

1 **The fitness effects of spontaneous mutations nearly unseen by selection in a**  
2 **bacterium with multiple chromosomes**

3  
4 Marcus M. Dillon\*, and Vaughn S. Cooper<sup>§</sup>

5  
6 \*Microbiology Graduate Program, University of New Hampshire, Durham, NH, 03824

7  
8 <sup>§</sup>Department of Microbiology and Molecular Genetics, University of Pittsburgh School of  
9 Medicine, Pittsburgh, PA, 15219

10 **Running Title:** Effects of spontaneous mutations

11

12 **Keywords:** fitness effects, deleterious mutation, mutation accumulation, genetic drift,  
13 *Burkholderia cenocepacia*

14

15 **Corresponding Author:**

16 Vaughn Cooper

17 425 Bridgeside Point II, 450 Technology Drive

18 Pittsburgh, PA 15219

19 Phone: (412)-624-1265

20 Email: [vaughn.cooper@pitt.edu](mailto:vaughn.cooper@pitt.edu)

21

22 **Author Contributions:** M.D., and V.C. designed the research; M.D. performed the  
23 research; M.D. analyzed the data; and M.D., and V.C. wrote the paper.

24

## ABSTRACT

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

Mutation accumulation (MA) experiments employ the strategy of minimizing the population size of evolving lineages to greatly reduce effects of selection on newly arising mutations. Thus, most mutations fix within MA lines independently of their fitness effects. This approach, more recently combined with genome sequencing, has detailed the rates, spectra, and biases of different mutational processes. However, a quantitative understanding of the fitness effects of mutations virtually unseen by selection has remained an untapped opportunity. Here, we analyzed the fitness of 43 sequenced MA lines of the multi-chromosome bacterium *Burkholderia cenocepacia* that had each undergone 5554 generations of MA and accumulated an average of 6.73 spontaneous mutations. Most lineages exhibited either neutral or deleterious fitness in three different environments in comparison with their common ancestor. The only mutational class that was significantly overrepresented in lineages with reduced fitness was the loss of the plasmid, though nonsense mutations, missense mutations, and coding insertion-deletion mutations were also overrepresented in MA lineages whose fitness had significantly declined. Although the overall distribution of fitness effects was similar between the three environments, the magnitude and even the sign of the fitness of a number of lineages changed with the environment, demonstrating that the fitness of some genotypes was environmentally dependent. These results present an unprecedented picture of the fitness effects of spontaneous mutations in a bacterium with multiple chromosomes and provide greater quantitative support of the theory that the vast majority of spontaneous mutations are neutral or deleterious.

46

## INTRODUCTION

47       The extent to which spontaneous mutations contribute to evolutionary change  
48 largely depends on their rates and fitness effects. Both parameters are fundamental to  
49 several evolutionary problems, including the preservation of genetic variation  
50 (Charlesworth *et al.* 1993, 2009; Charlesworth and Charlesworth 1998), the evolution of  
51 recombination (Muller 1964; Kondrashov 1988; Otto and Lenormand 2002; Roze and  
52 Blanckaert 2014), the evolution of mutator alleles (Sniegowski *et al.* 1997; Tenaillon *et*  
53 *al.* 1999), and deleterious mutation accumulation in small populations (Lande 1994;  
54 Lynch *et al.* 1995, 1999; Schwander and Crespi 2009). Many studies have now  
55 obtained direct and robust estimates of mutation rates and spectra across diverse  
56 organisms, but our understanding of the fitness effects of spontaneous mutations  
57 remains limited to mostly indirect estimates in classic model organisms (Eyre-Walker  
58 and Keightley 2007).

59       Mutation accumulation (MA) experiments provide the opportunity to quantify  
60 properties of the fitness of spontaneous mutations that have not been exposed to the  
61 sieve of natural selection. Specifically, MA experiments limit the efficiency of natural  
62 selection by passaging replicate lineages through repeated single cell bottlenecks.  
63 These lineages accumulate mutations independently over several thousand  
64 generations, and the magnitude and variance in fitness between lineages can be used  
65 to estimate several properties of the distribution of fitness effects (Halligan and  
66 Keightley 2009). MA studies have been used to characterize the fitness effects of  
67 spontaneous mutations in *Drosophila melanogaster* (Bateman 1959; Mukai 1964;  
68 Keightley 1994; Fry *et al.* 1999), *Arabidopsis thaliana* (Schultz *et al.* 1999; Shaw *et al.*

69 2000, 2002), *Caenorhabditis elegans* (Keightley and Caballero 1997; Vassilieva *et al.*  
70 2000; Estes *et al.* 2004; Katju *et al.* 2015), *Saccharomyces cerevisiae* (Wloch *et al.*  
71 2001; Zeyl and de Visser 2001; Dickinson 2008; Jasmin and Lenormand 2015),  
72 *Escherichia coli* (Kibota and Lynch 1996; Trindade *et al.* 2010), and other microbes  
73 (Heilbron *et al.* 2014; Kraemer *et al.* 2015). Results from these studies have  
74 occasionally been inconsistent, but the majority of results suggest that most  
75 spontaneous mutations have mild effects (Eyre-Walker and Keightley 2007; Halligan  
76 and Keightley 2009; Agrawal and Whitlock 2012; Heilbron *et al.* 2014), that deleterious  
77 mutations far outnumber beneficial mutations (Keightley and Lynch 2003; Silander *et al.*  
78 2007; Eyre-Walker and Keightley 2007), and that the distribution of effects of  
79 deleterious mutations is complex and multimodal (Zeyl and de Visser 2001; Eyre-  
80 Walker and Keightley 2007).

81 A more powerful approach for studying the fitness effects of spontaneous  
82 mutations is to pair MA experiments with whole-genome sequencing (MA-WGS), so that  
83 both the genetic basis and fitness effects of a collection of mutations can be known.  
84 MA-WGS studies have been conducted in a diverse array of bacteria, generating a  
85 growing database of naturally accumulated mutations that has dramatically improved  
86 estimates of mutation rates and spectra (Lee *et al.* 2012; Sung *et al.* 2012, 2015;  
87 Heilbron *et al.* 2014; Long *et al.* 2014, 2015; Dillon *et al.* 2015; Foster *et al.* 2015;  
88 Dettman *et al.* 2016). Yet, only one of these studies has also characterized the fitness of  
89 MA-WGS lines (Heilbron *et al.* 2014), and this study was conducted with mutator  
90 lineages, which have altered base-substitution and indel biases (Lee *et al.* 2012; Sung  
91 *et al.* 2015), and produce hundreds of mutations per line. Our understanding of the

92 fitness effects of spontaneous mutations would benefit greatly from more direct  
93 estimates of fitness derived from MA lineages that harbor fewer known mutations.

94 Here, we measured the relative fitness of 43 fully sequenced MA lineages  
95 derived from *B. cenocepacia* HI2424 in three laboratory environments after they had  
96 been evolved in the near absence of natural selection for 5554 generations. Following  
97 the MA experiment, each lineage harbored a total mutational load of 2 to 14  
98 spontaneous mutations, including base-substitution mutations (bpsms), insertion-  
99 deletion mutations (indels), and whole-plasmid deletions. By correlating the relative  
100 fitness of these MA lineages with the particular mutations that they harbor, we present a  
101 detailed picture of the fitness effects of spontaneous mutations, and precise estimates  
102 of deleterious mutation rates and fitness effects in a clinically important gamma-  
103 proteobacterium with multiple chromosomes.

104

105

## MATERIALS AND METHODS

106 **Bacterial strains and culture conditions.** All mutation accumulation experiments were  
107 founded from a single colony of *Burkholderia cenocepacia* HI2424, which was isolated  
108 from the soil and only passaged in the laboratory during isolation (Coenye and LiPuma  
109 2003). As a member of the diverse *B. cepacia* complex, *Burkholderia cenocepacia* can  
110 form highly resistant biofilms and has been associated with persistent lung infections in  
111 patients with cystic fibrosis (Mahenthiralingam *et al.* 2005; Traverse *et al.* 2013). The  
112 genome of *B. cenocepacia* HI2424 has been fully sequenced and is composed of three  
113 chromosomes (Chr1: 3.48-Mb, 3253 genes; Chr2: 3.00-Mb, 2709 genes; Chr3: 1.06-  
114 Mb, 929 genes) and a plasmid (0.164-Mb, 159 genes), though the third chromosome

115 can be eliminated under some conditions in some strains (Agnoli *et al.* 2012). To  
116 facilitate relative fitness assays, we competed all *B. cenocepacia* HI2424 strains derived  
117 from the MA experiment with a *B. cenocepacia* HI2424 Lac<sup>+</sup> strain, which is isogenic to  
118 *B. cenocepacia* HI2424, except for the introduction of the *lacZ* gene at the *attTn7* site,  
119 which causes colonies to turn blue when exposed to 5-bromo-4-chloro-indolyl- $\beta$ -  
120 galactopyranoside (X-gal) (Choi *et al.* 2005).

121 MA experiments were conducted on tryptic soy agar plates (TSA) (30 g/liter  
122 tryptic soy broth powder, 15 g/liter agar) and were incubated at 37°. At the conclusion of  
123 the MA experiment, frozen stocks were prepared by growing a single colony from each  
124 lineage overnight in 5ml of tryptic soy broth (TSOY) (30 g/liter tryptic soy broth powder)  
125 at 37° and freezing at -80° in 8% DMSO. All relative fitness assays were conducted in  
126 18 x 150mm glass capped tubes with 5ml of liquid medium and were maintained at 37°  
127 in a roller drum (30 rpm). Relative fitness of each lineage was assayed in three different  
128 environments. First, we conducted relative fitness assays in TSOY, a medium that  
129 mimics the conditions of the MA experiment and is expected to be very permissive.  
130 Second, we conducted relative fitness assays in M9 Minimal Medium supplemented  
131 with 0.3% casamino acids (M9MM+CAA) (3 g/liter casamino acid powder, 1 g/liter  
132 glucose, 6 g/liter sodium phosphate dibasic anhydrous, 3 g/liter potassium phosphate  
133 monobasic, 1 g/liter ammonium chloride, 0.5 g/liter sodium chloride, 0.1204 g/liter  
134 magnesium sulfate, 0.0147 g/liter calcium chloride), a medium that is more nutrient  
135 restrictive than TSOY, but contains all essential amino acids except tryptophan. Lastly,  
136 we conducted relative fitness assays in M9 Minimal Medium (M9MM) (1 g/liter glucose,  
137 6 g/liter sodium phosphate dibasic anhydrous, 3 g/liter potassium phosphate

138 monobasic, 1 g/liter ammonium chloride, 0.5 g/liter sodium chloride, 0.1204 g/liter  
139 magnesium sulfate, 0.0147 g/liter calcium chloride), which is a fully defined medium that  
140 is more restrictive than either TSOY or M9MM+CAA. Serial passaging during fitness  
141 assays was performed using 100-fold dilutions, so all relative fitness assays were  
142 conducted over the same number of generations, despite the moderate different  
143 carrying capacities in these mediums. All dilutions were performed using phosphate  
144 buffer saline (PBS) (80 g/liter NaCl, 2 g/liter KCl, 14.4 g/liter Na<sub>2</sub>HPO<sub>4</sub> • 2H<sub>2</sub>O, 2.4  
145 g/liter KH<sub>2</sub>PO<sub>4</sub>) in 96-well plates.

146  
147 **MA-WGS process.** The mutation accumulation experiment that generated the  
148 mutational load for this study has been reported previously (Dillon *et al.* 2015). Briefly,  
149 seventy-five independent lineages were founded from a single colony of *B. cenocepacia*  
150 HI2424 and independently propagated every 24 hours onto a fresh TSA plate for 217  
151 days. Daily generations were estimated monthly by taking a single representative  
152 colony from each lineage following 24 hours of growth, placing it in 2 ml of PBS, then  
153 serially diluting and spread plating it on TSA. The number of viable cells in each colony  
154 was then used to calculate the number of generations elapsed between each transfer,  
155 and the average number of generations across all lineages was used as the number of  
156 generations per day for that entire month. By multiplying the number of generations per  
157 day for each month by the number of days in that month, then summing these totals  
158 over the course of the whole experiment, we calculated the total number of generations  
159 elapsed per MA lineage over the course of the MA-experiment.



160 As we have described previously, genomic DNA was extracted from 1 ml of  
161 overnight TSB culture founded by 47 of the *B. cenocepacia* isolates that were stored at  
162 the conclusion of the MA experiment (Dillon *et al.* 2015). We used the Wizard Genomic  
163 DNA Purification kit for DNA extraction (Promega), all libraries were prepared using a  
164 modified Illumina Nextera protocol (Baym *et al.* 2015), and sequencing was performed  
165 with the 151-bp paired end platform on the Illumina HiSeq at the Hubbard Center for  
166 Genomic Studies at the University of New Hampshire. Following fastQC analysis, all  
167 reads were mapped to the *B. cenocepacia* HI2424 reference genome (LiPuma *et al.*  
168 2002) with both the Burrows-Wheeler aligner (BWA) (Li and Durbin 2009) and  
169 Novoalign ([www.novocraft.com](http://www.novocraft.com)). The average depth of coverage across all 47 lines was  
170 43x, but the average depth of coverage in the 43 lines used in this study was 46x.

171  
172 **Spontaneous mutation identification.** All bpsms were identified as described  
173 previously (Dillon *et al.* 2015). Briefly, after using a combination of SAMtools and in  
174 house perl scripts to produce all read alignments for each position in each line (Li *et al.*  
175 2009), a three step process was used to detect putative bpsms. First, pooled reads  
176 across all lines were used to generate an ancestral consensus base at each site in the  
177 reference genome, allowing us to correct differences between the published reference  
178 genome and the ancestral colony of our MA experiment. Second, reads from the  
179 individual lines were used to generate a lineage specific consensus base at each site in  
180 the reference genome for each lineage, as long as the site was covered by at least two  
181 forward and two reverse reads, and at least 80% of the reads identified the same base.  
182 Sites that did not meet these criteria were not analyzed in the respective lineage. Third,

183 lineage specific consensus bases for each lineage were compared to the ancestral  
184 consensus base at each site, and a putative bpsm was identified if they differed. This  
185 three-step process was carried out independently using both the BWA and Novoalign  
186 alignments, and putative bpsms were considered genuine only if both pipelines  
187 independently identified the bpsm. Despite these lenient criteria, all of the bpsms that  
188 were identified at analyzed sites in this study had considerably greater coverage and  
189 consensus than the minimal criteria (see File S1), demonstrating that these bpsms are  
190 not merely false positives in low coverage regions. The frequency of sites that were not  
191 analyzed in each lineage varied from 0.005 to 0.177. Putative bpsms in these regions  
192 were estimated by multiplying the number of unanalyzed sites in each lineage by the  
193 overall bpsm rate calculated in this study ( $1.31 (0.08) \cdot 10^{-10}$  /bp/generation) and the  
194 number of generations of mutation accumulation in each lineage (5554) (see Table S1).

195 Indels and large structural variants are inherently more difficult to identify than  
196 bpsms because gaps and simple sequence repeats (SSRs) reduce the accuracy of  
197 short-read alignment algorithms. To overcome these issues, we extracted all putative  
198 indels if at least 30% of the reads that covered the site identified the exact same indel  
199 (size and motif), and the site was covered by at least two forward and two reverse reads  
200 (Dillon *et al.* 2015). These putative indels were then subject to a series of more  
201 strenuous filters based on the complexity of the region in which they were identified and  
202 the consensus between the BWA and Novoalign alignments (see File S1). Specifically,  
203 all putative indels where more than 80% of the reads identified the exact same indel in  
204 both the BWA and Novoalign alignments were considered genuine indels. For putative  
205 indels where only 30-80% of the reads identified the exact same indel, we parsed out

206 only reads that had bases covering both the upstream and downstream region of the  
207 indel (if it was not in an SSR), and both the upstream and downstream region of the  
208 SSR (if it was in an SSR). Using this subset of reads, we reassessed the frequency of  
209 reads that identified the exact same indel, allowing us to more accurately identify indels  
210 involving the gain or loss of a single repeat within a SSR. These indels were considered  
211 genuine if more than 80% of the parsed reads identified the exact same indel and were  
212 discarded if they did not. Putative indels and other structural variants were also  
213 extracted using PINDEL, which uses paired-end information to identify insertions,  
214 deletions, inversions, tandem duplications and other structural variants (Ye *et al.* 2009).  
215 Here, indels were considered genuine if they were covered by at least six forward and  
216 six reverse reads, and at least 80% of the reads identified the exact same indel. Lastly,  
217 we analyzed the distribution of coverage between chromosomes and the 0.164-Mb  
218 plasmid to detect any chromosomal copy number variants. As with bpsms, putative  
219 indels in regions that were not analyzed were estimated by multiplying the number of  
220 unanalyzed sites in each lineage by the overall indel rate calculated in this study ( $2.39$   
221  $(0.34) \cdot 10^{-11}$  /bp/generation) and the number of generations of mutation accumulation in  
222 each lineage (5554) (see Table S1).

223

224 **Quantifying relative fitness.** To quantify the selection coefficients of each of the 43  
225 derived MA lineages, we conducted three-day competitions (over  $\approx 19.93$  generations)  
226 between each MA lineage and our *B. cenocepacia* HI2424 Lac<sup>+</sup> strain. These  
227 competitions were carried out independently in TSOY, M9MM+CAA, and M9MM, with  
228 four replicates being conducted for each lineage in each environment. Our MA ancestral

229 *B. cenocepacia* HI2424 strain was also competed against *B. cenocepacia* HI2424 Lac+  
230 as a control, with four replicates for each environment. Selection coefficients were  
231 estimated as described previously, using the relative growth of the focal MA lineage and  
232 the *B. cenocepacia* HI2424 Lac+ reference strain, normalized by the number of  
233 generations elapsed by the reference strain ( $G$ ) (Chevin 2011; Perfeito *et al.* 2014).  
234 First, the difference in growth ( $\Delta r_{ab}$ ) between the two strains was estimated as:

$$\Delta r_{ab} = \ln \left( \frac{N_{fa}}{N_{ia}} \right) - \ln \left( \frac{N_{fb}}{N_{ib}} \right)$$

235 where  $N_{ia}$  and  $N_{ib}$  were the initial numbers of test and reference bacteria, respectively,  
236 and  $N_{fa}$  and  $N_{fb}$  were the final numbers of test and reference bacteria, respectively.  
237 Selection coefficients ( $s_{ab}$ ) were then calculated as:

$$238 \quad s_{ab} = \Delta r_{ab} / G$$

239 where  $G$ , generations elapsed by the reference strain, is equal to:

$$G = \log_2 \left( \frac{N_{fb}}{N_{ib}} \right)$$

240 For each replicate, all 43 derived MA lineages, *B. cenocepacia* HI2424, and *B.*  
241 *cenocepacia* Lac+ were resurrected from frozen culture by inoculating them into 5 ml of  
242 TSOY broth and incubating overnight in a roller drum at 30 rpm. Depending on which  
243 environment was being assayed, each strain was then transferred to fresh TSOY,  
244 M9MM+CAA, or M9MM via a 10,000-fold dilution and acclimated for 24 hours at 37°  
245 and 30 rpm. Following acclimation, 44 competitions (43 MA lineages + control) were  
246 generated in the appropriate fresh medium at a 1:1 ratio via 100-fold dilution, and 30  $\mu$ l  
247 from each was extracted to quantify the initial frequency of each competitor ( $N_{ia}$ ,  $N_{ib}$ ).  
248 Competitions were then incubated for 72 hours at 37° and 30 rpm, being transferred to

249 fresh media every 24 hours via a 100-fold dilution. At the conclusion of the 72 hour  
250 competition, 30  $\mu$ l of the final culture was extracted to quantify the final frequency of  
251 each competitor ( $N_{fa}$ ,  $N_{fb}$ ).

252 To measure the initial frequency of each competitor, the extracted culture was  
253 diluted in PBS and 100  $\mu$ l of the diluted sample was plated on a TSA + X-Gal plate.  
254 Specifically, in the TSOY competitions the samples were diluted 30,000-fold, in the  
255 M9MM+CAA competitions the samples were diluted to 20,000-fold, and in the M9MM  
256 competitions the samples were diluted to 10,000-fold. Following a 48-hr incubation, the  
257 number of white and blue colonies were quantified and used to calculate  $N_{ia}$  and  $N_{ib}$ ,  
258 respectively, after accounting for the dilutions. Final frequencies were measured in the  
259 same way, except that an additional 100-fold dilution was required for each competition  
260 because the cultures were at carrying capacity. In addition, to calculate  $N_{fa}$  and  $N_{fb}$ , we  
261 had to account for the dilutions that were conducted prior to plating and the two 100-fold  
262 dilutions that were conducted during the three-day competition. Importantly, the  
263 selection coefficient of the *B. cenocepacia* HI2424 MA ancestor was not significantly  
264 different from 0 in any of environments (TSOY:  $s = -0.0002$  (0.0020), M9MM+CAA:  $s =$   
265  $+0.0075$  (0.0030), M9MM:  $s = -0.0016$  (0.0034) (SEM)).

266  
267 **Statistical analysis.** All statistical analyses were performed in R Version 0.98.1091  
268 using the Stats analysis package (R Development Core Team 2013). For independent  
269 two-tailed t-tests, all p-values were corrected for multiple comparisons using a  
270 Benjamini-Hochberg correction (Table S2), which ensures that our false positive rate  
271 remains below 5%, despite testing whether the selection coefficient differed significantly

272 from 0 for 43 lineages in each environment (Benjamini and Hochberg 1995). Corrected  
273 p-values that were below a threshold of 0.05 were considered significant. Linear  
274 regressions were used to evaluate the correlation between the number of mutations in a  
275 lineage and its selection coefficient, as well as the correlation between the selection  
276 coefficients of lineages in different environments. Lastly, to test for effects of replicate,  
277 genotype, environment, and genotype\*environment interaction on the fitness of each  
278 lineage, we performed an analysis of variance (ANOVA) on the cumulative dataset.

279

280 **Data Availability.** Illumina DNA sequences for the *B. cenocepacia* MA lines used in this  
281 study are available under the PRJNA326274 bioproject. File S1 contains all  
282 supplementary tables and figures referenced in text. File S2 contains all the information  
283 on the location and criteria used to identify each base-substitution and insertion-deletion  
284 mutation identified in this study. All strains are available upon request.

285

286

## RESULTS

287 We previously reported the rate and molecular spectrum of spontaneous  
288 mutations in wild-type *B. cenocepacia*, as determined from the cumulative results of a  
289 MA-WGS experiment involving 47 replicate lineages derived from *B. cenocepacia*  
290 HI2424 (Dillon *et al.* 2015). Each lineage was passaged through daily single-cell  
291 bottlenecks for 217 days, resulting in approximately 5554 generations of MA per  
292 lineage. The average number of generations of growth per day within a colony declined  
293 from 26.16 (0.06) to 24.92 (0.07) (SEM) over the course of the experiment, suggesting  
294 that some of the accumulated mutations had deleterious fitness effects. Here, we

295 present a detailed picture of the fitness effects of the spontaneous mutations  
296 accumulated at the conclusion of this MA-WGS experiment using 43 replicate lineages,  
297 as the remaining four lineages were discarded because of a lack of sufficient coverage  
298 in the WGS data.

299         The properties of the mutations found in these 43 lineages are consistent with  
300 constant mutation rates and limited selection over the course of our experiment. Neither  
301 the distribution of bpsms or indels across lineages differed significantly from a Poisson  
302 distribution (bpsms:  $\chi^2 = 3.46$ ,  $p = 0.94$ ; indels:  $\chi^2 = 0.28$ ,  $p = 0.96$ ), signifying that  
303 mutation rates did not vary across the lineages. The ratio of synonymous to  
304 nonsynonymous bpsms also did not differ from the expected ratio based on the codon-  
305 usage and %GC content at synonymous and nonsynonymous sites in *B. cenocepacia*  
306 HI2424 ( $\chi^2 = 0.78$ , d.f. = 1,  $p = 0.38$ ), which suggests minimal purifying selection.  
307 Further, the lack of genetic parallelism in the bpsm spectra across lineages (see File  
308 S1) is inconsistent with positive selection acting on these lines. Although both bpsms  
309 and indels were observed more frequently than expected in non-coding DNA (bpsms:  $\chi^2$   
310 = 2.19, d.f. = 1,  $p = 0.14$ ; indels:  $\chi^2 = 45.816$ , d.f. = 1,  $p < 0.0001$ ), a pattern consistent  
311 with purifying selection, this pattern could be generated by selection against coding  
312 mutations, preferential mismatch repair in coding regions, or the mutation prone nature  
313 of repetitive DNA in non-coding regions (Lee *et al.* 2012; Heilbron *et al.* 2014; Dillon *et*  
314 *al.* 2015). In any event, we estimate that the threshold selection coefficient below which  
315 genetic drift will overpower natural selection, as determined by  $N_E \times s = 1$  in haploid  
316 organisms, is 0.08 (Dillon *et al.* 2015). Thus, while a small class of adaptive or  
317 deleterious mutations with effects in excess of  $s = 0.08$  will be subject to the biases of

318 natural selection (Kimura 1983; Elena *et al.* 1998; Zeyl and de Visser 2001; Hall *et al.*  
319 2008), the vast majority of mutations that were observed in this study likely fixed  
320 irrespective of their fitness effects.

321  
322 **Genetic basis of spontaneous mutations.** The spontaneous bpsms and indels  
323 reported here are similar to those reported previously (Dillon *et al.* 2015), with a few  
324 exceptions. First, we allowed for bpsms to be called in more than one lineage, resulting  
325 in the addition of two bpsms. These bpsms are assumed to have occurred in the  
326 ancestral colony, but their presence in each lineage must be documented to accurately  
327 quantify the relationship between the fitness of each lineage and its mutational load.  
328 Second, we were able to confidently identify nine additional indels that occurred in  
329 simple sequence repeats by using an approach that considers only reads anchored on  
330 both sides of the repeat (see Methods; File S1). These indels may contribute  
331 substantively to the mutational load of the lineages in which they occur, so they were  
332 important to include in this study. Lastly, we did not analyze four of the lineages from  
333 the previous study because less than 80% of their genomes had sufficient coverage to  
334 be analyzed for the presence of bpsms and indels (see Methods). This low coverage  
335 would render us blind to a considerable portion of the mutational load in these lineages,  
336 which warranted their exclusion.

337       Among the 43 MA lineages, 233 bpsms, 42 short indels, and 4 plasmid-loss  
338 events were identified. The most common class of bpsms was missense bpsms (141),  
339 followed by synonymous bpsms (49), intergenic bpsms (37), and nonsense bpsms (6).  
340 Among indels, coding indels involving only a single gene (22) were slightly more



341 common than intergenic indels (20), while loss of the 0.16-Mb plasmid, which encodes  
342 157 genes, was observed in 4 lineages (see Figure 1). False-negative rates were  
343 estimated in each lineage as the number of sites that were not analyzed for mutations,  
344 multiplied by the product of experiment-wide bpsm and indel rates per base-pair per  
345 generation and the number of generations experienced by each lineage. Given that an  
346 average of 95.20 % (0.01) of each genome was sequenced to sufficient depth to  
347 analyze both bpsms and indels, we estimate that an average of only 0.25 (0.03)  
348 additional bpsms and 0.05 (0.01) additional indels would have been identified per  
349 lineage if the entire genome was analyzed (SEM) (see Table S1). Overall, mutations  
350 were not uniformly distributed across the 43 MA lineages, allowing us to analyze which  
351 mutation types are most likely to have fitness effects.

352

353 **Fitness effects of spontaneous mutations.** Fitness of the final isolate from each  
354 lineage was measured by direct competition with the ancestral *B. cenocepacia* HI2424  
355 strain in three different broth culture conditions. TSOY broth is a very permissive  
356 medium used to mimic the conditions of the MA experiment, M9MM+CAA is an amino-  
357 acid supplemented minimal medium that is less permissive than TSOY but more  
358 permissive than a strictly minimal medium, and M9MM is a fully defined minimal  
359 medium that is the least permissive of the three environments. In TSOY, 17 lineages  
360 had significantly reduced fitness and no lineages had significantly increased fitness (see  
361 Figure 2A). The average fitness across all MA lineages in TSOY was -0.024 (0.005)  
362 (SEM), with a range of -0.111 to +0.037. Similarly, 13 lineages had significantly reduced  
363 fitness in M9MM+CAA and none had significantly increased fitness (see Figure 2B). The

364 average fitness decline and the range across all MA lineages in M9MM+CAA were  
365 similar to those observed in TSOY (Average: -0.020 (0.005) (SEM); Range: -0.116 to  
366 +0.006). Lastly, we observed 13 lineages with significantly reduced fitness in M9MM,  
367 but here, 4 other lineages had significantly increased in fitness (see Figure 2C). Thus,  
368 the average fitness decline of the MA lineages in M9MM was only -0.013 (0.005) (SEM),  
369 and the range across all lineages was shifted to the right (-0.090 to +0.026). Because  
370 each lineage harbors multiple mutations, the distributions presented in Figure 2 cannot  
371 be used directly to elucidate the distribution of effects of individual spontaneous  
372 mutations. However, it is notable that all of the distributions are significantly non-normal  
373 (Shapiro Wilk's Test; TSOY:  $W = 0.95$ ,  $p = 0.04$ ; M9MM+CAA:  $W = 0.68$ ,  $p = 2.02 \cdot 10^{-8}$ ;  
374 M9MM:  $W = 0.73$ ,  $p = 1.44 \cdot 10^{-7}$ ) and the basic properties of the distributions outlined  
375 above are similar across environments. Specifically, the majority of lineages have  
376 neutral or moderately deleterious fitness and all three distributions have an extended  
377 left tail including lineages whose fitnesses have declined more dramatically.

378         Significant positive correlations between the selection coefficients of individual  
379 MA lineages across environments suggest that effects of experiment-wide mutational  
380 load interacted with these external environments only modestly (see Figure 3).  
381 Specifically, linear regressions between the selection coefficients of each lineage in  
382 TSOY and both M9MM+CAA and M9MM produced significantly positive relationships  
383 (TSOY-M9MM+CAA:  $F = 17.18$ ,  $df = 41$ ,  $p = 0.0002$ ,  $r^2 = 0.30$ ; TSOY-M9MM:  $F = 8.61$ ,  
384  $df = 41$ ,  $p = 0.0054$ ,  $r^2 = 0.17$ ). Selection coefficients in M9MM+CAA and M9MM were  
385 also significantly correlated and explained a greater fraction of the variance than either  
386 of the TSOY regressions ( $F = 124.00$ ,  $df = 41$ ,  $p < 0.0001$ ,  $r^2 = 0.75$ ), which was

387 expected given that these environments are more similar to each other than either is to  
388 TSOY (see Figure 3). However, the fitness of certain MA lineages declined significantly  
389 in one environment but not others, suggesting that some mutations in these lineages  
390 produced environment-dependent effects. A total of nine lineages were significantly less  
391 fit in a single environment (four TSOY, two M9MM+CAA, three M9MM), eight MA  
392 lineages were significantly less fit in two of the environments, and six MA lineages were  
393 significantly less fit in all three environments. Overall, fitness was significantly influenced  
394 by replicate, genotype, environment, and genotype-by-environment interaction (see  
395 Table 1). Yet, as noted above, the general properties of the distribution of fitness effects  
396 of these MA lineages remained similar across the three tested environments (see  
397 Figure 2).

398         Despite acquiring multiple mutations, a number of MA lineages did not have  
399 significant fitness differences from the ancestral strain. Further, the number of  
400 spontaneous mutations in a line did not correlate with their absolute selection  
401 coefficients in any environment (TSOY:  $F = 1.40$ ,  $df = 41$ ,  $p = 0.2434$ ,  $r^2 = 0.03$ ;  
402 M9MM+CAA:  $F = 1.35$ ,  $df = 41$ ,  $p = 0.2513$ ,  $r^2 = 0.03$ ; M9MM:  $F = 2.96$ ,  $df = 41$ ,  $p =$   
403  $0.0930$ ,  $r^2 = 0.07$ ) (see Figure S1). After adding the 11 additional mutations presumed to  
404 have been missed in the unanalyzed genomic regions across all lines, we estimate that  
405 a total of 290 spontaneous mutations occurred in the experiment, with an average of  
406 6.73 (0.36) (SEM) mutations per lineage. In combination, these results suggest that the  
407 fitness effects of a majority of spontaneous mutations were near neutral, or at least  
408 undetectable with plate-based laboratory fitness assays. Given the average selection  
409 coefficient of each line and the number of mutations that it harbors, we can estimate

410 that the average fitness effect ( $s$ ) of a single mutation was -0.0040 (0.0008) in TSOY, -  
411 0.0031 (0.0007) in M9MM+CAA, and -0.0017 (0.0007) (SEM) in M9MM.

412 Because the fitness of many lineages with multiple mutations did not significantly  
413 differ from the ancestor, and because mutation number and fitness were not correlated,  
414 this study suggests that most of the significant losses and gains in fitness were caused  
415 by rare, single mutations with large fitness effects. Based on this conjecture, we can use  
416 the number of lineages that experienced significantly reduced fitness (17 in TSOY, 13 in  
417 M9MM+CAA, and 13 in M9MM) divided by the total generations of MA across all lines to  
418 estimate the deleterious mutation rate per genome per generation ( $U_D$ ) in each  
419 environment. Estimates of  $U_D$  based on this approach are  $7.12 \times 10^{-5}$   
420 /genome/generation in TSOY,  $5.44 \times 10^{-5}$  /genome/generation in M9MM+CAA, and  
421  $5.44 \times 10^{-5}$  /genome/generation in M9MM. The fitness effects of these single  
422 deleterious mutations ( $s_D$ ) in some lineages can also be estimated directly from the  
423 fitness of these select lineages relative to the *B. cenocepacia* HI2424 ancestor, and  
424 suggest an average  $s_D$  of -0.048 (0.007) in TSOY, -0.053 (0.011) in M9MM+CAA, and -  
425 0.048 (0.009) in M9MM (SEM). Although we observe no significantly beneficial  
426 mutations in TSOY or M9MM+CAA, our data suggest that the beneficial mutation rate  
427 ( $U_B$ ) is  $1.68 \times 10^{-5}$  /genome/generation and the average significantly beneficial mutation  
428 has a selection coefficient ( $s_B$ ) of 0.013 (0.005) in M9MM. The primary caveats to these  
429 estimates are that they assume that only one of the mutations in each lineage  
430 determines its fitness, that none of the lineages that are statistically indistinguishable  
431 from the ancestral strain harbor deleterious mutations, and that they do not account for  
432 epistasis between the mutations accumulated in each lineage. To attain more precise

433 estimates of  $U_D$ ,  $s_D$ ,  $U_B$ , and  $s_B$ , studying spontaneous mutations in different  
434 combinations and using more precise fitness assays based on flow-cytometry and/or  
435 barcoded sequencing will be especially valuable.

436

437 **Genetic basis of deleterious load.** Without sequencing and measuring fitness at  
438 intermediate time-points or genetically engineering *B. cenocepacia* HI2424 strains that  
439 harbor only single spontaneous mutations, it is difficult to pinpoint which mutations  
440 generate the fitness declines in our MA lineages. However, we can examine  
441 relationships between the forms of mutational load harbored by each lineage and the  
442 fitness of those lineages (Figure 1). The only mutation type that was significantly  
443 overrepresented in lineages with reduced fitness in TSOY was the loss of the 0.164-Mb  
444 plasmid ( $\chi^2 = 6.12$ , d.f. = 1,  $p = 0.0130$ ). All four lineages that lost the plasmid were  
445 significantly less fit in TSOY, with an average selection coefficient of -0.060 (0.007)  
446 (SEM). These same four lineages were also less fit in M9MM ( $\chi^2 = 9.23$ , d.f. = 1,  $p =$   
447 0.0020; mean  $s = -0.043$  (0.018)), but the deleterious effects of plasmid loss appear to  
448 be mitigated in M9MM+CAA, where only one of these lineages had significantly reduced  
449 fitness.

450 Although no other mutation types were significantly overrepresented among  
451 lineages that had reduced fitness, we note that there were more coding indels,  
452 nonsense bpsms, and missense bpsms than expected in the lineages with reduced  
453 fitness in all three environments, with the exception of coding indels in M9MM+CAA  
454 (see Table 2). In contrast, intergenic bpsms and indels appear to be evenly distributed  
455 between lineages with significantly reduced fitness and those where  $s$  was not

456 significantly different from 0, suggesting that few if any intergenic mutations from this  
457 study have deleterious fitness effects. Similarly, the synonymous bpsms observed in  
458 this study do not appear to have deleterious effects, as they are less frequent than  
459 expected in lineages with reduced fitness in TSOY and are evenly distributed between  
460 neutral and reduced fitness lineages in M9MM+CAA and M9MM (see Table 2). Overall,  
461 these results support the expectation that coding indels, nonsense bpsms, and  
462 missense bpsms are more likely to have deleterious effects than intergenic and  
463 synonymous bpsms.

464

465

## DISCUSSION

466 Nearly all prior research on the fitness effects of mutations has studied mutants  
467 that have been screened by selection, which presumably purged deleterious variants  
468 and enriched beneficial ones. MA experiments are designed to minimize effects of  
469 selection to the greatest extent possible, thus capturing most mutations independent of  
470 the biases of natural selection. Combining this MA approach with whole genome  
471 sequencing and fitness measurements provides the potential to dramatically advance  
472 our understanding of the fitness effects of spontaneous mutations in diverse organisms.  
473 Here we analyzed the fitness effects of 43 lineages of *B. cenocepacia* that underwent 7  
474 months or >5500 generations of evolution under greatly limited selection. Direct  
475 competitions with the common ancestor were conducted in three environments to  
476 quantify effects of different forms of mutational load. Despite the duration of MA, many  
477 lineages evidently suffered no fitness decline, as they had selection coefficients that did  
478 not significantly differ from  $s = 0$ . Given that each lineage accumulated between 2-14

479 mutations, the most likely explanation for these findings is that the vast majority of  
480 spontaneous mutations have minimal effects on fitness across this range of  
481 environments. Under the assumption that the significant reductions in lineage fitness  
482 were driven mostly by single deleterious mutations (Davies *et al.* 1999; Heilbron *et al.*  
483 2014), we also obtain new estimates of the deleterious mutation rate ( $U_D$ ) and the  
484 average effect of deleterious mutations ( $s_D$ ) in all three environments. These measures  
485 reveal that the general features of the distribution of fitness effects are similar between  
486 these conditions. The most consistently deleterious mutational event involved loss of  
487 the 0.164-Mb plasmid, which reduced fitness in TSOY and M9MM but not M9MM+CAA.  
488 Further, nonsense bpsms, missense bpsms, and coding indels were more likely to have  
489 contributed to the deleterious mutational load than synonymous bpsms, intergenic  
490 bpsms, and intergenic indels.

491         Although a few select studies have claimed that a substantial fraction of  
492 spontaneous mutations are beneficial under certain conditions (Shaw *et al.* 2002;  
493 Silander *et al.* 2007; Dickinson 2008), evidence from diverse sources strongly suggests  
494 that the effect of most spontaneous mutations is to reduce fitness (Kibota and Lynch  
495 1996; Keightley and Caballero 1997; Fry *et al.* 1999; Vassilieva *et al.* 2000; Wloch *et al.*  
496 2001; Zeyl and de Visser 2001; Keightley and Lynch 2003; Trindade *et al.* 2010;  
497 Heilbron *et al.* 2014). Our measurements of selection coefficients in TSOY also suggest  
498 that most spontaneous mutations are neutral or deleterious because all lineages whose  
499 fitness differs significantly from the ancestor are less fit ( $s = -0.112$  to  $-0.014$ ). The  
500 overall distribution of selection coefficients of our lineages in TSOY also has a clear  
501 mode near  $s = 0$  and among lineages whose selection coefficients cannot be statistically

502 distinguished from  $s = 0$ , most are clearly negative (Chi-square test;  $\chi^2 = 7.54$ ,  $df = 1$ ,  $p$   
503  $= 0.0060$ ) (see Figure 2). However, whether this is the only mode in the distribution, or  
504 stems from the inability of our MA experiments to fix deleterious mutations with fitness  
505 effects below the selection threshold of  $s = -0.078$  remains uncertain.

506 Whether the environment affects the fitness effects of spontaneous mutations  
507 has also been the subject of considerable debate. Specifically, some studies have  
508 shown that larger declines in fitness are experienced in harsher environments, while  
509 others have not (Martin and Lenormand 2006; Halligan and Keightley 2009; Kraemer *et*  
510 *al.* 2015). In M9MM+CAA and M9MM, we can statistically distinguish fitness effects  
511 from  $s = 0$  with greater precision ( $s < -0.03$  in M9MM+CAA and  $s < -0.01$  or  $s > 0.01$  in  
512 M9MM) likely because the formulations for these media are more defined. These media  
513 are also expected to be more stringent for growth than TSOY because nutrients are  
514 more limited. Yet, a similar distribution of fitness effects was observed across our MA  
515 lineages in these environments as we observed in TSOY (see Figure 2). Again, most  
516 lineages whose selection coefficients are statistically different from 0 in M9MM+CAA  
517 and M9MM have reduced fitness, and the only clear mode in the distribution of fitness  
518 effects occurs near  $s = 0$ . However, in M9MM there are four lineages that are  
519 significantly more fit than the ancestor, and unlike in TSOY and M9MM+CAA, lineages  
520 whose selection coefficients are not significantly different from 0 in M9MM are no more  
521 likely to be negative than they are to be positive (Chi-square test;  $\chi^2 = 0$ ,  $df = 1$ ,  $p = 1$ )  
522 (see Figure 2). This suggests that fewer spontaneous mutations are deleterious for  
523 fitness in M9MM, possibly because a greater proportion of genes are unused when  
524 metabolizing only a single carbon substrate. Furthermore, a number of lineages whose



525 selection coefficients were significantly different from  $s = 0$  in one environment, were  
526 near neutral in other environments. Overall, these data suggest that the fitness effects  
527 of some individual spontaneous mutations interact with the growth environment, despite  
528 the minimal differences in the properties of the distribution of fitness effects among the  
529 three environments assayed in this study (see Figure 2; Figure 3).

530 Quantitative estimates of the rates and fitness effects of spontaneous mutations  
531 have received considerable interest because of their fundamental importance for the  
532 population genetics of evolving systems (Keightley and Lynch 2003; Eyre-Walker and  
533 Keightley 2007; Lynch 2008a; b, 2010; Halligan and Keightley 2009). Specifically, in an  
534 effort to unveil some general properties of the distribution of fitness effects of  
535 spontaneous mutations, estimates of the average selection coefficient of spontaneous  
536 mutations ( $s$ ) (Lynch *et al.* 2008; Halligan and Keightley 2009; Trindade *et al.* 2010;  
537 Heilbron *et al.* 2014), the rate of deleterious mutations ( $U_D$ ) (Kibota and Lynch 1996;  
538 Trindade *et al.* 2010), and the average selection coefficient of deleterious mutations ( $s_D$ )  
539 (Kibota and Lynch 1996; Trindade *et al.* 2010; Heilbron *et al.* 2014) have been derived  
540 using both direct and indirect approaches. Here, we estimate that  $s \cong 0$  in all three  
541 environments, largely because the vast majority of mutations appear to have near  
542 neutral effects on fitness. These estimates are remarkably similar to estimates from  
543 studies of MA lines with fully characterized mutational load in *P. aeruginosa* and *S.*  
544 *cerevisiae* (Lynch *et al.* 2008; Heilbron *et al.* 2014), but are lower than estimates  
545 derived from unsequenced MA lineages (Halligan and Keightley 2009; Trindade *et al.*  
546 2010). However, it is important to consider that while our data suggest that the vast  
547 majority of spontaneous mutations in *B. cenocepacia* have very low selection

548 coefficients in the laboratory, it should not imply that all of these mutations are  
549 effectively neutral in natural conditions. In fact, sequence analyses in enteric bacteria  
550 have revealed that fewer than 2.8% of amino-acid changing mutations are evolving  
551 neutrally, and this may be an overestimate due to the presence of adaptive mutations  
552 (Charlesworth and Eyre-Walker 2006; Eyre-Walker and Keightley 2007).

553 From this study, we were also able to estimate the deleterious mutation rates  
554 ( $U_D$ ) and the mean fitness effects of deleterious mutations ( $s_D$ ) in each environment  
555 under the presumption that only one mutation is responsible for the fitness declines  
556 experienced by a subset of our lineages. This assumption is warranted given that more  
557 than 5000 generations of MA had no significant effects on lineage fitness in the majority  
558 of our lineages, and that there was no significant correlation between the number of  
559 mutations in a lineage and its fitness. Further, time-series data from prior studies have  
560 supported that single, large-effect mutations disproportionately dictate the realized  
561 fitness in the majority of MA lineages (Davies *et al.* 1999; Heilbron *et al.* 2014). As was  
562 the case with our estimates of  $s$ , our estimates of  $U_D$  and  $s_D$  in all three environments  
563 are similar to prior estimates in *E. coli* (Kibota and Lynch 1996; Trindade *et al.* 2010)  
564 (TSOY:  $U_D = 7.12 \times 10^{-5}$  /genome/generation,  $s_D = -0.048$ ; M9MM+CAA:  $U_D = 5.44 \times$   
565  $10^{-5}$ ,  $s_D = -0.053$ ; M9MM:  $U_D = 5.44 \times 10^{-5}$ ,  $s_D = -0.048$ ). Interestingly, our  $U_D$  estimates  
566 are all slightly lower than the estimates from previous studies and our  $s_D$  estimates are  
567 somewhat greater, which is consistent with a failure to differentiate some moderately  
568 deleterious mutations from  $s = 0$  in this study. It is likely that both the difficulty of  
569 resolving the fitness of mutations of small effects and the effects of selection in our MA  
570 experiments on mutations with  $s > 0.078$  interfere with our estimates of  $U_D$  and  $s_D$ .

571 While the latter selective bias is an inevitable consequence of the MA approach, future  
572 studies can enhance resolution of subtle fitness effects by employing flow-cytometry  
573 and/or barcoding techniques (Gallet *et al.* 2012; Levy *et al.* 2015; Dillon *et al.* 2016).  
574 Furthermore, by extending these methods to a growing database of archived MA-WGS  
575 studies in diverse organisms and analyzing the effects of spontaneous mutations in  
576 different combinations, we will soon be able to evaluate the extent to which properties of  
577 the distribution of fitness effects vary between species and the epistatic deviation  
578 associated with mutations in different genetic backgrounds (Elena and Lenski 1997;  
579 Silander *et al.* 2007; Dickinson 2008; Schaack *et al.* 2013; Jasmin and Lenormand  
580 2015; Behringer and Hall 2016).

581       Significantly beneficial mutations were only observed in M9MM and in no lineage  
582 did the selective coefficient exceed  $s = +0.03$ . These limited observations prevent us  
583 from performing any detailed analyses on the rate and distribution of effects of  
584 beneficial mutations, but they support the notion that beneficial mutations are rare  
585 relative to deleterious mutations (Keightley and Lynch 2003). Furthermore, they suggest  
586 that most beneficial mutations likely provide only moderate benefits, even though the  
587 beneficial mutations that often fix in experimental populations can have large beneficial  
588 effects (Lenski *et al.* 1991; Ostrowski *et al.* 2005; Lang *et al.* 2013; Levy *et al.* 2015). As  
589 was the case for deleterious mutations, the number and average effects of beneficial  
590 mutations observed in this study are likely to be downwardly biased due to the impacts  
591 of other mutations in the same MA background that are mostly deleterious, which will  
592 further increase the value of sequencing and measuring the fitness of our lineages at  
593 intermediate time-points in future studies.

594           It is a well-established dogma in evolutionary biology that mutations that disrupt  
595 coding sequences are most likely to affect fitness, but this has never been quantitatively  
596 tested with naturally accumulated mutations. Specifically, mutations that frequently  
597 generate non-functional proteins, like nonsense bpsms or coding indels, are expected to  
598 have the most deleterious effects, followed by missense bpsms that mostly generate  
599 modified proteins, then synonymous and non-coding mutations that do not alter protein  
600 sequences. The fitness effects of plasmid gain and loss are less certain, as the size and  
601 genetic content of plasmids vary, but they may be energetically expensive to maintain  
602 (Smith and Bidochka 1998). Consequently, plasmids may be selectively lost in  
603 permissive laboratory environments where maintenance of the plasmid has a fitness  
604 cost (Lenski and Bouma 1987; Smith and Bidochka 1998). Our data suggest that  
605 although loss of the 0.164-Mb plasmid in *B. cenocepacia* occurs at an appreciable rate  
606 in the absence of selection during our MA experiments, it is universally deleterious to  
607 lose the plasmid in TSOY and M9MM. However, these effects appear to be mitigated in  
608 M9MM+CAA, suggesting that these fitness losses are related to amino acid synthesis.  
609 Therefore, these data suggest that in permissive laboratory conditions, the loss of some  
610 plasmids can be deleterious and not just advantageous for growth as is widely  
611 presumed. Other mutation types were not significantly overrepresented in lineages with  
612 significantly reduced fitness (see Table 2), but we do find that there are slightly more  
613 nonsense bpsms, missense bpsms, and coding indels than expected in lineages with  
614 significantly reduced fitness. This supports expectations that protein-modifying  
615 mutations are more likely to affect fitness than synonymous or intergenic mutations, and  
616 that most synonymous and intergenic mutations do not measurably affect fitness, even

617 though some synonymous and intergenic mutations can be under selective constraints  
618 (Eyre-Walker and Keightley 2007; Bailey *et al.* 2014).

619         The rate and fitness effects of spontaneous mutations are fundamental quantities  
620 that will help explain a number of evolutionary phenomena, including the origin and  
621 maintenance of genetic variation in natural populations (Charlesworth *et al.* 1993, 2009;  
622 Charlesworth and Charlesworth 1998), the evolution of recombination (Muller 1964;  
623 Kondrashov 1988; Otto and Lenormand 2002; Roze and Blanckaert 2014), the evolution  
624 of mutator alleles (Sniegowski *et al.* 1997; Tenaillon *et al.* 1999), and the mutational  
625 meltdown of small populations (Lande 1994; Lynch *et al.* 1995, 1999; Schwander and  
626 Crespi 2009). Although our methods show that the majority of spontaneous mutations in  
627 *B. cenocepacia* do not produce detectable fitness effects in any of three laboratory  
628 environments, natural selection may be operating on many of these sites in natural  
629 populations (Charlesworth and Eyre-Walker 2006). Still, our data suggest that purifying  
630 selection may act weakly on much of the variation generated by spontaneous mutation,  
631 allowing mutation pressure to substantively influence the genetic variation in natural  
632 populations under some conditions. The low frequencies of highly deleterious or  
633 beneficial alleles observed in this study may also have implications for our  
634 understanding of the evolution of recombination and mutation rates. Specifically, the  
635 benefits of reassorting genomes to separate deleterious and beneficial mutations (to  
636 avoid Muller's ratchet), or to combine beneficial mutations and accelerate adaptation  
637 (the Fisher-Muller hypothesis) might be reduced under conditions where both  
638 deleterious and beneficial mutations are rare (Zeyl and Bell 1997; Otto and Lenormand  
639 2002; de Visser and Elena 2007), like those observed in this study. Mutator alleles may

640 also be tolerated for longer periods if they fail to produce detectable deleterious load,  
641 thereby giving them more time to hitchhike to fixation by producing a highly beneficial  
642 allele. Indeed, mutator alleles have been frequently observed in experimental  
643 (Sniegowski *et al.* 1997), clinical (Mena *et al.* 2008; Oliver 2010; Silva *et al.* 2016), and  
644 environmental studies (Hall and Henderson-Begg 2006; Hazen *et al.* 2009). Lastly, the  
645 bias for spontaneous mutations to be deleterious rather than beneficial in *B.*  
646 *cenocepacia* is consistent with studies in a number of species (Kibota and Lynch 1996;  
647 Keightley and Caballero 1997; Fry *et al.* 1999; Vassilieva *et al.* 2000; Wloch *et al.* 2001;  
648 Zeyl and de Visser 2001; Keightley and Lynch 2003; Trindade *et al.* 2010; Heilbron *et*  
649 *al.* 2014), suggesting that gradual mutational meltdown may be inevitable for  
650 populations where the efficiency of natural selection is reduced (Sniegowski and Lenski  
651 1995; Lynch *et al.* 1999).

652 By measuring the fitness effects of MA lineages with fully characterized  
653 mutational load, we have performed a uniquely systematic study of the rate and fitness  
654 effects of spontaneous mutations with known genetic bases, demonstrating that the vast  
655 majority of mutations accumulated in *B. cenocepacia* MA lines are neutral or deleterious  
656 for fitness, and that the fitness of individual mutations can be environmentally  
657 dependent, even though the general features of the distribution of fitness effects are  
658 similar in different environments. Furthermore, we have provided new estimates for  
659 several parameters of the distribution of fitness effects of spontaneous mutations  
660 derived from genotypes with known mutational load. Although considerable uncertainty  
661 remains with respect to the shape and parameters that define the distribution of fitness  
662 effects, by extending these methods to a growing database of MA studies, many of

663 which have archived time-series data, we are now poised to dramatically enhance our  
664 understanding of the true nature of the distribution of fitness effects of spontaneous  
665 mutations.

666

667

## **ACKNOWLEDGMENTS**

668 This work was supported by the National Science Foundation Career Award (DEB-  
669 0845851 to VSC).

TABLES

**Table 1.** An experiment-wide ANOVA of the fitness of the 43 final mutation accumulation lines in each of 3 environments, measured in 4 replicates per treatment.

Effect	df	SS	F	p
Replicate	3	0.0037	10.02	< 0.0001
Line	42	0.3302	63.72	< 0.0001
Environment	2	0.0190	76.90	< 0.0001
Genotype • Environment	84	0.1166	11.25	< 0.0001



**Table 2.** Chi-tests comparing the observed number of mutations in lineages with significantly reduced fitness, relative to the expected amounts given the proportion of the total lineages that had significantly reduced fitness.

Mutation Type	Environment	Observed	Expected	$\chi^2$	df	p
Intergenic Bpsm	TSOY	13	14.63	0.2996	1	0.5841
Synonymous Bpsm	TSOY	16	19.37	0.9708	1	0.3245
Missense Bpsm	TSOY	60	55.74	0.5374	1	0.4635
Nonsense Bpsm	TSOY	4	2.37	1.8477	1	0.1741
Intergenic Indel	TSOY	8	7.91	0.0018	1	0.9661
Coding Indel	TSOY	12	8.70	2.0736	1	0.1499
Plasmid Loss	TSOY	4	1.58	6.1176	1	0.0134
<b>Total Mutations</b>	<b>TSOY</b>	<b>117</b>	<b>110.30</b>	<b>0.6726</b>	<b>1</b>	<b>0.4121</b>
Intergenic Bpsm	M9MM+CAA	12	11.19	0.0849	1	0.7708
Synonymous Bpsm	M9MM+CAA	15	14.81	0.0033	1	0.9539
Missense Bpsm	M9MM+CAA	47	42.63	0.6427	1	0.4227
Nonsense Bpsm	M9MM+CAA	3	1.81	1.1115	1	0.2917
Intergenic Indel	M9MM+CAA	6	6.05	0.0005	1	0.9819
Coding Indel	M9MM+CAA	6	6.65	0.0914	1	0.7624
Plasmid Loss	M9MM+CAA	1	1.21	0.0519	1	0.8198
<b>Total Mutations</b>	<b>M9MM+CAA</b>	<b>90</b>	<b>84.35</b>	<b>0.5427</b>	<b>1</b>	<b>0.4613</b>
Intergenic Bpsm	M9MM	10	11.19	0.1803	1	0.6712
Synonymous Bpsm	M9MM	15	14.81	0.0033	1	0.9539
Missense Bpsm	M9MM	50	42.63	1.8274	1	0.1764
Nonsense Bpsm	M9MM	2	1.81	0.0274	1	0.8686
Intergenic Indel	M9MM	6	6.05	0.0005	1	0.9819
Coding Indel	M9MM	8	6.65	0.3921	1	0.5312
Plasmid Loss	M9MM	4	1.21	9.2308	1	0.0024
<b>Total Mutations</b>	<b>M9MM</b>	<b>95</b>	<b>84.35</b>	<b>1.9278</b>	<b>1</b>	<b>0.1650</b>

## FIGURE LEGENDS

680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698

**Figure 1.** Distribution of base-substitution mutations (bpsms) and insertion-deletion mutations (indels) across the 43 mutation accumulation lineages derived from *Burkholderia cenocepacia* HI2424 analyzed in this study.

**Figure 2.** Distribution of the selection coefficients of each *Burkholderia cenocepacia* MA lineage relative to the ancestral *B. cenocepacia* HI2424 strain in tryptic soy broth (A), M9 minimal medium supplemented with casamino acids (B), and M9 minimal medium (C). Significance was determined from independent two-tailed t-tests on four replicate fitness assays for each lineage. P-values were corrected for multiple comparisons using a Benjamini-Hochberg correction, and corrected p-values that remained below 0.05 were considered significant.

**Figure 3.** Relationship between selection coefficients of all MA lineages in each of the different pairs of environments. All linear regressions are significant, but much of the variance is unexplained (A:  $F = 17.18$ ,  $df = 41$ ,  $p = 0.0002$ ,  $r^2 = 0.2953$ ; B:  $F = 8.613$ ,  $df = 41$ ,  $p = 0.0054$ ,  $r^2 = 0.1736$ ; C:  $F = 124.00$ ,  $df = 41$ ,  $p < 0.0001$ ,  $r^2 = 0.7515$ ).

699

## LITERATURE CITED

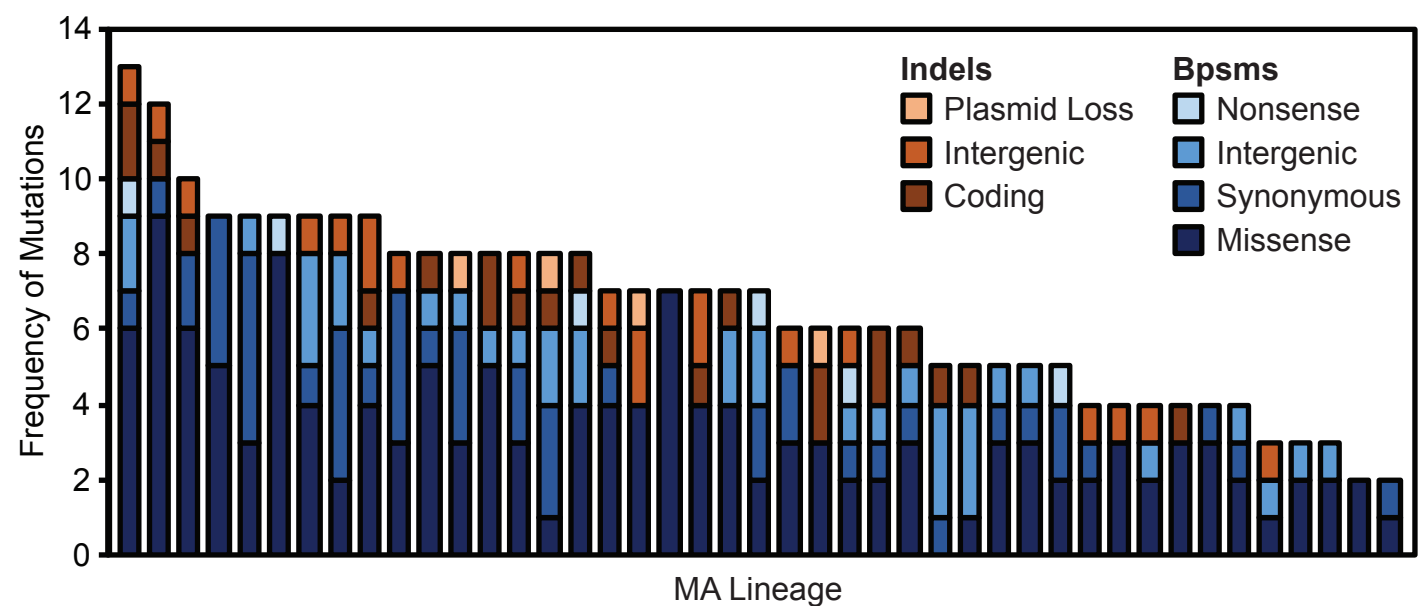
- 700 Agnoli K., Schwager S., Uehlinger S., Vergunst A., Viteri D. F., Nguyen D. T., Sokol P. A., Carlier A.,  
701 Eberl L., 2012 Exposing the third chromosome of *Burkholderia cepacia* complex strains as a  
702 virulence plasmid. *Mol. Microbiol.* **83**: 362–378.
- 703 Agrawal A. F., Whitlock M. C., 2012 Mutation load: The fitness of individuals in populations where  
704 deleterious alleles are abundant. *Annu. Rev. Ecol. Evol. Syst.* **43**: 115–135.
- 705 Bailey S. F., Hinz A., Kassen R., 2014 Adaptive synonymous mutations in an experimentally evolved  
706 *Pseudomonas fluorescens* population. *Nat. Commun.* **5**: 4076.
- 707 Bateman A. J., 1959 The viability of near-normal irradiated chromosomes. *Int. J. Radiat. Biol.* **1**: 170–180.
- 708 Baym M., Kryazhimskiy S., Lieberman T. D., Chung H., Desai M. M., Kishony R., 2015 Inexpensive  
709 multiplexed library preparation for megabase-sized genomes. *PLoS One* **10**: e0128036.
- 710 Behringer M. G., Hall D. W., 2016 The repeatability of genome-wide mutation rate and spectrum  
711 estimates. *Curr. Genet.* **1**: 1–6.
- 712 Benjamini Y., Hochberg Y., 1995 Controlling the false discovery rate: a practical and powerful approach  
713 to multiple testing. *J. R. Stat. Soc.* **57**: 289–300.
- 714 Charlesworth B., Morgan M. T., Charlesworth D., 1993 The effect of deleterious mutations on neutral  
715 molecular variation. *Genetics* **134**: 1289–1303.
- 716 Charlesworth B., Charlesworth D., 1998 Some evolutionary consequences of deleterious mutations.  
717 *Genetica* **102-103**: 3–19.
- 718 Charlesworth J., Eyre-Walker A., 2006 The rate of adaptive evolution in enteric bacteria. *Mol. Biol. Evol.*  
719 **23**: 1348–1356.
- 720 Charlesworth B., Betancourt A. J., Kaiser V. B., Gordo I., 2009 Genetic recombination and molecular  
721 evolution. In: *Cold Spring Harbor Symposia on Quantitative Biology*, pp. 177–186.
- 722 Chevin L.-M., 2011 On measuring selection in experimental evolution. *Biol. Lett.* **7**: 210–213.
- 723 Choi K.-H., Gaynor J. B., White K. G., Lopez C., Bosio C. M., Karkhoff-Schweizer R. R., Schweizer H. P.,  
724 2005 A Tn7-based broad-range bacterial cloning and expression system. *Nat. Methods* **2**: 443–448.
- 725 Coenye T., LiPuma J. J., 2003 Population structure analysis of *Burkholderia cepacia* genomovar III:  
726 varying degrees of genetic recombination characterize major clonal complexes. *Microbiology-Sgm*  
727 **149**: 77–88.
- 728 Davies E. K., Peters a D., Keightley P. D., 1999 High frequency of cryptic deleterious mutations in  
729 *Caenorhabditis elegans*. *Science* **285**: 1748–1751.
- 730 Dettman J. R., Sztepanacz J. L., Kassen R., 2016 The properties of spontaneous mutations in the  
731 opportunistic pathogen *Pseudomonas aeruginosa*. *BMC Genomics* **17**: 27–41.
- 732 Dickinson W. J., 2008 Synergistic fitness interactions and a high frequency of beneficial changes among  
733 mutations accumulated under relaxed selection in *Saccharomyces cerevisiae*. *Genetics* **178**: 1571–  
734 1578.
- 735 Dillon M. M., Sung W., Lynch M., Cooper V. S., 2015 The rate and molecular spectrum of spontaneous  
736 mutations in the GC-rich multichromosome genome of *Burkholderia cenocepacia*. *Genetics* **200**:  
737 935–946.
- 738 Dillon M. M., Rouillard N. P., Dam B. Van, Gallet R., Cooper V. S., 2016 Diverse phenotypic and genetic  
739 responses to short-term selection in evolving *Escherichia coli* populations. *Evolution (N. Y.)* **70**:  
740 586–599.
- 741 Elena S. F., Lenski R. E., 1997 Test of synergistic interactions among deleterious mutations in bacteria.  
742 *Nature* **390**: 395–398.
- 743 Elena S. F., Ekuwe L., Hajela N., Oden S. A., Lenski R. E., 1998 Distribution of fitness effects caused by  
744 random insertion mutations in *Escherichia coli*. *Genetica* **102**: 349–358.

- 745 Estes S., Phillips P. C., Denver D. R., Thomas W. K., Lynch M., 2004 Mutation accumulation in  
746 populations of varying size: The distribution of mutational effects for fitness correlates in  
747 *Caenorhabditis elegans*. *Genetics* **166**: 1269–1279.
- 748 Eyre-Walker A., Keightley P. D., 2007 The distribution of fitness effects of new mutations. *Nat. Rev.*  
749 *Genet.* **8**: 610–8.
- 750 Foster P. L., Lee H., Popodi E., Townes J. P., Tang H., 2015 Determinants of spontaneous mutation in  
751 the bacterium *Escherichia coli* as revealed by whole-genome sequencing. *Proc. Natl. Acad. Sci.*  
752 **112**: E5990–E5999.
- 753 Fry J. D., Keightley P. D., Heinsohn S. L., Nuzhdin S. V., 1999 New estimates of the rates and effects of  
754 mildly deleterious mutation in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U. S. A.* **96**: 574–579.
- 755 Gallet R., Cooper T. F., Elena S. F., Lenormand T., 2012 Measuring selection coefficients below 10<sup>(-3)</sup>:  
756 Method, questions, and prospects. *Genetics* **190**: 175–186.
- 757 Hall L. M. C., Henderson-Begg S. K., 2006 Hypermutable bacteria isolated from humans - a critical  
758 analysis. *Microbiology* **152**: 2505–2514.
- 759 Hall D. W., Mahmoudizad R., Hurd A. W., Joseph S. B., 2008 Spontaneous mutations in diploid  
760 *Saccharomyces cerevisiae*: another thousand cell generations. *Genet. Res. (Camb.)* **90**: 229–241.
- 761 Halligan D. L., Keightley P. D., 2009 Spontaneous mutation accumulation studies in evolutionary  
762 genetics. *Annu. Rev. Ecol. Evol. Syst.* **40**: 151–172.
- 763 Hazen T. H., Kennedy K. D., Chen S., Yi S. V., Sobecky P. A., 2009 Inactivation of mismatch repair  
764 increases the diversity of *Vibrio parahaemolyticus*. *Environ. Microbiol.* **11**: 1254–1266.
- 765 Heilbron K., Toll-Riera M., Kojadinovic M., Maclean R. C., 2014 Fitness is strongly influenced by rare  
766 mutations of large effect in a microbial mutation accumulation experiment. *Genetics* **197**: 981–990.
- 767 Jasmin J., Lenormand T., 2015 Accelerating mutational load is not due to synergistic epistasis or mutator  
768 alleles in mutation accumulation lines of yeast. *Genetics*.
- 769 Katju V., Packard L. B., Bu L., Keightley P. D., Bergthorsson U., 2015 Fitness decline in spontaneous  
770 mutation accumulation lines of *Caenorhabditis elegans* with varying effective population sizes.  
771 *Evolution (N. Y.)* **69**: 104–116.
- 772 Keightley P. D., 1994 The distribution of mutation effects on viability in *Drosophila melanogaster*.  
773 *Genetics* **138**: 1315–1322.
- 774 Keightley P. D., Caballero A., 1997 Genomic mutation rates for lifetime reproductive output and lifespan  
775 in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. U. S. A.* **94**: 3823–3827.
- 776 Keightley P. D., Lynch M., 2003 Toward a realistic model of mutations affecting fitness. *Evolution (N. Y.)*.  
777 **57**: 683–685.
- 778 Kibota T. T., Lynch M., 1996 Estimate of the genomic mutation rate deleterious to overall fitness in *E. coli*.  
779 *Nature* **381**: 694–696.
- 780 Kimura M., 1983 *The Neutral Theory of Molecular Evolution*. Cambridge University Press, Cambridge,  
781 New York.
- 782 Kondrashov A. S., 1988 Deleterious mutations and the evolution of sexual reproduction. *Nature* **336**:  
783 435–440.
- 784 Kraemer S. A., Morgan A. D., Ness R. W., Keightley P. D., Colegrave N., 2015 Fitness effects of new  
785 mutations in *Chlamydomonas reinhardtii* across two stress gradients. *J. Evol. Biol.* **44**: n/a–n/a.
- 786 Lande R., 1994 Risk of population extinction from fixation of new deleterious mutations. *Evolution (N. Y.)*.  
787 **48**: 1460–1469.
- 788 Lang G. I., Rice D. P., Hickman M. J., Sodergren E., Weinstock G. M., Botstein D., Desai M. M., 2013  
789 Pervasive genetic hitchhiking and clonal interference in forty evolving yeast populations. *Nature* **500**:  
790 571–574.
- 791 Lee H., Popodi E., Tang H. X., Foster P. L., 2012 Rate and molecular spectrum of spontaneous mutations

- 792 in the bacterium *Escherichia coli* as determined by whole-genome sequencing. Proc. Natl. Acad.  
793 Sci. U. S. A. **109**: E2774–E2783.
- 794 Lenski R. E., Bouma J. E., 1987 Effects of segregation and selection on instability of plasmid pACYC184  
795 in *Escherichia coli* B. J. Bacteriol. **169**: 5314–5316.
- 796 Lenski R. E., Rose M. R., Simpson S. C., Tadler S. C., 1991 Long-term experimental evolution in  
797 *Escherichia coli*. 1. Adaptation and divergence during 2,000 generations. Am. Nat. **138**: 1315–1341.
- 798 Levy S. F., Blundell J. R., Venkataram S., Petrov D. A., Fisher D. S., Sherlock G., 2015 Quantitative  
799 evolutionary dynamics using high-resolution lineage tracking. Nature **519**: 181–186.
- 800 Li H., Durbin R., 2009 Fast and accurate short read alignment with Burrows-Wheeler transform.  
801 Bioinformatics **25**: 1754–1760.
- 802 Li H., Handsaker B., Wysoker A., Fennell T., Ruan J., Homer N., Marth G., Abecasis G., Durbin R., 2009  
803 The sequence alignment/map format and SAMtools. Bioinformatics **25**: 2078–2079.
- 804 LiPuma J. J., Spilker T., Coenye T., Gonzalez C. F., 2002 An epidemic *Burkholderia cepacia* complex  
805 strain identified in soil. Lancet **359**: 2002–2003.
- 806 Long H., Sung W., Miller S. F., Ackerman M. S., Doak T. G., Lynch M., 2014 Mutation rate, spectrum,  
807 topology, and context-dependency in the DNA mismatch repair (MMR) deficient *Pseudomonas*  
808 *fluorescens* ATCC948. Genome Biol. Evol. **7**: 262–271.
- 809 Long H., Kucukyildirim S., Sung W., Williams E., Lee H., Ackerman M., Doak T. G., Tang H., Lynch M.,  
810 2015 Background mutational features of the radiation-resistant bacterium *Deinococcus radiodurans*.  
811 Mol. Biol. Evol. **32**: 2383–2392.
- 812 Lynch M., Conery J., Burger R., 1995 Mutation accumulation and the extinction of small populations. Am.  
813 Nat. **146**: 489–518.
- 814 Lynch M., Blanchard J., Houle D., Kibota T., Schultz S., Vassilieva L., Willis J., 1999 Perspective:  
815 Spontaneous deleterious mutation. Evolution (N. Y). **53**: 645–663.
- 816 Lynch M., 2008a Estimation of Nucleotide Diversity, Disequilibrium Coefficients, and Mutation Rates from  
817 High-Coverage Genome-Sequencing Projects. Mol. Biol. Evol. **25**: 2409–2419.
- 818 Lynch M., 2008b The Cellular, Developmental and Population-Genetic Determinants of Mutation-Rate  
819 Evolution. Genetics **180**: 933–943.
- 820 Lynch M., Sung W., Morris K., Coffey N., Landry C. R., Dopman E. B., Dickinson W. J., Okamoto K.,  
821 Kulkarni S., Hartl D. L., Thomas W. K., 2008 A genome-wide view of the spectrum of spontaneous  
822 mutations in yeast. Proc. Natl. Acad. Sci. U. S. A. **105**: 9272–9277.
- 823 Lynch M., 2010 Evolution of the mutation rate. Trends Genet. **26**: 345–352.
- 824 Mahenthiralingam E., Urban T. A., Goldberg J. B., 2005 The multifarious, multireplicon *Burkholderia*  
825 *cepacia* complex. Nat. Rev. Microbiol. **3**: 144–156.
- 826 Martin G., Lenormand T., 2006 The fitness effect of mutations across environments: a survey in light of  
827 fitness landscape models. Evolution **60**: 2413–2427.
- 828 Mena A., Smith E. E., Burns J. L., Speert D. P., Moskowitz S. M., Perez J. L., Oliver A., 2008 Genetic  
829 adaptation of *Pseudomonas aeruginosa* to the airways of cystic fibrosis patients is catalyzed by  
830 hypermutation. J. Bacteriol. **190**: 7910–7917.
- 831 Mukai T., 1964 Genetic Structure of Natural Populations of *Drosophila Melanogaster*. 1. Spontaneous  
832 Mutation Rate of Polygenes Controlling Viability. Genetics **50**: 1–&.
- 833 Muller H. J., 1964 The relation of recombination to mutational advance. Mutat. Res. Mol. Mech. Mutagen.  
834 **1**: 2–9.
- 835 Oliver A., 2010 Mutators in cystic fibrosis chronic lung infection: Prevalence, mechanisms, and  
836 consequences for antimicrobial therapy. Int J Med Microbiol **300**: 563–572.
- 837 Ostrowski E. A., Rozen D. E., Lenski R. E., 2005 Pleiotropic effects of beneficial mutations in *Escherichia*  
838 *coli*. Evolution (N. Y). **59**: 2343–2352.

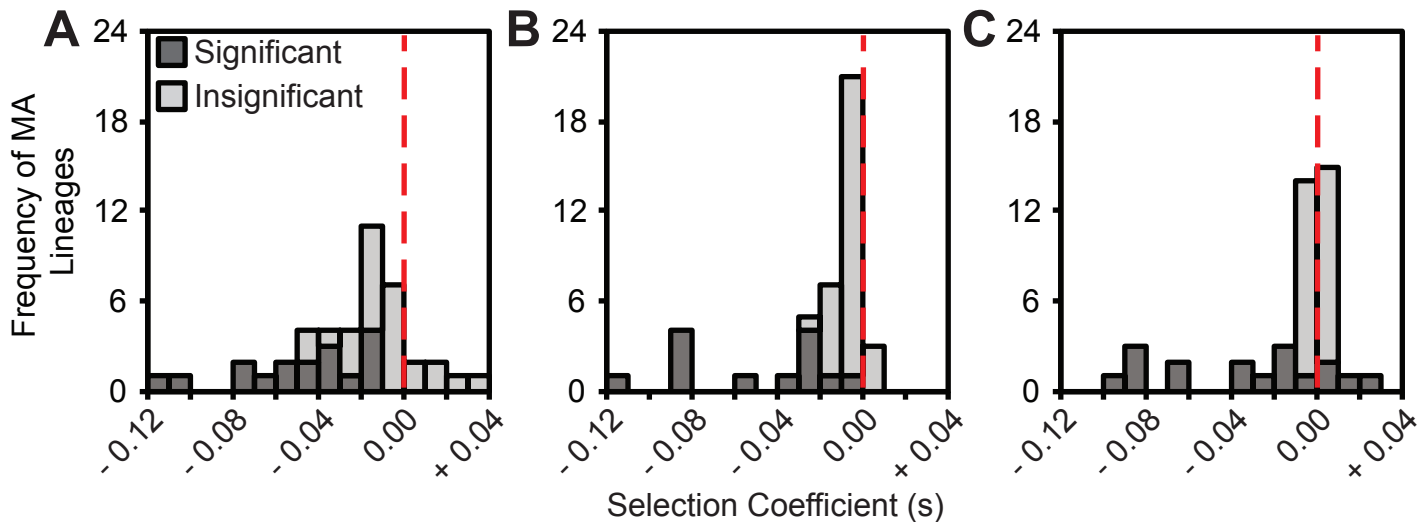
- 839 Otto S. P., Lenormand T., 2002 Resolving the paradox of sex and recombination. *Nat Rev Genet* **3**: 252–  
840 261.
- 841 Perfeito L., Sousa a., Bataillon T., Gordo I., 2014 Rates of fitness decline and rebound suggest pervasive  
842 epistasis. *Evolution* (N. Y). **68**: 150–162.
- 843 R Development Core Team, 2013 R: A Language and Environment for Statistical Computing.
- 844 Roze D., Blanckaert A., 2014 Epistasis, pleiotropy, and the mutation load in sexual and asexual  
845 populations. *Evolution* (N. Y). **68**: 137–149.
- 846 Schaack S., Allen D. E., Latta L. C., Morgan K. K., Lynch M., 2013 The effect of spontaneous mutations  
847 on competitive ability. *J. Evol. Biol.* **26**: 451–6.
- 848 Schultz S. T., Lynch M., Willis J. H., 1999 Spontaneous deleterious mutation in *Arabidopsis thaliana*.  
849 *Proc. Natl. Acad. Sci. U. S. A.* **96**: 11393–8.
- 850 Schwander T., Crespi B. J., 2009 Twigs on the tree of life? Neutral and selective models for integrating  
851 macroevolutionary patterns with microevolutionary processes in the analysis of asexuality. *Mol.*  
852 *Ecol.* **18**: 28–42.
- 853 Shaw R. G., Byers D. L., Darmo E., 2000 Spontaneous mutational effects on reproductive traits of  
854 *Arabidopsis thaliana*. *Genetics* **155**: 369–378.
- 855 Shaw F. H., Geyer C. J., Shaw R. G., 2002 A comprehensive model of mutations affecting fitness and  
856 inferences for *Arabidopsis thaliana*. *Evolution* **56**: 453–463.
- 857 Silander O. K., Tenaillon O., Chao L., 2007 Understanding the evolutionary fate of finite populations: the  
858 dynamics of mutational effects. *PLoS Biol.* **5**: e94.
- 859 Silva I. N., Santos P. M., Santos M. R., Zlosnik J. E. A., Speert D. P., Buskirk S., Waters C., Cooper V. S.,  
860 Moreira L. M., 2016 Long-term evolution of *Burkholderia multivorans* during a chronic cystic fibrosis  
861 infection reveals shifting forces of selection. *mSystems* **1**: e00029–16.
- 862 Smith M. A., Bidochka M. J., 1998 Bacterial fitness and plasmid loss: the importance of culture conditions  
863 and plasmid size. *Can. J. Microbiol.* **44**: 351–355.
- 864 Sniegowski P. D., Lenski R. E., 1995 Mutation and adaptation - the directed mutation controversy in  
865 evolutionary perspective. *Annu. Rev. Ecol. Syst.* **26**: 553–578.
- 866 Sniegowski P. D., Gerrish P. J., Lenski R. E., 1997 Evolution of high mutation rates in experimental  
867 populations of *E. coli*. *Nature* **387**: 703–705.
- 868 Sung W., Ackerman M. S., Miller S. F., Doak T. G., Lynch M., 2012 Drift-barrier hypothesis and mutation-  
869 rate evolution. *Proc. Natl. Acad. Sci. U. S. A.* **109**: 18488–18492.
- 870 Sung W., Ackerman M. S., Gout J.-F., Miller S. F., Williams E., Foster P. L., Lynch M., 2015 Asymmetric  
871 context-dependent mutation patterns revealed through mutation-accumulation experiments. *Mol.*  
872 *Biol. Evol.* **32**: 1672–1683.
- 873 Tenaillon O., Toupance B., Nagard H. Le, Taddei F., Godelle B., 1999 Mutators, population size, adaptive  
874 landscape and the adaptation of asexual populations of bacteria. *Genetics* **152**: 485–493.
- 875 Traverse C. C., Mayo-Smith L. M., Poltak S. R., Cooper V. S., 2013 Tangled bank of experimentally  
876 evolved *Burkholderia* biofilms reflects selection during chronic infections. *Proc. Natl. Acad. Sci. U. S.*  
877 *A.* **110**: E250–E259.
- 878 Trindade S., Perfeito L., Gordo I., 2010 Rate and effects of spontaneous mutations that affect fitness in  
879 mutator *Escherichia coli*. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **365**: 1177–1186.
- 880 Vassilieva L. L., Hook a M., Lynch M., 2000 The fitness effects of spontaneous mutations in  
881 *Caenorhabditis elegans*. *Evolution* (N. Y). **54**: 1234–46.
- 882 Visser J. A. G. M. de, Elena S. F., 2007 The evolution of sex: empirical insights into the roles of epistasis  
883 and drift. *Nat. Rev. Genet.* **8**: 139–149.
- 884 Wloch D. M., Szafraniec K., Borts R. H., Korona R., 2001 Direct estimate of the mutation rate and the  
885 distribution of fitness effects in the yeast *Saccharomyces cerevisiae*. *Genetics* **159**: 441–452.

- 886 Ye K., Schulz M. H., Long Q., Apweiler R., Ning Z. M., 2009 Pindel: a pattern growth approach to detect  
887 break points of large deletions and medium sized insertions from paired-end short reads.  
888 *Bioinformatics* **25**: 2865–2871.
- 889 Zeyl C., Bell G., 1997 The advantage of sex in evolving yeast populations. *Nature* **388**: 465–468.
- 890 Zeyl C., Visser J. A. de, 2001 Estimates of the rate and distribution of fitness effects of spontaneous  
891 mutation in *Saccharomyces cerevisiae*. *Genetics* **157**: 53–61.
- 892

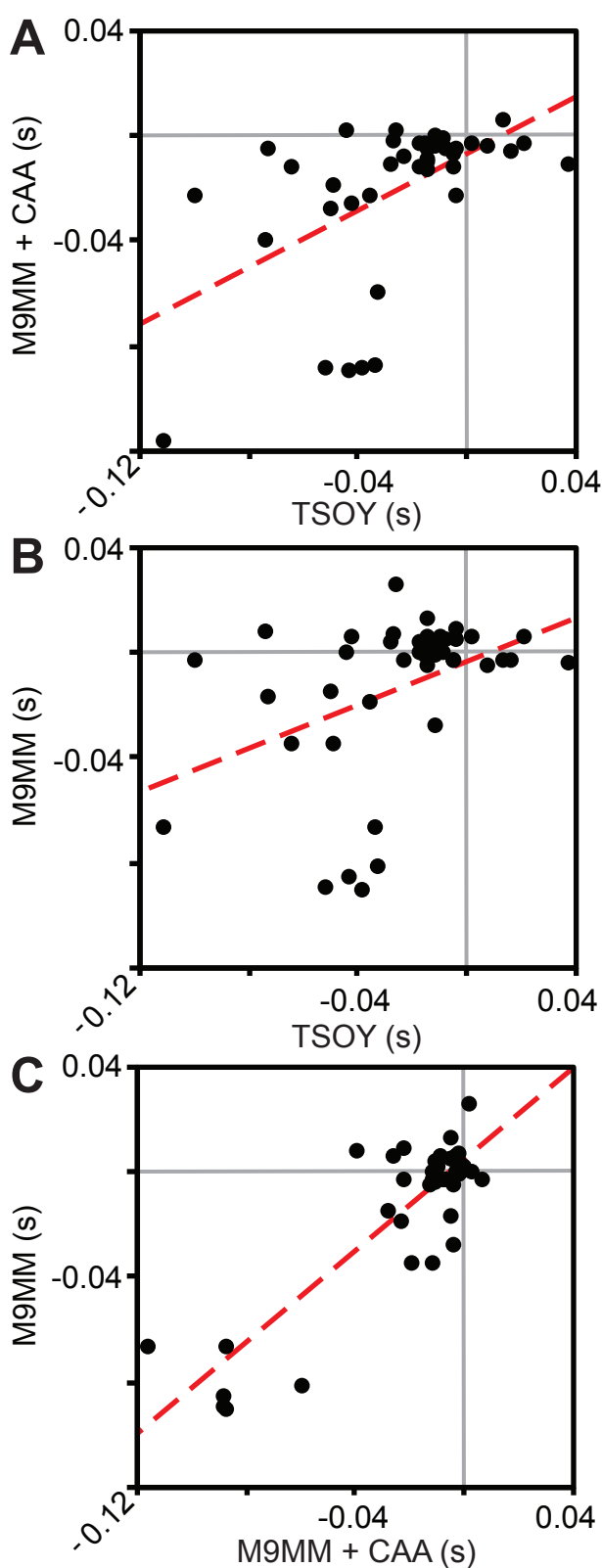


**Figure 1.** Distribution of base-substitution mutations (bpsms) and insertion-deletion mutations (indels) across the forty-three mutation accumulation lineages derived from *Burkholderia cenocepacia* HI2424 analyzed in this study.





**Figure 2.** Distribution of the selection coefficients of each *Burkholderia cenocepacia* MA lineage relative to the ancestral *B. cenocepacia* HI2424 strain in tryptic soy broth (A), M9 minimal medium supplemented with casamino acids (B), and M9 minimal medium (C). Significance was determined from independent two-tailed t-tests on four replicate fitness assays for each lineage. P-values were corrected for multiple comparisons using a Benjamini-Hochberg correction, and corrected p-values that remained below 0.05 were considered significant.



**Figure 3.** Relationship between selection coefficients of all MA lineages in each of the different pairs of environments. All linear regressions are significant, but much of the variance is unexplained (A:  $F = 17.18$ ,  $df = 41$ ,  $p = 0.0002$ ,  $r^2 = 0.2953$ ; B:  $F = 8.613$ ,  $df = 41$ ,  $p = 0.0054$ ,  $r^2 = 0.1736$ ; C:  $F = 124.00$ ,  $df = 41$ ,  $p < 0.0001$ ,  $r^2 = 0.7515$ ).