-function mutants
74 (e.g., [4] and [5] among others). Adaptive genetics also expands the possibilities for discovering
75 constraints on protein structure and function and for discerning the architecture and malleability
76 of networks that regulate nutrient-sensing and cell division.

77 Microbial populations were once thought to evolve by periodic selection as a succession
78 of adaptive clones, each fitter than its antecedent, replacing one another over time [6-9]. This
79 model was consistent with Muller and Haldane's view of how beneficial mutations spread in
80 large asexual populations [10-12] under conditions governed by competitive exclusion [13].
81 Today we know that large, initially clonal populations rapidly accumulate and retain genetic
82 variation, much of which is beneficial [14-18]. In fact, the amount of adaptive genetic variation
83 observed in EME populations can be enormous, owing large population sizes with a continuous
84 input of beneficial mutations and the subsequent competition among new adaptive lineages,
85 which gives rise to clonal interference [14,16,19,20].

86 Clonal interference can occur within a larger framework of stable subpopulation structure
87 [21] when microbial lineages come under balancing selection [22-25] or specialize to exploit
88 niches created either by the culture conditions [23,26,27], or by the organisms themselves [28-
89 30]. In a simple constant environment like a chemostat the persistence of subpopulations likely
90 depends on founder genotype, the emergence of specific key mutations, and availability of the
91 limiting nutrient [31]. Ferenci and colleagues never observed stable subpopulation structure in
92 glucose-limited evolutions originating from *E. coli* K12 strain BW2952 [32], whereas Adams
93 and colleagues, using a different strain, often did [30,33]. Unlike BW2952, the K12-derived
94 ancestor used by Adams, JA122 [30] harbors a supE44 *glnX* tRNA nonsense suppressor as well
95 as nonsense mutations in housekeeping and stationary-phase transcription factors, RpoD and
96 RpoS respectively, and mismatch repair enzyme MutY. The JA122 ancestor's defect in DNA
97 repair increases mutational load on its descendants [28,34], while the nonsense suppressor
98 mitigates the effect of mutations that create premature stop codons. Such a suppressor would
99 likely make the blunt instrument of *de novo* nonsense mutations a less effective agent of adaptive
100 change, possibly resulting in a more nuanced spectrum of beneficial mutations than would
101 otherwise occur among mutators.

102 To understand the impact that a mutator/suppressor founder has on the spectrum and fate
103 of new beneficial mutations, and on the dynamics of population structure, we repeated Adams *et*
104 *al.* classic evolution experiments using the same ancestral strain and culture conditions [30]. We
105 monitored, at 50-generation intervals, the incidence of mutations that reached at least 1%
106 frequency over the course of 300-500 generations, identifying mutations that were either
107 transiently beneficial or hitch-hiking with mutations that were. To determine which mutations
108 co-occurred within a given lineage we sequenced 96 clones from each population at the time-

109 point where we observed greatest allelic diversity. We uncovered no evidence for stable sub-
110 population structure, but instead saw pervasive clonal interference, with only 17 out of 3,346
111 mutations going to near fixation across replicate experiments. The temporal order in which
112 certain mutations rose to high frequency was predictable, reflecting a high degree of parallelism
113 both within and between replicates. In general, mutations that enhanced glucose assimilation
114 arose early, followed by mutations in global regulators and mutations that either increase
115 efficiency of limiting resource utilization or mitigate the deleterious effects of certain earlier
116 mutations. Altogether, our results show that even in bacterial populations founded by an ancestor
117 having a high mutation rate and the capacity to tolerate many *de novo* mutations, the spectrum of
118 genomic changes that rise to appreciable frequency and the adaptive outcome of replicate
119 evolutions are limited when those populations evolve in a simple constant environment.

120

121 **Results**

122 **Experimental design.** Evolution experiments were carried out in triplicate under continuous
123 nutrient limitation using Davis Minimal Medium [30], with glucose (0.0125% w/v) as the sole
124 source of carbon for energy and growth. Chemostats (300 mL working volume) were run under
125 aerobic conditions for 300-500 generations at constant temperature (30°C) and at constant
126 dilution rate ($D=0.2 \text{ hr}^{-1}$). Under these conditions, population density reaches $\sim 10^8 \text{ cells mL}^{-1}$ at
127 steady state. The *E. coli* strain used to initiate these experiments, JA122, is distinguished from *E.*
128 *coli* K12 by alleles likely to influence the spectrum of mutations arising during adaptive
129 evolution (**Table S1**; [28]). Among these is a nonsense mutation in MutY (Leu299*) that results
130 in a 10-fold greater mutation rate and GC→TA transversion bias [28], nonsense mutations in the
131 genes that encode stationary phase sigma factor RpoS (Gln33*) [35] and ‘housekeeping’ sigma

132 factor RpoD (Glu26*), as well as a suppressor mutation in the *glnX* tRNA known to suppress
133 amber, ochre and opal mutations (**Table S1**) [36].

134 To identify the mutations that arose during the evolutions, we performed whole genome,
135 whole population sequencing every 50 generations on each of the three chemostat populations.
136 We generated approximately 50 million 2x100bp paired end reads per sample, yielding coverage
137 of up to ~1000x for each time point (inserts were selected to be short enough such that forward
138 and reverse reads overlapped, which while reducing coverage, increases quality; see Methods).
139 Based on this level of coverage, we were able to identify mutations that rose to an allele
140 frequency of ~1% or greater. Given an effective population size of $>10^{10}$ and 300-500
141 generations of selection it is highly improbable that any allele could reach such a frequency by
142 drift alone [23]. We can therefore assume that every mutation recovered was either under
143 positive selection or hitch-hiking along with one that was.

144 ***Population sequencing shows general patterns of mutation that are consistent across***
145 ***independent evolutions.*** Across all samples, 3,326 SNPs were detected in 2,083 unique genes or
146 **intergenic regions** (**File S1**). The overwhelming majority (97.5%) of these SNPs were GC→TA
147 transversions, as expected given the ancestral strain's defect in the mismatch repair protein
148 MutY, which encodes adenine glycosylase [37]. Consistent with the protein coding density of *E.*
149 *coli* (87.8%) [38], 85% (2,854) of SNPs occurred in coding regions. On average, 69.2% of these
150 created a missense mutation, 23.4% resulted in a synonymous mutation and 7.4% caused a
151 nonsense mutation (**Fig. 1**). Relative to proportions observed in mutation accumulation
152 experiments carried out using wild-type *E. coli* [39], we observed more nonsynonymous and
153 nonsense mutations. Small deletions were rarely detected (one single-nucleotide deletion in each
154 of vessel 1 and vessel 2, and none detected in vessel 3), but we observed a single large ~150kb

155 duplication in vessel 2. The overall number of mutations in each population increased linearly
156 over time and at approximately the same rate across replicates (**Fig. 1**), as would be expected
157 with a mutator phenotype.

158 ***Comparison of population level mutations reveals clonal interference and widespread***
159 ***parallelism.*** Despite the large number of SNPs detected, only 17 alleles arose above a frequency
160 of 98% across replicate evolutions ranging from 300-500 generations. Moreover, the maximum
161 frequency of most alleles never exceeded 10% (**Fig. 2A**), and the vast majority of alleles were
162 present at lower frequency in the final time-point than they were at their maximum (**Fig. 2B**).
163 Together, the foregoing observations suggest that in each evolution experiment population
164 dynamics was largely driven by clonal interference [40]. A small number of loci were recurrently
165 mutated above what would be expected by chance, indicating that variants at these loci were
166 likely beneficial (**Table 1**, **Table S2**). For example, a total of 212 mutations arose in the 10 most
167 significantly mutated genes in the population sequencing data, with each gene receiving at least
168 five mutations (**Table 1**). Moreover, 30 and 14 distinct allelic variants were discovered in just
169 two: the genes encoding the DNA binding repressor GalS and the RNA-binding protein Hfq,
170 respectively (**Table S3**). High-resolution population sequencing also revealed that 13 SNPs not
171 present at the start of the experiment reached at least 1% frequency in all three vessels at various
172 time-points, while 52 SNPs recurred in two out of three chemostats (**Table S4**). Thus, our data
173 also provide compelling evidence for substantial parallel evolution at the genic level.

174 ***Clonal sequencing further clarifies lineage relationships and parallelism*** To determine
175 linkage relationships between the novel alleles, we sequenced 96 individual clones from each
176 vessel. In each case, the 96 clones were isolated at random from the time-point at which we
177 detected the greatest number of mutant alleles at $\geq 5\%$ frequency. To assess whether the

178 frequency estimates from population sequencing were reasonable, and whether the isolated
179 clones constituted a reasonable subsample, we compared frequencies of mutations identified in
180 both datasets at the corresponding time-point and found that they correlate well (**Fig. 3**).

181 For each set of 96 clones, we constructed a phylogeny to represent their putative
182 evolutionary relationships (**Fig. 4**). Inspection of the mutations and trees from each vessel (i.e.
183 each independent evolution) revealed several instances in which exactly the same mutation arose
184 not only in different vessels, but often more than once in the same vessel on distinct branches of
185 a given tree. In the most extreme case, 6 of the 11 *hfq* alleles detected via clone sequencing were
186 identified in clones from different vessels, indicating independent parallel origins (**Fig. 4, File**
187 **S2**). Furthermore, 7 of the 11 appear to have arisen more than once within the same vessel.

188 ***Clonal dynamics are shaped by relationships among de novo alleles, hard and soft selective***
189 ***sweeps, and absence of periodic selection*** Combining population allele frequency data with
190 linkage information derived from clonal sequencing makes it possible to depict lineage dynamics
191 using Muller diagrams (**Fig. 5, Files S3-S5**). In general, we observe early, hard sweeps of highly
192 beneficial mutations related to limiting nutrient influx, followed by soft sweeps [41-43] and
193 multiple-origin soft sweeps that may fine-tune glucose uptake or utilization later in the
194 experiment when diversity was higher [44-46]. Hard sweeps consistently involved mutations in
195 regulators (*gals* in chemostat 1, transcriptional terminator *rho* in chemostats 1 and 3) or
196 regulatory regions (upstream of *dnaG* in chemostat 1, upstream of *mglB* in chemostats 1, 2 and 3
197 – See supplementary Files 2, 3 and 4 for detailed dynamics), while soft sweeps were comprised
198 of both regulatory and operon-specific mutations (e.g. *hfq* and *opgH* in chemostats 1, 2 and 3,
199 upstream of *adhE* in chemostat 1, *pgi* in chemostat 3) (**Fig. 5, S1, Files S3-S5**) [42,47]. Here, we
200 note that multiple-origin soft sweeps may be especially prevalent in our experiments due to the

201 ancestral mutator allele at *mutY*, as the likelihood of concurrent identical mutations in the same
202 gene should increase with mutation rate.

203 With regard to periodic selection, rather than favorable alleles arising within a set of lineages
204 that successively replace one another over time, we observe groups or cohorts of mutations co-
205 evolving, with widespread clonal interference among lineages that carry different beneficial
206 mutations [48]. For example, in chemostat 1, a spreading lineage with a cohort of mutations
207 upstream of *mglB/lptA/opgH* (pink) is checked by the emergence of lineages carrying mutations
208 in *hfq* (green) (**File S3**). All of these phenomena – hard and soft sweeps, cohorts of mutations
209 that increase or decrease in frequency together, and clonal interference – have been observed in
210 yeast [14,19,49] and *E. coli* [23] populations evolving in the laboratory, as well as in
211 *Pseudomonas aeruginosa* evolving in the cystic fibrosis lung [50].

212 **Early sweeps are related to influx of the limiting nutrient glucose.** For specific growth rates
213 between $\sim \mu = 0.1 \text{ hr}^{-1}$ and $\mu = 0.9 \text{ hr}^{-1}$, glucose is most efficiently transported using a
214 combination of the maltoporin LamB and the galactose transporter MglBAC, and glucose
215 limitation frequently selects for mutations that increase expression of these proteins [45,51-59].
216 As expected, 7 of the top 10 frequently mutated genes/gene regions we observed (*galS*, upstream
217 *mglB*, *malT*, *malK*, *hfq*, *rho* and upstream *dnaG*) play a role in transcriptional regulation of *lamB*
218 or *mglBAC*, either directly or through their interactions with global regulators (**Table 1, Fig. 6**).

219 **MUTATIONS IN *GALS* AND UPSTREAM OF *MGLB*.** Thirty different alleles of *galS* (encoding GalS,
220 a negative regulator of *mglBAC* transcription) were detected over the course of our experiments.
221 These spanned the length of the gene, and the majority caused missense amino acid changes
222 likely to disrupt *mglBAC* transcriptional repression and augment glucose flux across the inner
223 membrane (**Fig. 6**) [60]. Despite the large number of alleles we observed, few persisted beyond

224 generation 50 or attained a final frequency greater than 5%, demonstrating high clonal diversity
225 early in the experiment. In chemostats 2 and 3, no clear “winner” *galS* genotype emerged, though
226 in chemostat 1, a GalS allele (Arg146Leu) swept to near fixation (89.6% of the population at
227 generation 50).

228 Early-arising *GalS* mutant genotypes were rapidly displaced by clones carrying highly-
229 beneficial mutations in the *mgl* operator sequence upstream of *mglB* (**Fig. 5**). This sequence of
230 events, like the mutations themselves, has been observed elsewhere [45,61]. The most successful
231 mutation upstream of *mglB* (bp 2,238,647 C→A) occurred early in every vessel, and in every
232 case increased in frequency to over 90% of the population (**Table S5, File S3-S5**). Notably, this
233 same mutation was observed in chemostat-grown *E. coli* by Notley-McRobb et al. (1999) as well
234 as in the experimental population described by Helling et al. (1987), where it was found to be the
235 only SNP shared by all members of a cross-feeding consortium [34,54].

236 The dynamics of *galS* replacement illustrates the effect that clonal interference can have
237 on the fate of different alleles. In chemostat 1, clones carrying GalS Arg146Leu rapidly dropped
238 in frequency when lineages emerged with a mutation upstream of *mglB* (position 2,238,647), but
239 were not completely displaced until generation 400 and even enjoyed brief periods of expansion.
240 In chemostat 2, clones with the same mutation upstream of *mglB* were present by generation 50,
241 but did not surpass a 90% threshold for another 250 generations due to competition from 22
242 different *galS* lineages and a lineage carrying a different upstream of *mglB* allele (2,238,648
243 G→T) (**Fig. 5, S1B, File S4**). By contrast, in chemostat 3, a lineage with the upstream *mglB*
244 mutation (2,238,647) experienced little competition and was almost fixed by generation 150
245 (**Table S5**).

246 Over the remainder of the experiment only three other mutations upstream of *mglB*
247 mutations reached the threshold for detection: two were within 2 base-pairs of the first mutation
248 and did not rise to high frequency, while the third (chemostat 1, 2,238,630 C→A) located in the
249 CRP activator binding site, co-occurred with 2,238,647 C→A and increased to ca. 80%
250 frequency by generation 500 (**Figs. 5, S1A, File S3**). This dynamic suggests additional mutations
251 that affect GalS repressor binding are not of great benefit after the preferred allele has swept the
252 population, whereas mutations that modulate the activity of other regulators (i.e. CRP) can act
253 synergistically.

254 THE DYNAMICS OF LAMB REGULATION. LamB glycoporin overexpression is a hallmark
255 feature of *E. coli* populations adapted to glucose-limited chemostat growth [34,45,53,55,56,62].
256 Previous experiments have shown that under glucose limitation, overexpression of LamB can be
257 the result of any one of the following: constitutive activation of transcriptional regulator MalT,
258 disruption of the MalT inhibitor MalK, mutation of the RNA chaperone Hfq, alteration of sigma
259 factor dynamics (σ^S/σ^D ratio), or mutation of the *malT* repressor Mlc, (**Fig. 6**) [34,45,53-
260 56,62,63].

261 Across the three replicate evolutions, we observed 19 unique *malT* alleles and 14 unique
262 *malK* alleles (**Figs. 6, S2, Table S3**). Over half of the mutations in *malT* (10 out of 19) are
263 known either to cause MalT to become constitutively active, or to occur in amino acids involved
264 in MalT/MalK interaction [45,64,65]. A single MalK mutation (Ala296Asp) rose to high
265 frequency early (94% by generation 100) in chemostat 1 (**Fig. S2, Table S5**). This SNP is in a
266 regulatory domain likely to be the site of MalK/MalT interaction [66]. Alteration of a
267 neighboring residue (Asp297) has been previously shown to allow unregulated transcription of
268 the *mal* operon [66].

269 *malT/malK* allele dynamics differ among experimental populations. As mentioned above
270 and shown in **Fig. S2**, MalK Ala296Asp sweeps early in chemostat 1, whereas in chemostat 3,
271 early MalK mutations (blue) are displaced by later mutations in MalT (green). In chemostat 2,
272 the picture is quite different: clones with either *malK* or *malT* mutations co-exist through all 500
273 generations. The reason for this contrast in allele frequency dynamics cannot be attributed to
274 emergence of a single “most fit” allele, as the majority types from chemostats 1 and 3 arose
275 independently in chemostat 2, but did not sweep. Despite the importance of MalT and MalK as
276 high-value targets for selection during adaptation to glucose limitation, other advantageous
277 mutations (upstream *mglB*, *rho* and *hfq*, discussed below) may have ultimately carried “winning”
278 *mal* alleles in chemostats 1 and 3 to higher frequency, purging allelic diversity at this locus.
279 Interestingly, although we observed 30 *malT* and 22 *malK* mutations in the population
280 sequencing data (**Table 1**), in only 5 out of the 288 sequenced clones do mutant alleles of these
281 two genes co-occur, suggesting that there may be no additional advantage or even some
282 disadvantage to having both. In the Helling et al. evolution experiments [30], which were
283 founded by the same ancestor used here, secondary resource specialists share a mutation in
284 MalT, whereas the primary resource (glucose) specialist that feeds those clones carries a
285 mutation in MalK [34].

286 SELECTION OF MUTATIONS IN RNA CHAPERONE HFQ THAT AFFECT TRANSLATION OF LAMB
287 AND STATIONARY PHASE TRANSCRIPTION FACTOR RPOS. Hfq is a global regulatory protein that
288 facilitates translation and/or RNA degradation by mediating ncRNA-mRNA interactions. It
289 participates in a diverse range of cellular processes including nutrient uptake, motility and
290 metabolism [67]. *hfq* mutations identified in other glucose-limited evolution experiments exhibit
291 pleiotropic physiological effects: they appear to increase translation of LamB glycoporin, reduce

292 levels of stationary phase transcription factor RpoS, inhibit cellular aggregation, and enhance
293 glucose transport via PtsG [63,68].

294 *Hfq* is one of the most frequently mutated genes observed in our experiments: 24 *hfq*
295 mutations, resulting in 14 distinct *hfq* alleles, were detected via population sequencing; by the
296 end of our evolutions >50% of each population carried a mutation in *hfq* (**Table S3**). (**Table 1**).
297 Two of these alleles arose independently in all three vessels (same nucleotide position, same
298 SNP), and six additional alleles were observed in two of three vessels (**Table S4**). The frequency
299 of and parallelism exhibited in *hfq* mutations is particularly curious in the context of experiments
300 by Maharjan et al. in which *hfq* mutations arise, but are at low frequency and subject to negative
301 frequency-dependent selection and epistatic interaction with mutations in *rpoS* [32,52,62,63].

302 The dynamics of *hfq* mutations are variable across evolutions and may depend on which
303 other beneficial alleles are present in the same lineage or in the same population. In chemostats 1
304 and 2, a large number of *hfq* alleles (10 in chemostat 1 and 11 in chemostat 2) appear after
305 generation 250 and are preceded by mutations in *malK* or *malE* and the *opg* operon. The most
306 successful Hfq allele in chemostat 1, Val62Phe, occurred in a sweeping lineage with a secondary
307 mutation upstream of *mglB* (discussed above) and may have been carried along by association.
308 In chemostat 3, a single *hfq* mutation arises early (Ser60Tyr, present by generation 100), sweeps
309 to near fixation alongside MalT Met311Ile and is closely followed by mutations in *opgH* (**Fig.**
310 **S1, File S5**).

311 RECURRENT MUTATIONS OCCUR IN RHO. Early-arising mutations in the *rho* termination
312 factor are a conspicuous feature of chemostats 1 and 3 (**Figs. 5, S1**). Rho is required for
313 transcriptional termination of up to 50% of cellular mRNAs [69,70] and can participate in gene
314 regulation via intragenic terminators [71]. Mutagenesis and ChIP-chip analyses have identified

315 Rho-dependent terminators within multiple genes relevant to glucose limitation, specifically
316 *lamB*, *mglA*, and *mglC* and downstream of *malT* and *mglC* [71,72]. In fact, it has long been
317 known that defective LamB expression in MalT activator mutants can be restored via
318 compensatory mutations in *rho* [73]. In chemostats 1 and 3, Rho mutations fix or nearly fix early
319 and do so in concert with mutations in MalK (chemostat 1 Ala296Asp) and mutations upstream
320 of *mglB* (chemostats 1 and 3, bp 2,238,647) (**Figs. 5, S1, Table S3, Table S5, Supplementary**
321 **Files 2 and 4**). Conversely, in chemostat 2 only three *rho* alleles were detected, none of which
322 rose in frequency to >6% of the population (**Fig. S1, Table S3**).

323 ***Mutations that impact energy conservation, membrane biogenesis and cell adhesion are***
324 ***late arising targets of selection*** PHOSPHOGLUCOSEISOMERASE (PGI) is an abundantly expressed
325 central metabolic enzyme responsible for converting glucose-6-phosphate into fructose-6-
326 phosphate. Knockdown of *pgi* mRNA alleviates catabolite repression [74], favoring increased
327 expression of CRP-regulated genes such as *lamB* and *mglBAC*. Twenty-four unique *pgi* alleles
328 were detected over the course of our three replicate evolutions. However, few rose to appreciable
329 frequency before generation 200, suggesting their benefit may be contingent on the presence of
330 other mutations or some aspect of the chemostat environment that consistently changed after this
331 time point. *Pgi* alleles were least successful in chemostat 1, which was also the only replicate in
332 which a large fraction of clones (79% by generation 500) acquired a second mutation upstream
333 of *mglB*. This observation suggests that *pgi* mutations and mutations in the CRP-binding site of
334 the *mglBAC* promoter may be functionally redundant.

335 MEMBRANE GLYCOSYLTRANSFERASE OPGH is involved in the synthesis of periplasmic
336 glucans, highly branched oligosaccharides made from β -linked glucose monomers. While we do
337 not observe *opgH* mutations earlier than generation 100, they rapidly increase in frequency once

338 they appear, usually either just before or just after *hfq* mutations (**Fig. S1, S4, File S3-S5**). Novel
339 *opgH* alleles, especially the nonsense mutations that we frequently observe, may constrain
340 glucan production and serve as a glucose conservation measure. A “moonlighting” function has
341 also recently been reported for OpgH: the glucosyltransferase interacts with the tubulin-like cell
342 division protein FtsZ to delay cell division when levels of UDP-glucose are low [75]. Thus,
343 mutations in OpgH may augment the rate of cell division, and thereby provide a fitness
344 advantage under slow-growth chemostat conditions. The only *opg* operon mutation identified
345 among strains in previous Adams et al. experiments occurred in *opgG* of the glucose scavenger,
346 CV103 (E487*) [28].

347 MUTATIONS IN RHO-INDEPENDENT TERMINATOR T₁ THAT ALLOW RUN-THROUGH
348 TRANSCRIPTION MAY TIP THE BALANCE BETWEEN COMPETING SIGMA FACTORS. Sigma factor RpoD
349 (σ^{70}) is the predominant sigma factor associated with RNA polymerase during exponential
350 growth. As cells enter stationary phase, transcription of the gene for alternate sigma factor RpoS
351 (σ^S) increases [76]. *rpoS* mutations are often selected for under continuous glucose limitation as
352 they allow continued transcription from promoters negatively regulated by σ^S but required for
353 glucose uptake and metabolism (e.g. [77,78]).

354 In chemostat 1, a mutation in the *rpsU-dnaG-rpoD* macromolecular synthesis operon
355 upstream of *dnaG* (bp 3,209,081 G→T) was present in over 90% of the population by generation
356 50 (**Table S5**). This SNP decreases the stability of the rho-independent terminator T₁ situated
357 between *rpsU* and *dnaG*, and thus may be expected to increase expression of RpoD [79] and as a
358 result operons positively controlled by σ^{70} (e.g. *mglBAC* and *malK-lamB-malM*). A T1 mutation
359 (bp 3,209,075 C→A) was also shared among Helling et al. strains, defining the lineage that gave
360 rise to three of four consortium members [30,34]. In addition, in Chemostat 2, we observed an

361 ~150kb duplication that included *rpoD* and in chemostat 3, eight clones out of 96 carried
362 intragenic suppressor mutations of the ancestral nonsense allele (*26Asp and *26Gln) in RpoD.

363 FIMBRIAL PROTEIN GENES (*FIM*) Genes associated with production/function of type 1 fimbriae,
364 particularly *fimH* (fimbrial adhesion), were an unexpected and frequent target of mutation in all
365 three chemostats (**Table 1**, **Figs. 4, S1**, **Table S3**, **Files S3-S5**). Though novel *fim* alleles were
366 transient in vessels 2 and 3, in chemostat 1 a FimH Asn54Lys variant rose to a frequency of 70%
367 by generation 150, temporarily displacing high-fitness alleles in *rho*, *malK* and upstream *mglB*
368 (**File S3**). Because *fimH* mutants demonstrated an increased capacity for biofilm formation (data
369 not presented), a recurrent issue in chemostat experiments, but did not acquire any of the
370 mutations expected to enhance glucose metabolism, *fimH* mutations were likely related to
371 chemostat persistence rather than to competition for limiting substrate.

372

373 **Discussion and Conclusions**

374 ***History matters: ancestry influences evolutionary trajectory*** The tempo and trajectory of a
375 clonal population depend on its genetic point of departure. Our departure point was a founder
376 that harbored nonsense mutations in mismatch repair (MutY, Leu299*), and in housekeeping and
377 stationary phase sigma factors (RpoD, Glu26* and RpoS, Gln33*), but also carried an
378 amber/ochre/opal nonsense tRNA suppressor. Populations originating from such a founder
379 would not only have an increased mutational load but also the capacity to tolerate those
380 mutations, in particular nonsense mutations that would otherwise result in complete loss-of-
381 function.

382 Laboratory evolution studies have borne out the idea that loss-of-function mutations can
383 be significant drivers of adaptation [20,80-82]. Metabolic network re-programming via

384 modulation of existing function can occur much faster than the evolution of new pathways via
385 mutation [81], and in many cases nonsense mutations or deletions confer greater fitness benefit
386 than missense mutations affecting the same gene [82]. However, loss of function often comes at
387 the expense of metabolic flexibility, limiting the ability of evolved clones to compete in
388 alternative environments [20]. RpoS has been shown to be a high-value target of selection under
389 nutrient limitation: under low-nutrient conditions RpoS normally outcompetes RpoD for binding
390 to RNA polymerase, repressing genes required for growth and cell division and activating those
391 required to enter stationary phase [78,83]. *rpoS* mutants thus continue to divide under conditions
392 where wild-type cells arrest. In this respect, our genetic ‘point of departure’ could be viewed as
393 being pre-adapted to life under glucose limitation. However, the combined phenotypic effect of
394 ancestral *rpoS* and *rpoD* nonsense mutations in a suppressor background is murky and raises the
395 question of whether this combination of mutations is favorable under glucose limitation, merely
396 tolerated or detrimental. Despite the fact that many changes we observed (*galS*, upstream *mglB*,
397 *hfq*) enhance glucose assimilation, are predictable, occur repeatedly and rise to high frequency,
398 we also saw the persistence of clones with none of these mutations that instead carry intragenic
399 suppressors of the nonsense mutation in *rpoD* (*26→Asp and *26→Gln, chemostat 3 **Fig. 4**) or a
400 duplication that includes *rpoD*. If rapid adaptation can be driven by loss-of-function but occurs at
401 the expense of metabolic flexibility, nonsense mutations have a distinct advantage over deletions
402 in that reversion or suppression is possible should environmental conditions change [20].

403 Another ancestral allele that we expected to influence evolutionary trajectories was an
404 A→T CRP binding site mutation 224 bp upstream of the acetate scavenging enzyme, *acs* (acetyl-
405 CoA synthetase). This mutation alters regulation of the *acs-pta* operon such that the ancestor
406 poorly assimilates acetate excreted during growth under continuous glucose-limitation, opening

407 up a secondary resource for novel mutants that can [34]. Here, we uncovered no evidence for the
408 type of cross-feeding described in prior reports [29,30,33]. This result was not unanticipated, as
409 evidence for cross-feeding polymorphisms was observed in only half the evolution experiments
410 founded by this ancestor or its close relatives [33]. Moreover, a recent model [31] defining the
411 boundary conditions for cross-feeding to evolve in a chemostat showed that such an outcome is
412 sensitive to variation in dilution rate as well as to the relative fitness of *de novo* mutants that gain
413 access to secondary metabolites. Subtle differences in either of these parameters may account for
414 why we saw no evidence for acetate/glycerol/formate cross-feeding in our experiments. The
415 absence of such interactions may also be due to the fact that no variants arose at loci where
416 mutations have been implicated in cross-feeding evolution: *acs* (acetyl CoA synthetase), *lpd*
417 (lipoamide dehydrogenase) and *ptsI* (phosphoenolpyruvate phosphotransferase).

418 As expected, the ancestral *acs-pta* defect resulted in appreciable levels of residual acetate
419 (~45-90 μM) at the onset of our experiments (**Fig. S3**). While we uncovered no evidence for the
420 evolution of secondary resource specialists [30,31,33] and refs therein), residual acetate levels
421 consistently fell below detection limit by generation 200. Thus, adaptive mutants arising here
422 found other ways than cross-feeding to metabolize all available carbon. One possible work-
423 around may involve the *pgi* locus, which was second only to *galS* in the total number of
424 mutations recovered (**Table S2**). Generation 200 coincides with the emergence of mutant *pgi*
425 alleles in all three populations. In chemostat experiments with *pgi* deletion mutants, Yao et al.
426 found that in the absence of Pgi, glucose uptake rate drops slightly compared to wild-type, but no
427 overflow acetate is produced and biomass yield is unchanged [84].

428 ***Population and clone sequencing open up a detailed view of the full spectrum of***
429 ***beneficial mutations and how that spectrum changes over time*** High-coverage, whole-genome,

430 whole-population sequencing makes it possible to discover every new allele reaching $\geq 1\%$
431 frequency in a population of $>10^{10}$ cells. Because alleles are highly unlikely to reach such
432 frequencies by draft, all were either transiently beneficial or hitchhiking with alleles that were.
433 This depth of analysis opens up a richly-detailed view of the spectrum of beneficial mutations
434 arising in *E. coli* under constant resource limitation. Periodic whole-population sequencing
435 allows patterns to be discerned as to how these spectra change over time, while clone sequencing
436 makes it possible to establish linkage relations among novel alleles and represent their collective
437 fate as evolving lineages. Multiple patterns emerge from these analyses. *First*, new alleles
438 accumulate in replicate populations at similar rates, and the proportion of alleles that are
439 missense, nonsense, synonymous, or noncoding remains fairly constant; the great majority is
440 either missense (60-70%) or nonsense (5-10%). *Second*, the distribution of new mutations across
441 the genome is skewed, with only a few dozen of the more than 1,000 mutated genes having a
442 significant number of mutations; yet even among these most frequently mutated genes, few *de*
443 *novo* mutations fix. *Third*, by clonal sequencing we are able to determine that many, independent
444 lineages co-exist and compete within the culture. Thus, evolutionary dynamics in these
445 populations is governed by clonal interference rather than by clonal replacement or
446 reinforcement.

447 A *fourth* pattern to emerge is widespread parallelism in regulatory evolution. Both across
448 and within populations, the same genes are mutated again and again, often at exactly the same
449 nucleotide position in independent replicates, and sometimes in independent lineages co-
450 evolving in the same vessel. Many of these genes (*galS*, *malT*, *malK*, upstream *mglB*, *hfq*, *rho*)
451 act in processes related to the transport and assimilation of the limiting nutrient, glucose.

452 However, in most cases the mutations recovered alter regulation of these processes, and not the
453 structural proteins that carry them out.

454 A *fifth* pattern relates to the order of beneficial mutations and the influence that order has
455 on dynamics. Consistent with previous reports, mutations that increase glucose flux across the
456 inner membrane (*galS*, upstream *mglB*) occur early and precede those that increase flux across
457 the outer membrane (*malK/malT*, *hfq*, *rho*). In both cases, mutations in binding partners
458 (*GalS*/upstream *mglB* and *MalT/MalK*) rarely occur in the same clone, and the order in which
459 they occur can lead to either a sweep (upstream *mglB* clones quickly displace *galS* clones) or
460 clonal interference (*malT* and *malK* clones can co-exist). Other alleles appear to emerge later in
461 the experiment and nearly always together: clones with existing mutations in the *mal* operon
462 acquire subsequent mutations in *hfq* and *opgH*, regardless of which gene is altered first and
463 which alleles are already present in the population (Fig. S1). These patterns are echoed by the
464 genotypes reported by Kinnersley et al. [22] in which glucose scavenger CV103 has mutations in
465 *malK*, *opgG* and *hfq* while acetate specialist CV101 only carries a mutation in *malT*.

466 Similar experiments carried out by Maharjan et al. [32] using *E. coli* BW2952
467 demonstrated that under continuous glucose limitation population-level phenotypic changes are
468 often the result of multiple soft sweeps by combinations of beneficial mutations. While we did
469 not assay clone phenotypes, multiple alleles of *galS*, *hfq* and *opgH* appear to sweep our
470 populations in concert suggesting a similar pattern in which a phenotypic effect (reduced
471 expression of a particular gene) is favored, but has different genetic bases in co-existing lineages.
472 At the clone level, BW2952 also exhibits sign epistasis between mutations in *rpoS/hfq* and
473 *galS/malT* [32,52]. In our experiments, we did not uncover evidence of sign epistasis between
474 the ancestral *rpoS* allele and *hfq*: by generation 250, over 50% of clones in populations 1 and 3

475 carry mutations in both genes. Maharjan et al (2013) proposed that fitness deficits exhibited by
476 *rpoS/hfq* double mutants may be the result of altered cell division [62,85] and that *hfq* mutations
477 enhance glucose uptake during slow growth, but diminish viability when cells are dividing
478 rapidly. Hfq deletion mutants exhibit cell division anomalies due to elevated expression of cell
479 division proteins, including FtsZ [86,87]. Recent work by Hill et al. (2013) has shown that
480 during fast growth OpgH (which in our experiments is nearly always mutated alongside *hfq*)
481 binds FtsZ to postpone cell division. Thus, it may be that negative fitness effects experienced by
482 *hfq-rpoS* double mutants are the result of cell division anomalies mitigated by mutations in
483 *opgH*. It is noteworthy in this regard that whereas cells in the Maharjan et al. experiments
484 experienced a dilution rate of $D=0.1 \text{ hr}^{-1}$, those in evolutions performed by Adams, Helling and
485 colleagues were doubling twice as fast ($D=0.2 \text{ h}^{-1}$). Thus, this discrepancy may be a
486 manifestation of trade-offs between glucose uptake and cell viability.

487 Finally, some mutations occur repeatedly and are likely beneficial, but their dynamics are
488 unpredictable (e.g. beneficial mutations in transcriptional terminator *rho* sweep when they co-
489 occur with beneficial mutations upstream of *mglB*, but otherwise remain at low frequency (**Table**
490 **S3, Fig. S1**). This dependence on genetic context, or “quasi-hitchhiking”, of beneficial mutations
491 was previously observed by Lang et al. (2013) in evolving yeast populations and may be
492 consistent feature of microbial evolution experiments that becomes observable when populations
493 are sequenced at high depth of coverage and sufficient temporal resolution [14].

494 ***The evolution of population genetic complexity.*** Szostak, Hazen and others [88,89] argue
495 that a biological system’s complexity should be evaluated in terms of its functional information
496 content. Although the total number of alleles in an evolving population at any given time-point is
497 information content, it is functional only in how it is integrated among the lineages co-existing at

498 that time-point. Our approach of integrating population sequencing with clone sequencing makes
499 it possible to estimate the pace and extent with which complexity, measured as lineage-specific
500 functional information [88,89], emerges in replicate evolving populations originating from a
501 common ancestor. Implicit in our perspective is the assumption that the sequence differences by
502 which lineages can be distinguished have physiological and fitness consequences. For each
503 population, we calculated at 50-generation intervals three measures of complexity: Shannon's
504 Entropy (H), equitability (H/H_{\max}) and normalized population richness (Lineage counts) (**Fig.**
505 **S5**). All three measures increased during the course of evolution, but with a different tempo in
506 each population. Complexity increased in population 3 with no indication of reaching an
507 asymptote by generation 300 when the experiment was terminated. Populations 1 and 2 reached
508 asymptotes after ~400 generations, following a steady increase in population 1, but a nearly-300
509 generation period of stasis in Population 2. While longer experiments are clearly called for, our
510 finding that longer-term experiments appear to reach an asymptote in complexity is consistent
511 with theoretical [90] and empirical [91] observations that fitness plateaus when microbes are
512 cultured by serial transfer or in chemostats, even starting with mutator strains [92], and that
513 complexity itself may plateau when its evolution is simulated using RNA-like replicators [93]. It
514 is intriguing to contemplate the possibility that there may be a limit to the level of clonal
515 interference that can be sustained in asexual populations once all avenues for large fitness gains
516 have been exhausted. Indeed, it was recently shown using lineage tracking, that while fixation of
517 an adaptive mutant causes a stochastic crash in diversity, the generation of new adaptive mutants
518 within such a fixing lineage is expected to generate new diversity, such that a longer term steady
519 state level of diversity will be achieved [94].

520 Previous evolution experiments founded with the ancestor used here often resulted in
521 stable sub-populations supported by cross-feeding. In the present experiments neither the
522 spectrum of observed mutations nor the structure of clone phylogenies provides evidence for the
523 evolution of this type of trophic interaction. Instead, the possibility of a plateau in complexity,
524 coupled with the finding that every population has driven residual metabolites close to their
525 detection limit, suggest that these populations converge on an adaptive peak by diverse
526 mechanisms but that clonal interference keeps adaptive lineages off the summit, confined to
527 exploring the many roads by which the summit can be approached.

528

529 **Materials and Methods**

530 **Strains, media and culture conditions.** *Escherichia coli* JA122, population samples and clones
531 were maintained as permanent frozen stocks and stored at -80°C in 20% glycerol. Davis minimal
532 medium was used for all liquid cultures with 0.025% glucose added for batch cultures and
533 0.0125% for chemostats, as previously described [34]. Chemostat cultures were initiated using
534 colonies picked from Tryptone Agar (TA) plate inoculated with JA122, and outgrown in Davis
535 minimal batch medium overnight. Chemostats were maintained at 30°C with a dilution rate of ≈
536 0.2 hr⁻¹ for 300-500 generations. Every other day culture density was assessed by measuring
537 absorbance spectrophotometrically at A₅₅₀. Every other day, population samples were archived at
538 -80°C, and assayed for purity by plating serial dilutions on TA and examining Colony Forming
539 Units (CFU) that arose following 24-hr incubation at 30°C. When necessary, chemostats were re-
540 started from frozen stocks (chemostat 1: generation 217; chemostat 2: generation 410; chemostat
541 3 generation 251). At each sequencing time-point, 50 mL of culture was pelleted then stored at -

542 80°C for DNA extraction. For clone sequencing, entire colonies were picked from TA plates
543 inoculated from glycerol stocks, and re-archived in 96-well plate format.

544 **Metabolite assays.** 10 mL of sterile, cell-free chemostat filtrate was concentrated 20-fold by
545 lyophilization (Labconco 4.5 Liter Freeze Dry System), then re-suspended in 0.5 mL sterile
546 Millipore water. Residual glucose and residual acetate concentrations were determined on
547 concentrated filtrate. Glucose was assayed enzymatically using the High Sensitivity Glucose
548 Assay Kit (Sigma-Aldrich, Cat# MAK181), while acetate concentration was determined using
549 the Acetate Colorimetric Assay Kit (Sigma-Aldrich, Cat# MAK086). Results presented in **Fig.**
550 **S3** represent means ± SEM of duplicate assays.

551 **Population sequencing.** Bacterial DNA was prepared using the DNeasy Blood and Tissue Kit
552 (Qiagen, cat. 69504) following the manufacturer's guidelines. For population sequencing, 5 x
553 10^{10} cells, collected from every 50 generations in three chemostat vessels (up to 500 generations
554 in vessels 1 and 2, and up to 300 generations in vessel 3, 29 samples total) and frozen as pellets,
555 yielded 10-20 μ g of DNA. Following Proteinase K treatment, RNaseA treatment was used (20 μ L
556 10mg/mL RNase A, 2 min at room temperature) to avoid degraded RNA from visually
557 obscuring size selection during library preparation. Samples were split into two columns to avoid
558 overloading. Bacterial DNA was sheared to a 150-200bp fragment size using a Covaris S2 series
559 sonicator (6min, Duty=5%, Intensity=3, Cycles/Burst=200), and was then ligated to barcoded
560 adapters as described [95], except that 200bp fragments were size selected after adapter ligation
561 (to maximize the fidelity of sequencing, by reading each fragment in both directions). Six
562 barcoded libraries were combined and sequenced on each lane of HiSeq 2000 Illumina
563 Sequencer.

564 **Variant calling from population sequencing with CLC Genomics Workbench 7.5** Illumina
565 reads were trimmed (removing adapters on both ends) and stringently mapped (Mismatch cost 2,
566 Insertion cost 3, Deletion cost 3, Length fraction 1.0, Similarity fraction 0.97) to the reference
567 sequence (WIS_MG1655_m56). Variants were called with the following parameters: minimum
568 frequency 1%, minimal coverage 100, minimum count 2, and base quality filtering
569 (neighborhood radius 5, minimum central quality 15 and minimum neighborhood quality 20).
570 Sequencing data uncovered low-level contamination of whole population samples with *Serratia*
571 *liquifaciensis*. We therefore first determined proportion of contaminating reads by mapping
572 population sequencing to *S. liquifaciensis* genome and then removed SNPs with frequency
573 closely tracking percentage of contamination (between 1 and 5%) that matched *S. liquifaciensis*
574 sequence.

575 **Selection of clones for sequencing.** Allele frequencies for each chemostat were examined at
576 each time point, and the time-point at which there was the largest number of alleles present at 5%
577 or greater frequency was chosen for the isolation of clones for whole genome sequencing. The
578 rationale for this was that it would afford us the greatest opportunity to phase as many high
579 frequency alleles as possible.

580 **Clonal DNA preparation.** A colony was re-suspended in 300µL of sterile ddH₂O with 17%
581 glycerol and stored in three aliquots at -80°C. 100 µL of glycerol stock were used for DNA
582 preparation. After removing glycerol (using MultiScreen High Volume Filter Plates with 0.45
583 µm Durapore membrane, Millipore MVHVN4525), cell were resuspended in 500µL LB and
584 grown overnight at 30°C without shaking in deep well plates. Cells were collected again using
585 filter plates and subjected to DNeasy 96 Blood and Tissue Kit (Qiagen 69581) (yielding 4-15µg
586 per strain).

587 **Clonal libraries preparation and sequencing.** Multiplexed sequencing libraries from clones
588 were prepared using the Nextera kit (Illumina catalog # FC-121-1031 and # FC-121-1012) as
589 described in [96], starting with 1-4ng of genomic DNA. Resulting libraries from each 96-well
590 plate were pooled at an equal volume. Resulting pooled libraries were analyzed on the Qubit and
591 Bioanalyzer platforms and sequenced on HiSeq 2000 (one lane per 96 clone pool). All raw
592 sequencing data are available from the SRA under BioProject ID PRJNA517527.

593 **Variant calling from clonal sequencing with CLC Genomics Workbench 7.5** Short reads with
594 adapters removed were mapped to the reference with the same parameters as above, except the
595 length fraction was set to 0.5, and the similarity fraction to 0.8. Variants were called with a
596 minimum frequency 80%, minimum count 2 and the same base quality filtering as above.

597 **Generation of phylogenies.** For each chemostat, SNP and indel events for all 96 clones and the
598 ancestor JA122 were concatenated and re-coded as binary characters (i.e. presence/absence with
599 the ancestral state composed of all zeroes) and assembled into character matrices. PAUP ver.
600 4.0a149 was used to generate Camin-Sokal parsimony trees using the ancestor as the outgroup
601 under the assumption that reversions were extremely unlikely due to the extreme transversion
602 bias [97,98]. Tree files (.tre) were loaded into the Interactive Tree of Life (iTOL) web service for
603 character mapping and figure generation [99].

604 **Determining genes with an excess of mutations.** To identify genes with an excess of mutations,
605 we first determined the overall density of mutations as:

606 $\rho = M/L$, where M is the total number of mutations, and L is the length of the genome.

607 The probability of a given mutation landing in a segment of length l , is:

608 $\lambda = \rho \times l$

609 To calculate the p-value of n mutations landing in a segment of length l , we assume a Poisson
610 sampling process of a mutation landing in a given segment, and thus use:

$$p = \sum_{i=n}^{\infty} \frac{\lambda^i e^{-\lambda}}{i!}$$

611 though in practice, we capped i arbitrarily at 50, as continually summing at $i > 50$ does not
612 appreciably affect the calculated p-value. For a given segment, we calculated the number of
613 segments that would be expected to have p-value as good or better, as the number of tested
614 segments multiplied by the p-value. From this, we also determined a false positive rate.

615 ***Generation of Muller diagrams.*** Based on both the clonal sequencing we were able to determine
616 which mutations were in which lineages together, and from both the clonal and population
617 sequencing an approximate order of those mutations (though this was not exhaustive for all
618 mutations). Using these data, we developed a lineage file format that described which mutations
619 occurred in which lineages, and which lineages descended from one another, and used a custom
620 Perl script that combined this information with the allele frequencies over time from the
621 population sequencing to generate a graphical representation of the evolutionary dynamics, often
622 referred to as a Muller diagram.

623

624 **Acknowledgements**

625 The authors thank Matthew Herron and Eugene Kroll for their careful reading of the manuscript
626 and their thoughtful suggestions for its improvement.

627 **Author Contributions**

628 Conceived and designed the experiments: GS MK FR. Performed the experiments: MK and KS.
629 Analyzed the data: GS KS MK JB DY FR. Contributed reagents/materials/analysis tools: GS.

630 Wrote the paper: MK GS FR

631

632 **Main Figure legends**

633 **Fig. 1. Input of de novo mutations.** The rate at which novel alleles appear, and the proportion
634 of synonymous, missense and nonsense mutations, and non-coding mutations are consistent
635 across all three replicate evolutions.

636

637 **Fig. 2. Most *de novo* mutations only reach low allele frequencies, and experience pervasive
638 clonal interference.** (A) Histogram of maximum allele frequencies from three replicate
639 evolutions, (B) Final versus maximum allele frequency for each *de novo* mutation, shows most
640 mutations are at a lower frequency at the end of the experiment than they were at their
641 maximum.

642 **Fig. 3. Isolated clones are representative of the populations from which they are drawn.**
643 Mutation frequencies for population and clonal sequencing for mutations identified in both
644 datasets at the same time-point shows similar frequencies.
645

646 **Fig. 4. Clone phylogenies.** Phylogenies depicting relationships among sequenced clones isolated
647 from chemostats when allelic diversity attained its maximum. Distributions of different *malK*,
648 *malT*, *fimH*, *hfq* and *opgH* alleles are indicated by colored bars. For each gene, all alleles
649 observed in the dataset are numbered (see File S2 for details of which number corresponds to
650 which allele for each gene). Underlined numbers denote alleles independently observed in more
651 than one chemostat, while numbers marked with an asterisk appear to have arisen more than
652 once within the same vessel. Grey shading delineates clades comprised of clones that have not
653 acquired the standard mutations related to enhanced glucose uptake and instead carry variant
654 *fimH* alleles that contribute to biofilm formation. Bracketed clones in chemostat 3 exhibited

655 mutations expected to revert the ancestral nonsense mutations in the housekeeping gene
656 encoding sigma factor RpoD.

657 **Fig. 5. Muller diagrams.** Evolutionary dynamics of adaptive lineages, deduced from combining
658 whole-population whole-genome sequence data and whole-genome sequence data of individual
659 clones isolated from each chemostat at the time-point where allelic diversity reached its
660 maximum value. Select genes are indicated in the plots – see Figure S1 and Supplementary Files
661 S3-S5 for additional details. Also note, most mutations that went extinct by the sampling
662 timepoint are not shown. See Figure S1 for their relative frequencies.

663

664 **Fig. 6. Overview of pathways relating some of the most frequently mutated genes to glucose**
665 **transport and metabolism.** Numbers in parentheses next to protein/gene names denote the
666 number of mutant alleles found in each chemostat population over the course of 300-500
667 generations (also see Table S3).

668

669

670 **Supplementary Figure Legends**

671 **Fig. S1. Population-level dynamics of mutations in 10 frequently hit genes show consistent**
672 **patterns.** For each panel, elapsed generations are depicted on the x-axis and the height of each
673 grey box represents a frequency of 100%. Cumulative frequencies for all alleles of a given gene
674 present in the population at each time point were calculated and are represented as colored plots.
675 (A) chemostat 1 (B) chemostat 2 (C) chemostat 3.

676

677 **Fig. S2. MalK/MalT population dynamics.** Mutant alleles of both LamB regulators, *malT* and
678 *malK*, seldom co-occur in the same lineage, and when they do, those lineages fail to go to high
679 frequency.

680

681 **Fig. S3. Cell density and residual metabolite concentrations.** Chemostat populations at steady
682 state exhibit balanced growth, where population size remains constant and the limiting substrate,
683 glucose is near its detection limit. As expected, populations initially produce the overflow
684 metabolite acetate, as the founder carries a mutation that dysregulates acetyl CoA synthetase, the
685 chief route by which *E. coli* assimilates low levels of this metabolite.

686

687 **Fig. S4. Mutations in glucosyltransferase *opgH* occur repeatedly and go to high frequency.**

688

689 **Fig. S5. Patterns of change in population genetic complexity.** Shannon's Entropy [H],
690 Equitability [H/Hmax] and normalized population Richness [Lineage counts] were calculated at
691 50 generation intervals for each of three replicate evolution experiments. Shannon's entropy is an
692 effective metric of population diversity as it accounts for both lineage richness (the number of

693 lineages observed) and the relative abundance of different lineages (evenness). Lineage richness
694 was normalized between zero and one by dividing the number of observed lineages, S , by the
695 maximum S observed over the course of each experiment.

696

697 **Main Tables**

698 **Table 1. Characteristics of frequently mutated genes.** Each asterisk indicates an allele that
699 arose more than once independently, either within or between vessels.

700

701 **Supplementary Tables**

702 **Table S1.** Key mutations that distinguish ancestral strain JA122 from K12 (MG1655)

703 **Table S2.** Beneficial alleles

704 **Table S3.** Population allele frequencies for frequently mutated genes

705 **Table S4.** Identical mutations arise within and among replicate evolution experiments.

706 **Table S5.** Fixed alleles among replicate populations (“fixed” defined as >98% at any time point
707 between generation 50 and 500).

708

709 **Supplementary Data Files**

710 **File S1.** Identity and frequencies of mutations detected via population sequencing.

711 **File S2.** Alleles mapped onto clone phylogenies represented in main Fig. 4.

712 **File S3.** Muller diagrams for novel alleles arising in chemostat 1, showing details for each
713 lineage

714 **File S4.** Muller diagrams for novel alleles arising in chemostat 2, showing details for each
715 lineage

716 **File S5.** Muller diagrams for novel alleles arising in chemostat 3, showing details for each

717 lineage.

718

719 **Table 1 Frequently mutated genes.** Each asterisk indicates an allele that
720 arose more than once independently, either within or between vessels.

Rank	Gene	Observed Mutations	Expected Mutations	P-Value	FDR
Population sequencing					
1	gals***	38	0.78	6.55E-50	4.42E-45
2	hfq*****	24	0.23	6.91E-40	2.33E-35
3	pgi*****	35	1.23	4.54E-38	1.02E-33
4	opgH**	31	1.90	8.74E-27	1.48E-22
5	malT*****	30	2.02	8.10E-25	1.09E-20
6	malK*****	22	0.83	7.47E-24	8.40E-20
7	upstream mglB**	7	0.21	2.91E-09	2.81E-05
8	rho**	11	0.94	5.49E-09	4.64E-05
9	upstream dnaG	5	0.08	3.06E-08	2.30E-04
10	fimH***	9	0.68	4.38E-08	2.96E-04
Clonal sequencing					
1	hfq*****	26	0.1020	3.79E-53	3.68E-48
2	pgi****	17	0.5448	5.52E-20	2.68E-15
3	opgH***	17	0.8400	6.57E-17	2.13E-12
4	upstream mglB**	8	0.0925	1.22E-13	2.96E-09
5	fimH****	10	0.2982	1.17E-12	2.26E-08
6	ompR***	8	0.2377	2.05E-10	3.31E-06
7	upstream adhE*	6	0.1575	1.85E-08	0.000257
8	malT***	10	0.8935	3.98E-08	0.000482
9	proQ*	6	0.2308	1.72E-07	0.001858
10	pfkA**	6	0.3180	1.09E-06	0.010612

721

722

723

724 **References**

- 725 1. Lang GI, Desai MM (2014) The spectrum of adaptive mutations in experimental evolution.
726 *Genomics* 104: 412-416.
- 727 2. Cvijovic I, Nguyen Ba AN, Desai MM (2018) Experimental studies of evolutionary dynamics
728 in Microbes. *Trends Genet* 34: 693-703.
- 729 3. Cooper VS (2018) Experimental Evolution as a high-throughput screen for genetic
730 adaptations. *mSphere* 3.
- 731 4. Garay E, Campos SE, Gonzalez de la Cruz J, Gaspar AP, Jinich A, et al. (2014) High-
732 resolution profiling of stationary-phase survival reveals yeast longevity factors and their
733 genetic interactions. *PLoS Genet* 10: e1004168.
- 734 5. Paradis-Bleau C, Kritikos G, Orlova K, Typas A, Bernhardt TG (2014) A genome-wide screen
735 for bacterial envelope biogenesis mutants identifies a novel factor involved in cell wall
736 precursor metabolism. *PLoS Genet* 10: e1004056.
- 737 6. Novick A, Szilard L (1950) Experiments with the chemostat on spontaneous mutations of
738 bacteria. *Proc Natl Acad Sci U S A* 36: 708-719.
- 739 7. Novick A, Szilard L (1951) Experiments on spontaneous and chemically induced mutations of
740 bacteria growing in the Chemostat. *Cold Spring Harb Symp Quant Biol* 16: 337-343.
- 741 8. Atwood KC, Schneider LK, Ryan FJ (1951) Periodic selection in *Escherichia coli*. *Proc Natl
742 Acad Sci U S A* 37: 146-155.
- 743 9. Atwood KC, Schneider LK, Ryan FJ (1951) Selective mechanisms in bacteria. *Cold Spring
744 Harb Symp Quant Biol* 16: 345-355.
- 745 10. Haldane JBS (1927) A Mathematical Theory of Natural and Artificial Selection, Part V:
746 Selection and Mutation. *Mathematical Proceedings of the Cambridge Philosophical
747 Society* 23: 838-844.
- 748 11. Fisher RAS (1930) The genetical theory of natural selection. Oxford: Clarendon Press.
- 749 12. Muller HJ (1932) Some genetic aspects of sex. *The American Naturalist* 66: 118-138.
- 750 13. Gause GF (1934) Experimental analysis of Vito Volterra's mathematical theory of the
751 struggle for existence. *Science* 79: 16-17.
- 752 14. Lang GI, Rice DP, Hickman MJ, Sodergren E, Weinstock GM, et al. (2013) Pervasive
753 genetic hitchhiking and clonal interference in forty evolving yeast populations. *Nature*
754 500: 571-574.
- 755 15. Marad DA, Buskirk SW, Lang GI (2018) Altered access to beneficial mutations slows
756 adaptation and biases fixed mutations in diploids. *Nat Ecol Evol* 2: 882-889.
- 757 16. Levy SF, Blundell JR, Venkataram S, Petrov DA, Fisher DS, et al. (2015) Quantitative
758 evolutionary dynamics using high-resolution lineage tracking. *Nature* 519: 181-186.
- 759 17. Good BH, McDonald MJ, Barrick JE, Lenski RE, Desai MM (2017) The dynamics of
760 molecular evolution over 60,000 generations. *Nature* 551: 45-50.
- 761 18. Lauer S, Avecilla G, Spealman P, Sethia G, Brandt N, et al. (2018) Single-cell copy number
762 variant detection reveals the dynamics and diversity of adaptation. *PLoS Biol* 16:
763 e3000069.
- 764 19. Kao KC, Sherlock G (2008) Molecular characterization of clonal interference during
765 adaptive evolution in asexual populations of *Saccharomyces cerevisiae*. *Nat Genet* 40:
766 1499-1504.
- 767 20. Kvitek DJ, Sherlock G (2013) Whole genome, whole population sequencing reveals that loss
768 of signaling networks is the major adaptive strategy in a constant environment. *PLoS
769 Genet* 9: e1003972.

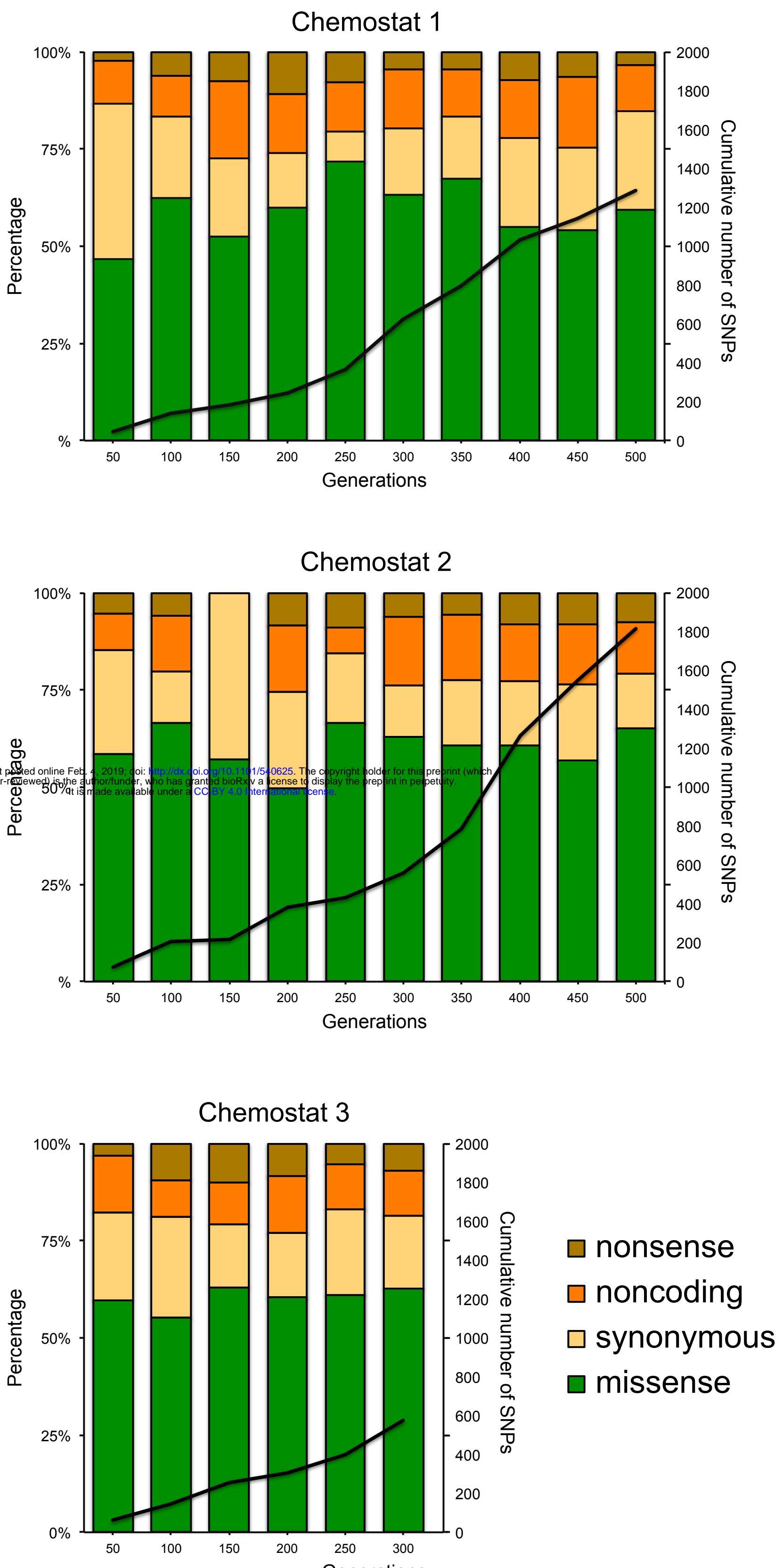
- 770 21. Behringer MG, Choi BI, Miller SF, Doak TG, Karty JA, et al. (2018) *Escherichia coli*
771 cultures maintain stable subpopulation structure during long-term evolution. *Proc Natl
772 Acad Sci U S A* 115: E4642-E4650.
- 773 22. Rozen DE, Lenski RE (2000) Long-term experimental evolution in *Escherichia coli*. VIII.
774 Dynamics of a balanced polymorphism. *Am Nat* 155: 24-35.
- 775 23. Herron MD, Doeblei M (2013) Parallel evolutionary dynamics of adaptive diversification in
776 *Escherichia coli*. *PLoS Biol* 11: e1001490.
- 777 24. Maddamsetti R, Lenski RE, Barrick JE (2015) Adaptation, clonal interference, and
778 frequency-dependent interactions in a long-term evolution experiment with *Escherichia*
779 *coli*. *Genetics* 200: 619-631.
- 780 25. Charlesworth D (2006) Balancing selection and its effects on sequences in nearby genome
781 regions. *PLoS Genet* 2: e64.
- 782 26. Rainey PB, Travisano M (1998) Adaptive radiation in a heterogeneous environment. *Nature*
783 394: 69-72.
- 784 27. Rozen DE, Schneider D, Lenski RE (2005) Long-term experimental evolution in *Escherichia*
785 *coli*. XIII. Phylogenetic history of a balanced polymorphism. *J Mol Evol* 61: 171-180.
- 786 28. Kinnersley M, Wenger J, Kroll E, Adams J, Sherlock G, et al. (2014) Ex uno plures: clonal
787 reinforcement drives evolution of a simple microbial community. *PLoS Genet* 10:
788 e1004430.
- 789 29. Rosenzweig RF, Sharp RR, Treves DS, Adams J (1994) Microbial evolution in a simple
790 unstructured environment: genetic differentiation in *Escherichia coli*. *Genetics* 137: 903-
791 917.
- 792 30. Helling RB, Vargas CN, Adams J (1987) Evolution of *Escherichia coli* during growth in a
793 constant environment. *Genetics* 116: 349-358.
- 794 31. Gudelj I, Kinnersley M, Rashkov P, Schmidt K, Rosenzweig F (2016) Stability of cross-
795 feeding polymorphisms in microbial communities. *PLoS Comput Biol* 12: e1005269.
- 796 32. Maharjan RP, Liu B, Feng L, Ferenci T, Wang L (2015) Simple phenotypic sweeps hide
797 complex genetic changes in populations. *Genome Biol Evol* 7: 531-544.
- 798 33. Treves DS, Manning S, Adams J (1998) Repeated evolution of an acetate-crossfeeding
799 polymorphism in long-term populations of *Escherichia coli*. *Mol Biol Evol* 15: 789-797.
- 800 34. Kinnersley MA, Holben WE, Rosenzweig F (2009) E Unibus Plurum: genomic analysis of
801 an experimentally evolved polymorphism in *Escherichia coli*. *PLoS Genet* 5: e1000713.
- 802 35. Atlung T, Nielsen HV, Hansen FG (2002) Characterisation of the allelic variation in the *rpoS*
803 gene in thirteen K12 and six other non-pathogenic *Escherichia coli* strains. *Mol Genet
804 Genomics* 266: 873-881.
- 805 36. Singaravelan B, Roshini BR, Munavar MH (2010) Evidence that the supE44 mutation of
806 *Escherichia coli* is an amber suppressor allele of *glnX* and that it also suppresses ochre
807 and opal nonsense mutations. *J Bacteriol* 192: 6039-6044.
- 808 37. Au KG, Clark S, Miller JH, Modrich P (1989) *Escherichia coli* *mutY* gene encodes an
809 adenine glycosylase active on G-A mispairs. *Proc Natl Acad Sci U S A* 86: 8877-8881.
- 810 38. Blattner FR, Plunkett G, 3rd, Bloch CA, Perna NT, Burland V, et al. (1997) The complete
811 genome sequence of *Escherichia coli* K-12. *Science* 277: 1453-1462.
- 812 39. Lee H, Popodi E, Tang H, Foster PL (2012) Rate and molecular spectrum of spontaneous
813 mutations in the bacterium *Escherichia coli* as determined by whole-genome sequencing.
814 *Proc Natl Acad Sci U S A* 109: E2774-2783.

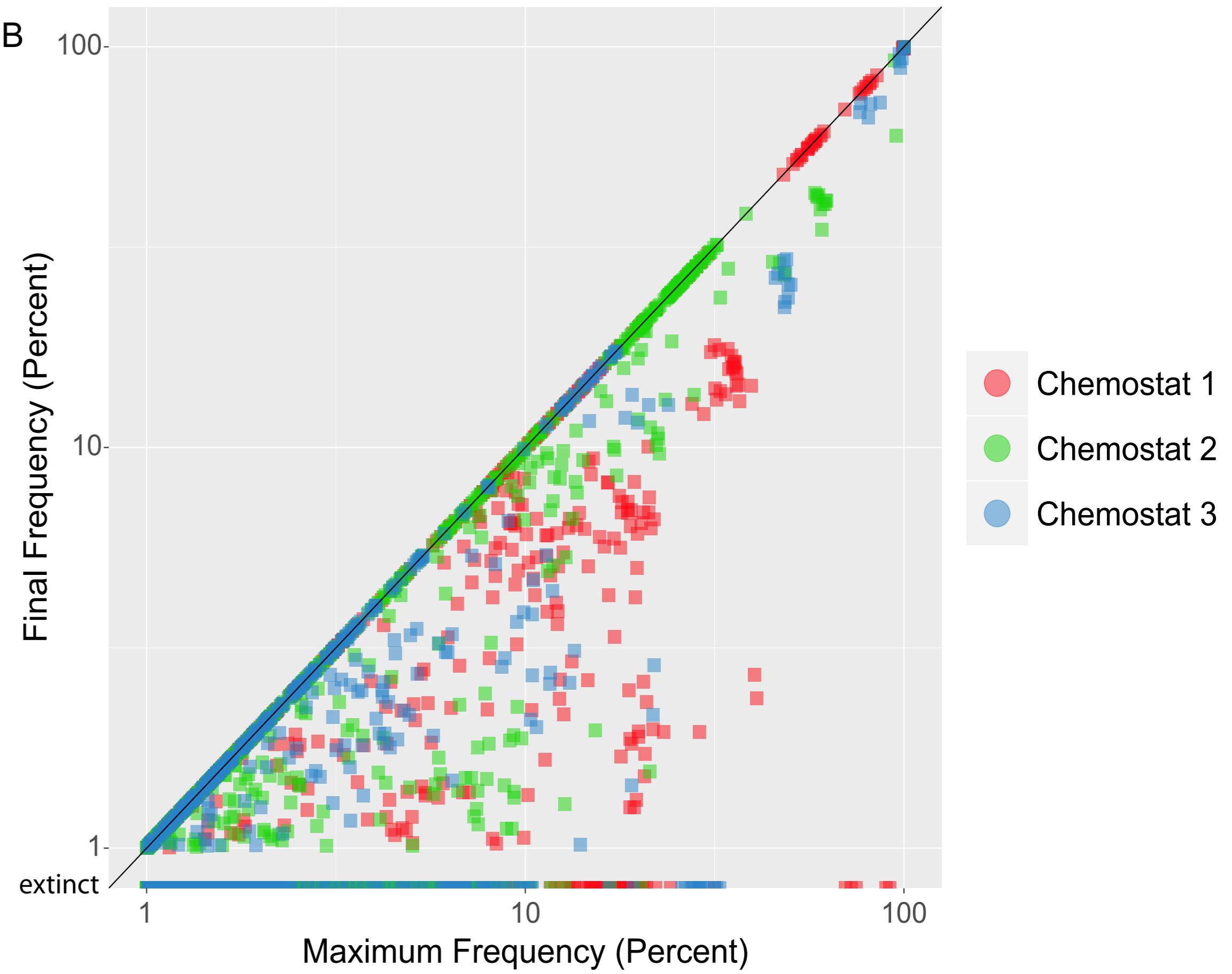
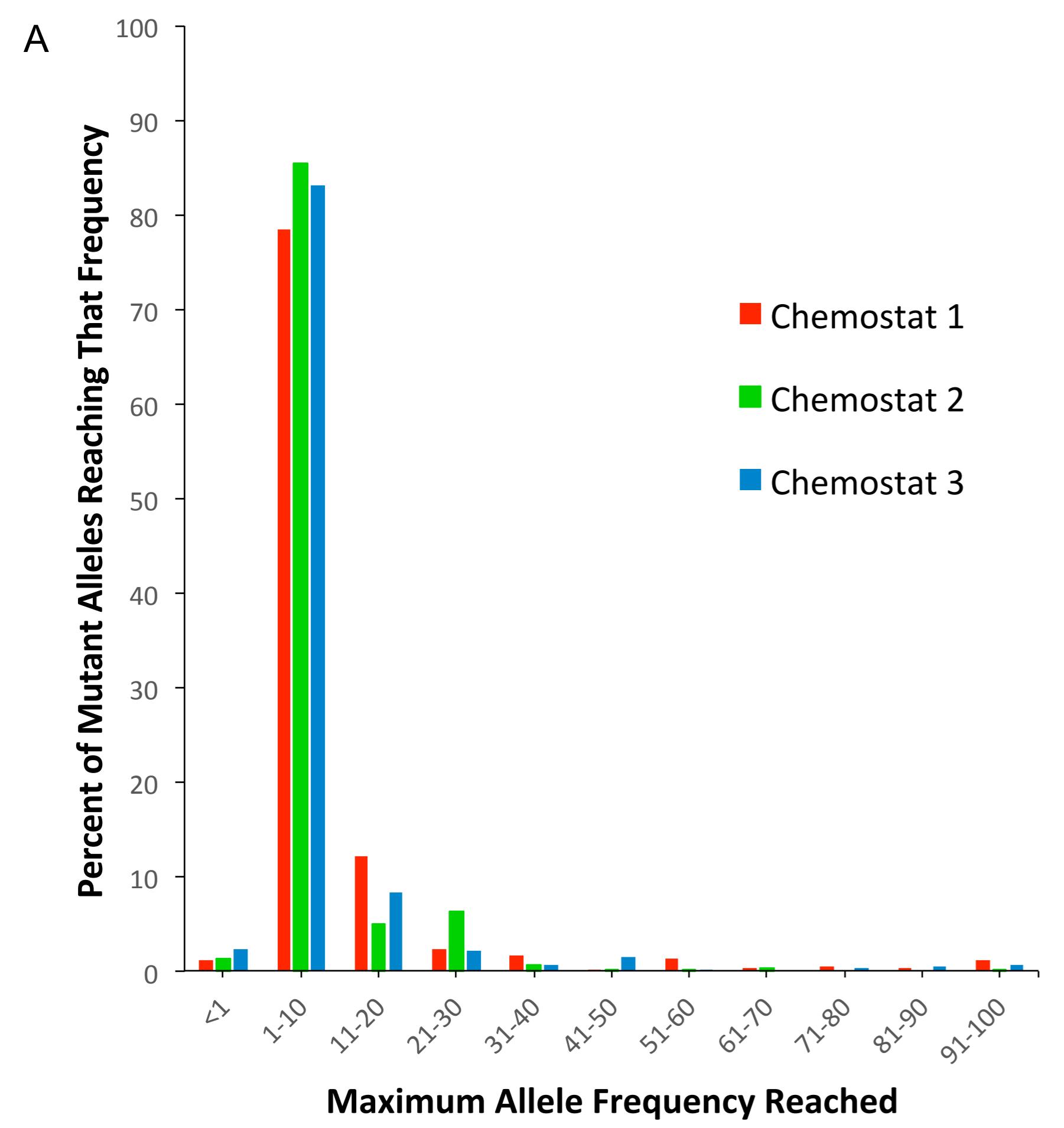
- 815 40. de Visser JA, Rozen DE (2006) Clonal interference and the periodic selection of new
816 beneficial mutations in *Escherichia coli*. *Genetics* 172: 2093-2100.
817 41. Hermisson J, Pennings PS (2005) Soft sweeps: molecular population genetics of adaptation
818 from standing genetic variation. *Genetics* 169: 2335-2352.
819 42. Pennings PS, Hermisson J (2006) Soft sweeps II--molecular population genetics of
820 adaptation from recurrent mutation or migration. *Mol Biol Evol* 23: 1076-1084.
821 43. Pennings PS, Hermisson J (2006) Soft sweeps III: the signature of positive selection from
822 recurrent mutation. *PLoS Genet* 2: e186.
823 44. Desai MM, Walczak AM, Fisher DS (2013) Genetic diversity and the structure of
824 genealogies in rapidly adapting populations. *Genetics* 193: 565-585.
825 45. Notley-McRobb L, Ferenci T (1999) The generation of multiple co-existing mal-regulatory
826 mutations through polygenic evolution in glucose-limited populations of *Escherichia coli*.
827 *Environ Microbiol* 1: 45-52.
828 46. Hermisson J, Pennings PS (2017) Soft sweeps and beyond: understanding the patterns and
829 probabilities of selection footprints under rapid adaptation. *Methods in Ecology and
830 Evolution* 8: 700-716.
831 47. Jensen JD (2014) On the unfounded enthusiasm for soft selective sweeps. *Nat Commun* 5:
832 5281.
833 48. Fogle CA, Nagle JL, Desai MM (2008) Clonal interference, multiple mutations and
834 adaptation in large asexual populations. *Genetics* 180: 2163-2173.
835 49. Buskirk SW, Peace RE, Lang GI (2017) Hitchhiking and epistasis give rise to cohort
836 dynamics in adapting populations. *Proc Natl Acad Sci U S A* 10.1073/pnas.1702314114.
837 50. Diaz Caballero J, Clark ST, Coburn B, Zhang Y, Wang PW, et al. (2015) Selective Sweeps
838 and Parallel Pathoadaptation Drive *Pseudomonas aeruginosa* Evolution in the Cystic
839 Fibrosis Lung. *MBio* 6: e00981-00915.
840 51. Ferenci T (2001) Hungry bacteria--definition and properties of a nutritional state. *Environ
841 Microbiol* 3: 605-611.
842 52. Maharjan RP, Ferenci T (2013) Epistatic interactions determine the mutational pathways and
843 coexistence of lineages in clonal *Escherichia coli* populations. *Evolution* 67: 2762-2768.
844 53. Manch K, Notley-McRobb L, Ferenci T (1999) Mutational adaptation of *Escherichia coli* to
845 glucose limitation involves distinct evolutionary pathways in aerobic and oxygen-limited
846 environments. *Genetics* 153: 5-12.
847 54. Notley-McRobb L, Ferenci T (1999) Adaptive mgl-regulatory mutations and genetic
848 diversity evolving in glucose-limited *Escherichia coli* populations. *Environ Microbiol* 1:
849 33-43.
850 55. Notley-McRobb L, Ferenci T (2000) Experimental analysis of molecular events during
851 mutational periodic selections in bacterial evolution. *Genetics* 156: 1493-1501.
852 56. Notley-McRobb L, Seeto S, Ferenci T (2003) The influence of cellular physiology on the
853 initiation of mutational pathways in *Escherichia coli* populations. *Proc Biol Sci* 270: 843-
854 848.
855 57. Ferenci T (1996) Adaptation to life at micromolar nutrient levels: the regulation of
856 *Escherichia coli* glucose transport by endoinduction and cAMP. *FEMS Microbiol Rev*
857 18: 301-317.
858 58. Death A, Ferenci T (1993) The importance of the binding-protein-dependent Mgl system to
859 the transport of glucose in *Escherichia coli* growing on low sugar concentrations. *Res
860 Microbiol* 144: 529-537.

- 861 59. Death A, Notley L, Ferenci T (1993) Derepression of LamB protein facilitates outer
862 membrane permeation of carbohydrates into *Escherichia coli* under conditions of nutrient
863 stress. *J Bacteriol* 175: 1475-1483.
- 864 60. Tweeddale H, Notley-McRobb L, Ferenci T (1999) Assessing the effect of reactive oxygen
865 species on *Escherichia coli* using a metabolome approach. *Redox Rep* 4: 237-241.
- 866 61. Geanacopoulos M, Adhya S (1997) Functional characterization of roles of GalR and GalS as
867 regulators of the gal regulon. *J Bacteriol* 179: 228-234.
- 868 62. Maharjan R, McKenzie C, Yeung A, Ferenci T (2013) The basis of antagonistic pleiotropy in
869 hfq mutations that have opposite effects on fitness at slow and fast growth rates. *Heredity*
870 (Edinb) 110: 10-18.
- 871 63. Maharjan R, Zhou Z, Ren Y, Li Y, Gaffe J, et al. (2010) Genomic identification of a novel
872 mutation in hfq that provides multiple benefits in evolving glucose-limited populations of
873 *Escherichia coli*. *J Bacteriol* 192: 4517-4521.
- 874 64. Richet E, Joly N, Danot O (2005) Two domains of MalT, the activator of the *Escherichia coli*
875 maltose regulon, bear determinants essential for anti-activation by MalK. *J Mol Biol* 347:
876 1-10.
- 877 65. Dardonville B, Raibaud O (1990) Characterization of malT mutants that constitutively
878 activate the maltose regulon of *Escherichia coli*. *J Bacteriol* 172: 1846-1852.
- 879 66. Bohm A, Diez J, Diederichs K, Welte W, Boos W (2002) Structural model of MalK, the
880 ABC subunit of the maltose transporter of *Escherichia coli*: implications for mal gene
881 regulation, inducer exclusion, and subunit assembly. *J Biol Chem* 277: 3708-3717.
- 882 67. Moller P, Overloper A, Forstner KU, Wen TN, Sharma CM, et al. (2014) Profound impact of
883 Hfq on nutrient acquisition, metabolism and motility in the plant pathogen
884 *Agrobacterium tumefaciens*. *PLoS One* 9: e110427.
- 885 68. Maharjan RP, Ferenci T, Reeves PR, Li Y, Liu B, et al. (2012) The multiplicity of divergence
886 mechanisms in a single evolving population. *Genome Biol* 13: R41.
- 887 69. Cardinale CJ, Washburn RS, Tadigotla VR, Brown LM, Gottesman ME, et al. (2008)
888 Termination factor Rho and its cofactors NusA and NusG silence foreign DNA in *E. coli*.
889 *Science* 320: 935-938.
- 890 70. Banerjee S, Chalissery J, Bandey I, Sen R (2006) Rho-dependent transcription termination:
891 more questions than answers. *J Microbiol* 44: 11-22.
- 892 71. Peters JM, Mooney RA, Kuan PF, Rowland JL, Keles S, et al. (2009) Rho directs widespread
893 termination of intragenic and stable RNA transcription. *Proc Natl Acad Sci U S A* 106:
894 15406-15411.
- 895 72. Ciampi MS (2006) Rho-dependent terminators and transcription termination. *Microbiology*
896 152: 2515-2528.
- 897 73. Colonna B, Hofnung M (1981) rho Mutations restore lamB expression in *E. coli* K12 strains
898 with an inactive malB region. *Mol Gen Genet* 184: 479-483.
- 899 74. Nakashima N, Ohno S, Yoshikawa K, Shimizu H, Tamura T (2014) A vector library for
900 silencing central carbon metabolism genes with antisense RNAs in *Escherichia coli*. *Appl*
901 *Environ Microbiol* 80: 564-573.
- 902 75. Hill NS, Buske PJ, Shi Y, Levin PA (2013) A moonlighting enzyme links *Escherichia coli*
903 cell size with central metabolism. *PLoS Genet* 9: e1003663.
- 904 76. Lange R, Hengge-Aronis R (1994) The cellular concentration of the sigma S subunit of RNA
905 polymerase in *Escherichia coli* is controlled at the levels of transcription, translation, and
906 protein stability. *Genes Dev* 8: 1600-1612.

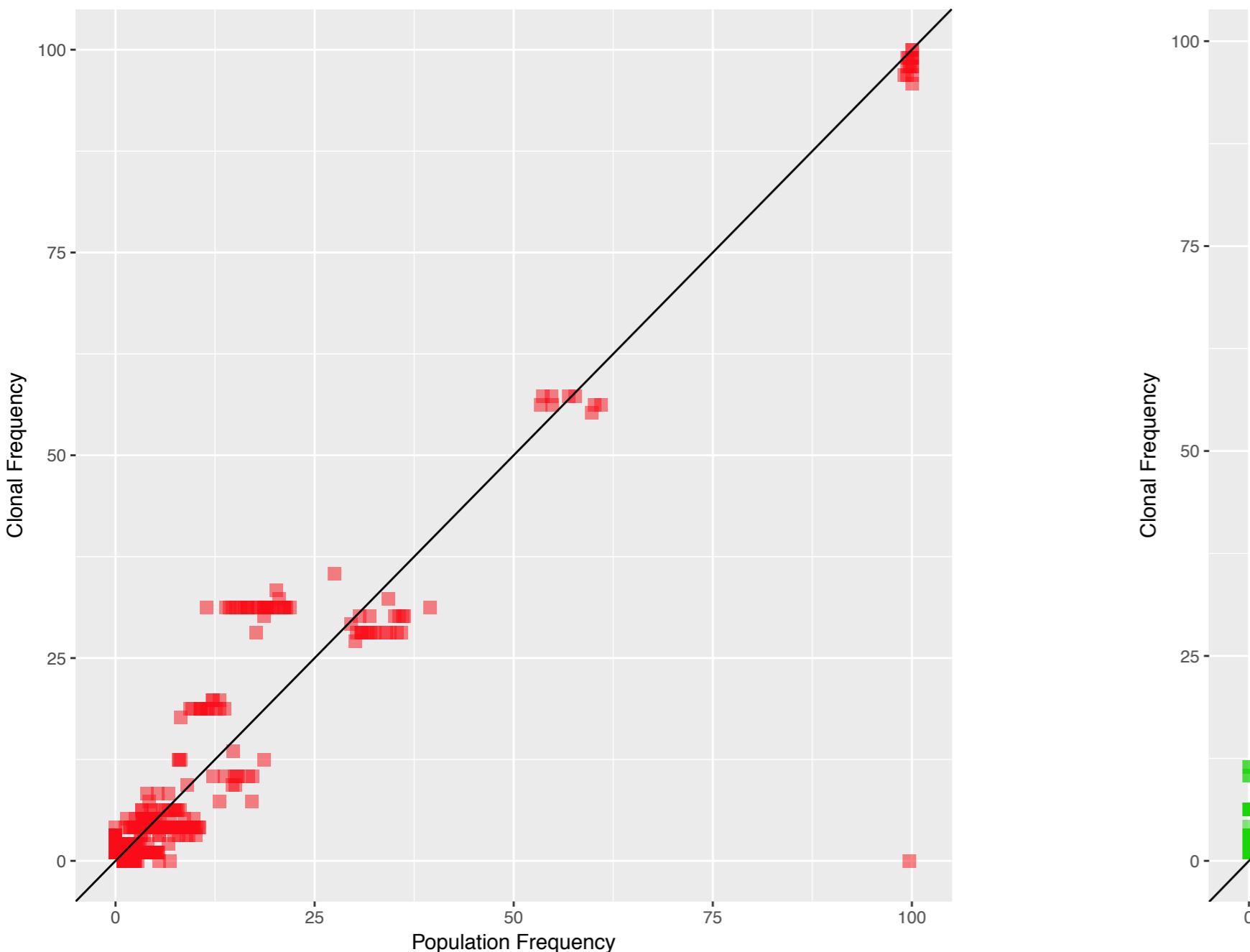
- 907 77. King T, Ishihama A, Kori A, Ferenci T (2004) A regulatory trade-off as a source of strain
908 variation in the species *Escherichia coli*. *J Bacteriol* 186: 5614-5620.
909 78. Notley-McRobb L, King T, Ferenci T (2002) *rpoS* mutations and loss of general stress
910 resistance in *Escherichia coli* populations as a consequence of conflict between
911 competing stress responses. *J Bacteriol* 184: 806-811.
912 79. Versalovic J, Koeuth T, Britton R, Geszvain K, Lupski JR (1993) Conservation and
913 evolution of the *rpsU-dnaG-rpoD* macromolecular synthesis operon in bacteria. *Mol*
914 *Microbiol* 8: 343-355.
915 80. Hutchins PR, Miller SR (2017) Genomics of variation in nitrogen fixation activity in a
916 population of the thermophilic cyanobacterium *Mastigocladus laminosus*. *ISME J* 11: 78-
917 86.
918 81. Hottes AK, Freddolino PL, Khare A, Donnell ZN, Liu JC, et al. (2013) Bacterial adaptation
919 through loss of function. *PLoS Genet* 9: e1003617.
920 82. Venkataram S, Dunn B, Li Y, Agarwala A, Chang J, et al. (2016) Development of a
921 Comprehensive Genotype-to-Fitness Map of Adaptation-Driving Mutations in Yeast.
922 *Cell* 166: 1585-1596 e1522.
923 83. Phan K, Ferenci T (2013) A design-constraint trade-off underpins the diversity in
924 ecologically important traits in species *Escherichia coli*. *ISME J* 7: 2034-2043.
925 84. Yao R, Hirose Y, Sarkar D, Nakahigashi K, Ye Q, et al. (2011) Catabolic regulation analysis
926 of *Escherichia coli* and its *crp*, *mlc*, *mgsA*, *pgi* and *ptsG* mutants. *Microb Cell Fact* 10:
927 67.
928 85. Vecerek B, Rajkowitsch L, Sonnleitner E, Schroeder R, Blasi U (2008) The C-terminal
929 domain of *Escherichia coli* Hfq is required for regulation. *Nucleic Acids Res* 36: 133-
930 143.
931 86. Takada A, Wachi M, Nagai K (1999) Negative regulatory role of the *Escherichia coli* *hfq*
932 gene in cell division. *Biochem Biophys Res Commun* 266: 579-583.
933 87. Zambrano N, Guichard PP, Bi Y, Cayrol B, Marco S, et al. (2009) Involvement of HFq
934 protein in the post-transcriptional regulation of *E. coli* bacterial cytoskeleton and cell
935 division proteins. *Cell Cycle* 8: 2470-2472.
936 88. Hazen RM, Griffin PL, Carothers JM, Szostak JW (2007) Functional information and the
937 emergence of biocomplexity. *Proc Natl Acad Sci U S A* 104 Suppl 1: 8574-8581.
938 89. Szostak JW (2003) Functional information: Molecular messages. *Nature* 423: 689.
939 90. Gordo I, Campos PR (2013) Evolution of clonal populations approaching a fitness peak. *Biol*
940 *Lett* 9: 20120239.
941 91. de Visser JA, Lenski RE (2002) Long-term experimental evolution in *Escherichia coli*. XI.
942 Rejection of non-transitive interactions as cause of declining rate of adaptation. *BMC*
943 *Evol Biol* 2: 19.
944 92. Maharjan RP, Liu B, Li Y, Reeves PR, Wang L, et al. (2013) Mutation accumulation and
945 fitness in mutator subpopulations of *Escherichia coli*. *Biol Lett* 9: 20120961.
946 93. Takeuchi N, Hogeweg P (2008) Evolution of complexity in RNA-like replicator systems.
947 *Biol Direct* 3: 11.
948 94. Blundell JR, Schwartz K, Francois D, Fisher DS, Sherlock G, et al. (2018) The dynamics of
949 adaptive genetic diversity during the early stages of clonal evolution. *Nat Ecol Evol*
950 10.1038/s41559-018-0758-1.

- 951 95. Schwartz K, Wenger JW, Dunn B, Sherlock G (2012) APJ1 and GRE3 homologs work in
952 concert to allow growth in xylose in a natural *Saccharomyces sensu stricto* hybrid yeast.
953 *Genetics* 191: 621-632.
- 954 96. Kryazhimskiy S, Rice DP, Jerison ER, Desai MM (2014) Microbial evolution. Global
955 epistasis makes adaptation predictable despite sequence-level stochasticity. *Science* 344:
956 1519-1522.
- 957 97. Blount ZD, Barrick JE, Davidson CJ, Lenski RE (2012) Genomic analysis of a key
958 innovation in an experimental *Escherichia coli* population. *Nature* 489: 513-518.
- 959 98. Swofford DL (2002) PAUP*. Phylogenetic Analysis Using Parsimony (*and Other
960 Methods). 4 ed. Sunderland, Massachusetts: Sinauer Associates.
- 961 99. Letunic I, Bork P (2007) Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree
962 display and annotation. *Bioinformatics* 23: 127-128.
- 963

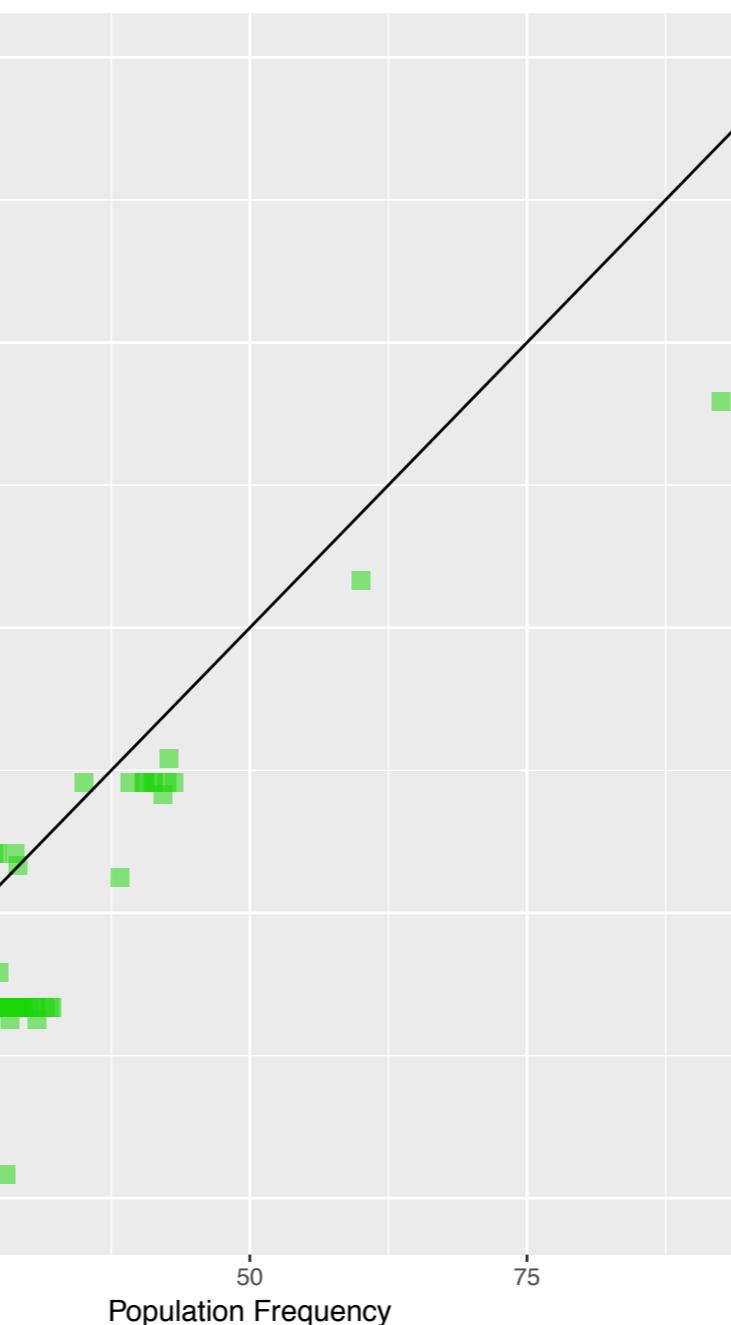




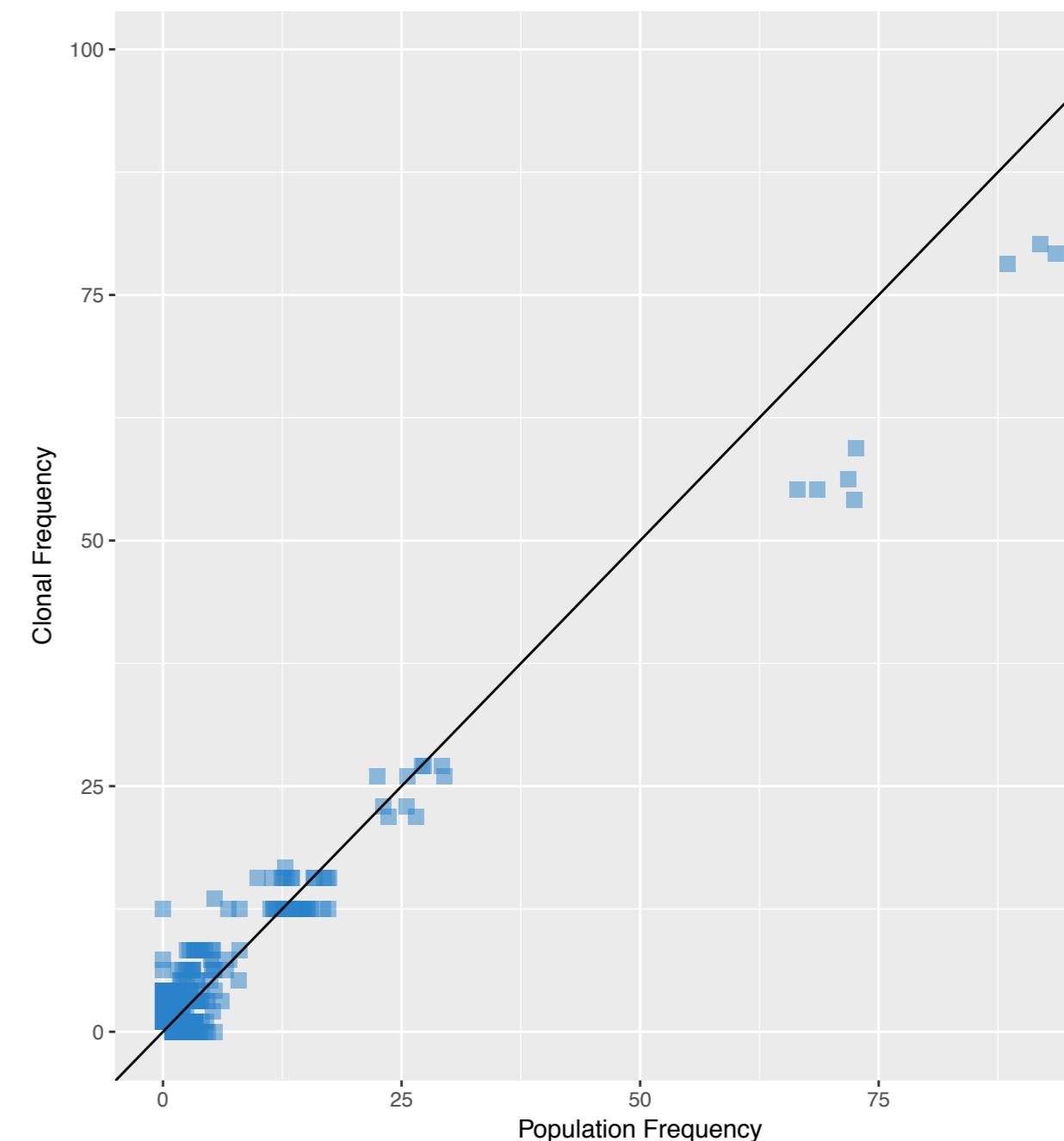
Chemostat 1

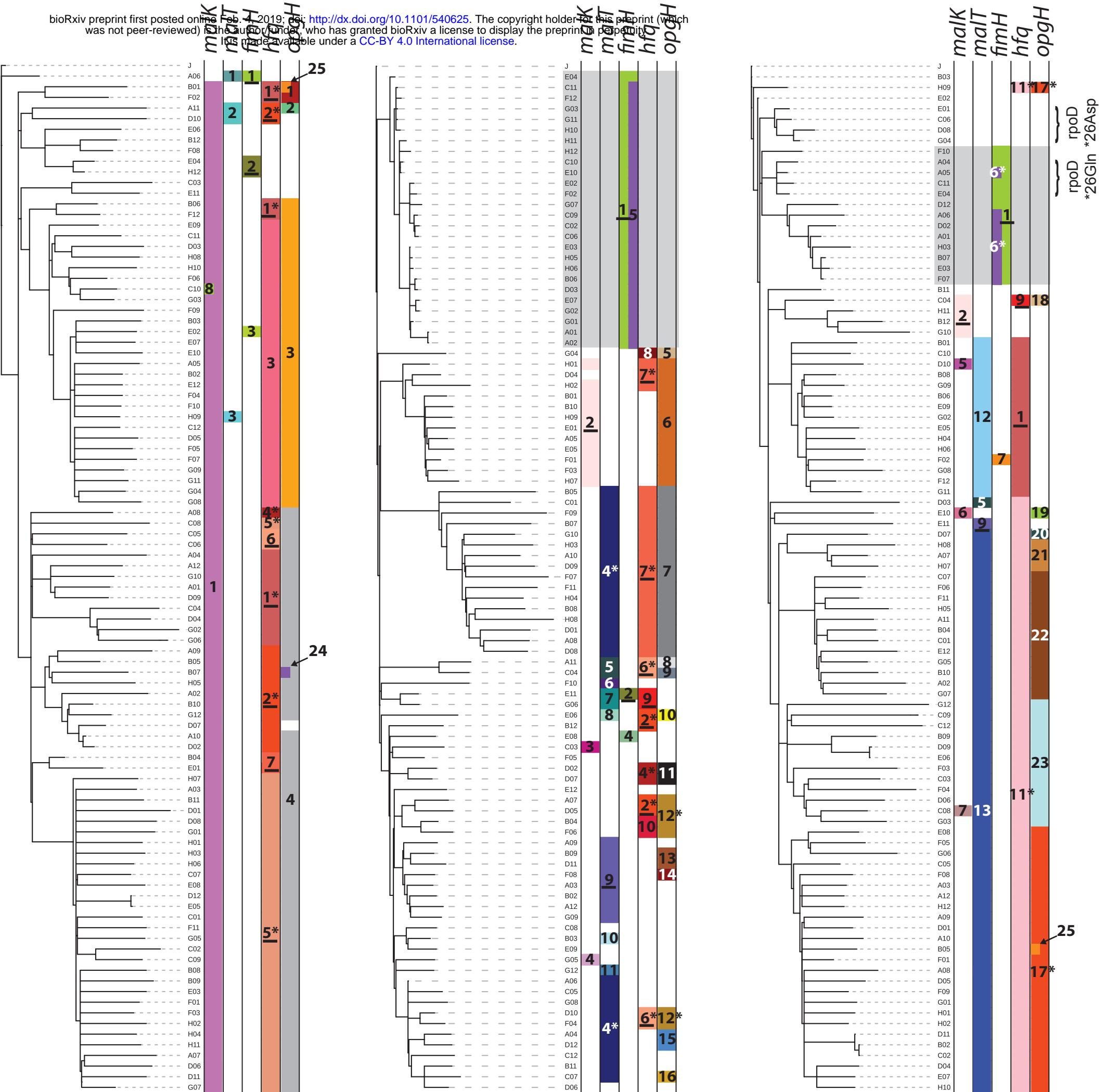


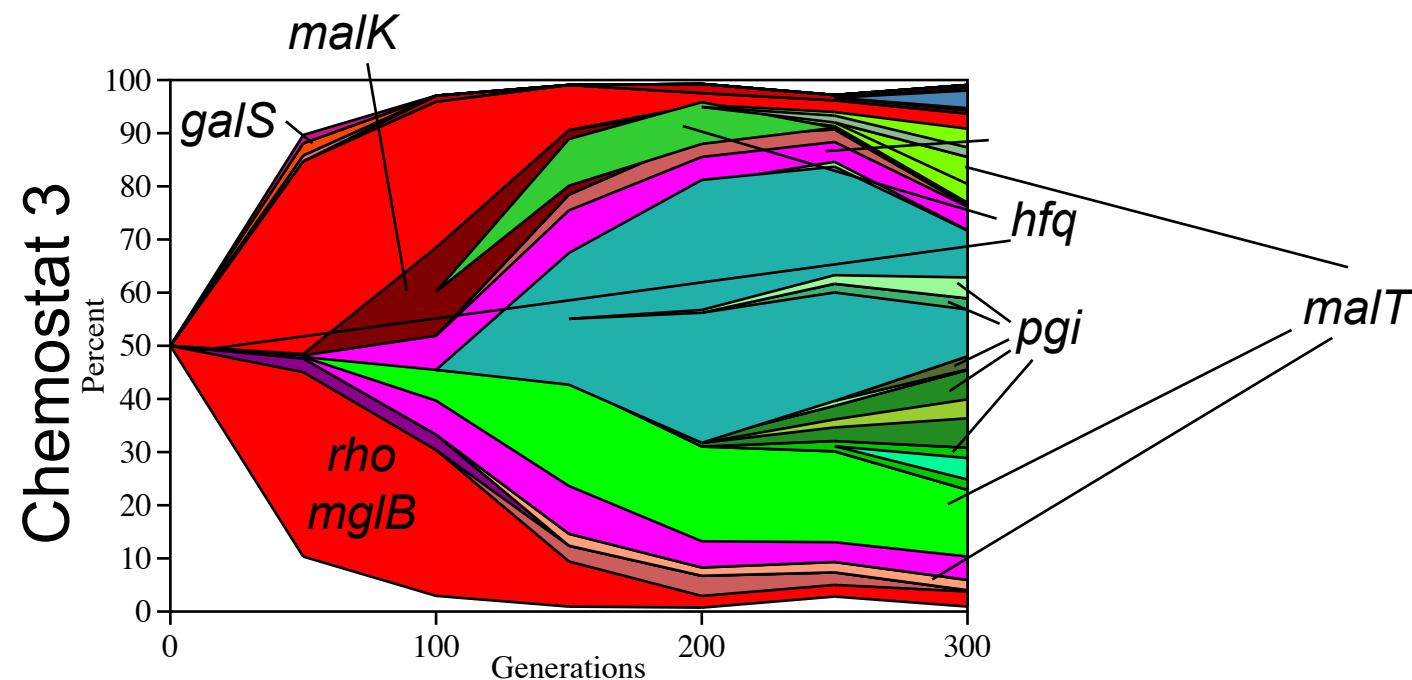
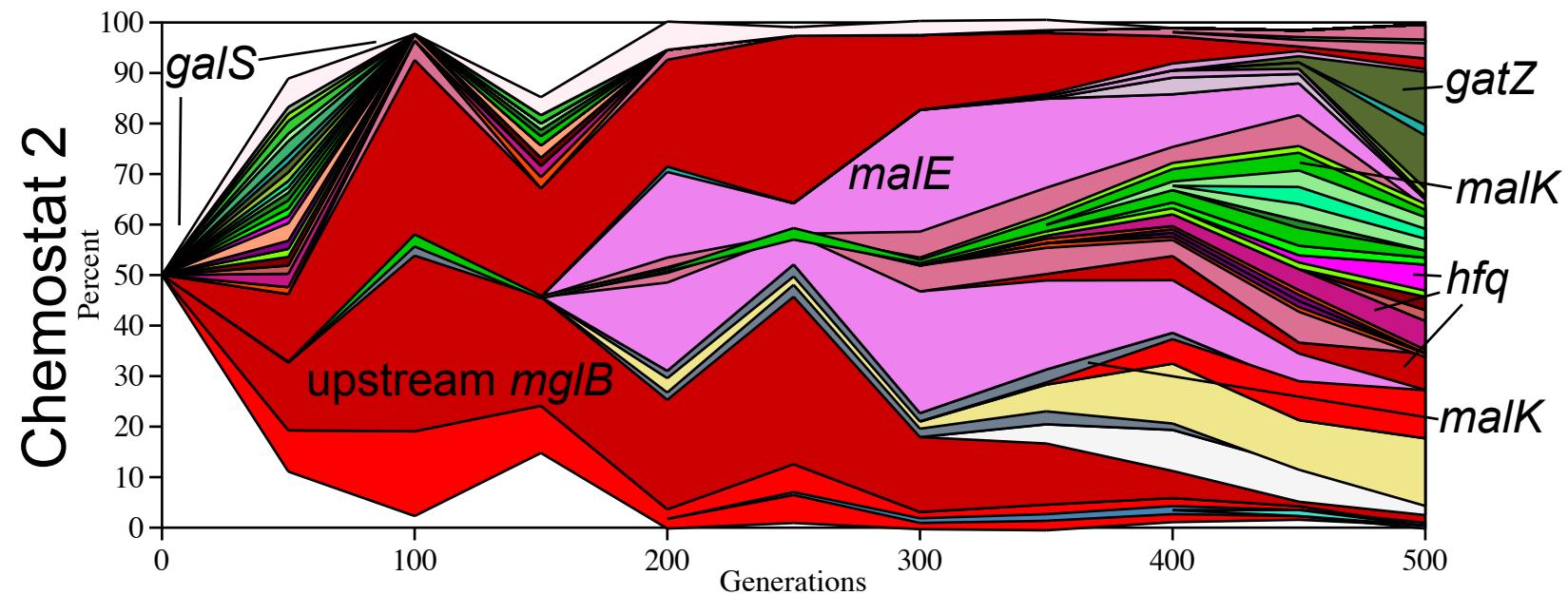
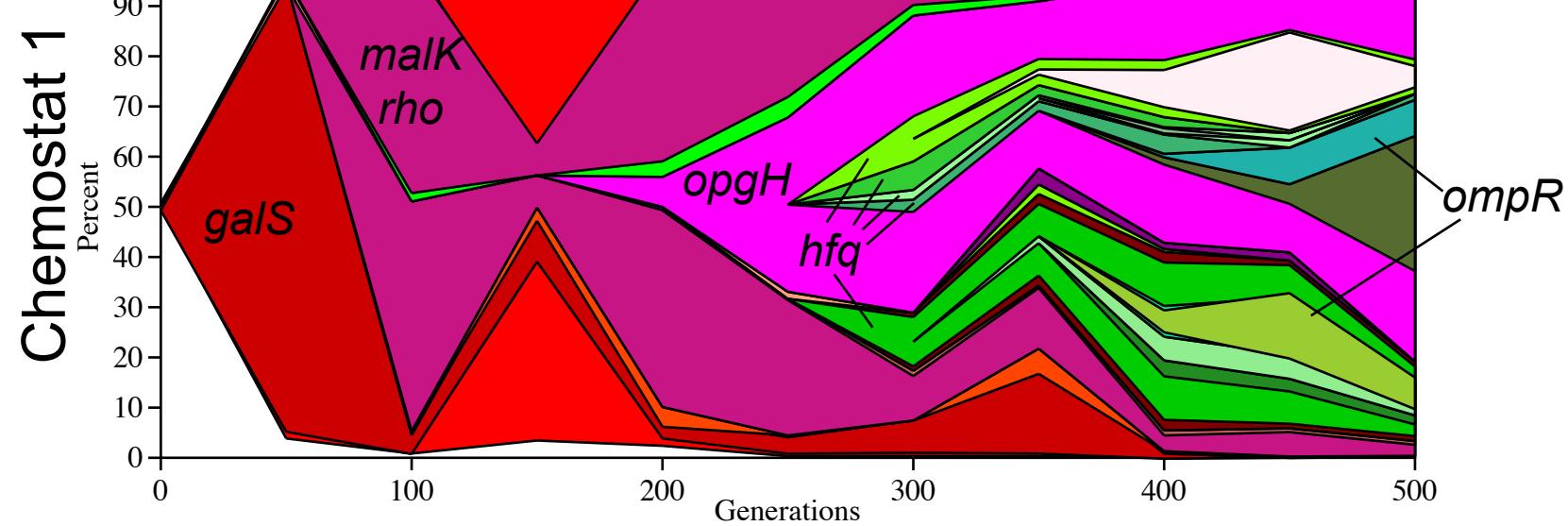
Chemostat 2



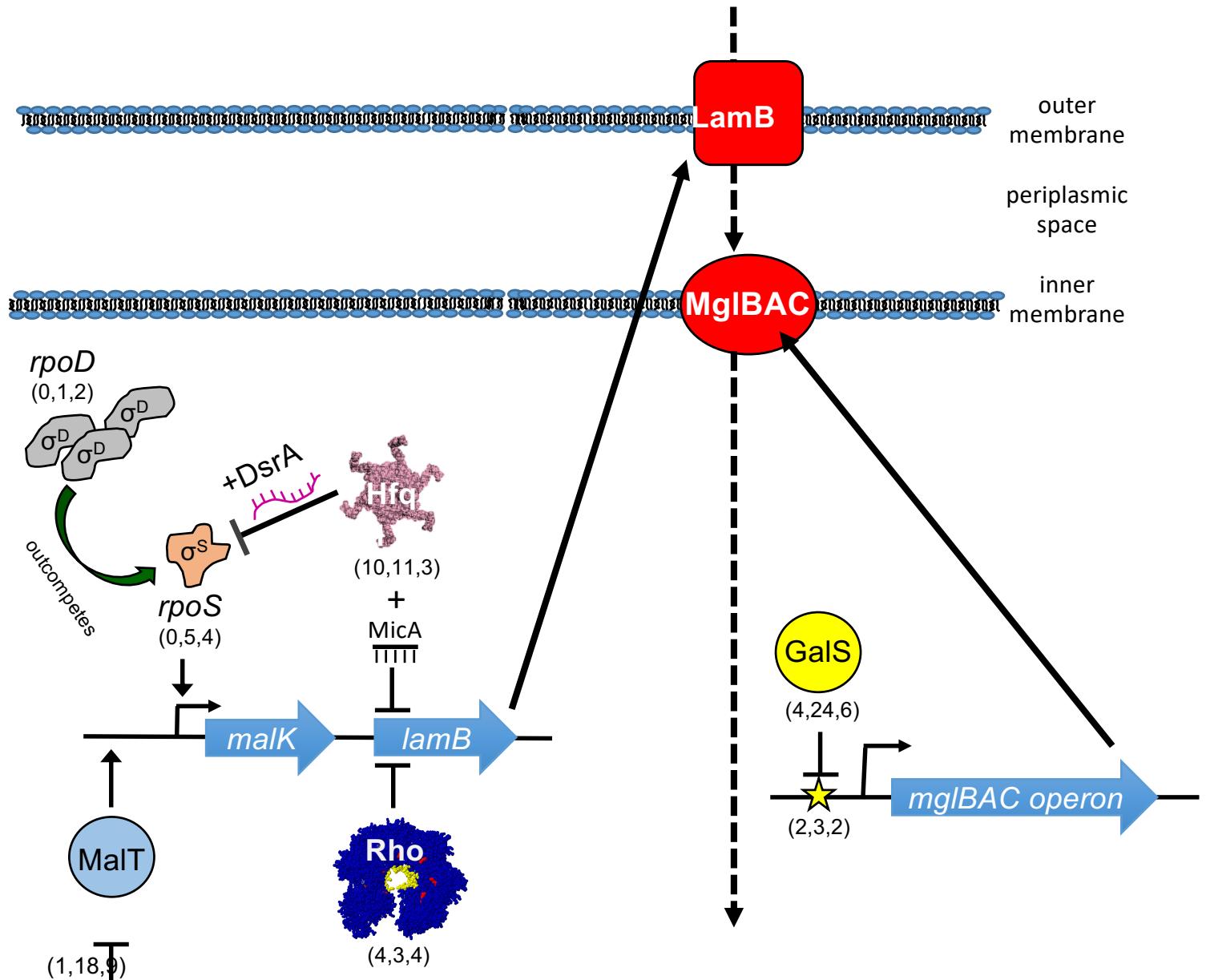
Chemostat 3







Glucose



Glucose-6-P

