

PROCEEDINGS B

**Divergent evolution peaks under intermediate Population
Bottlenecks during bacterial Experimental Evolution**

Journal:	<i>Proceedings B</i>
Manuscript ID	RSPB-2016-0749.R1
Article Type:	Research
Date Submitted by the Author:	14-Jun-2016
Complete List of Authors:	Vogwill, Tom; University of Oxford, Zoology; Imperial College London - Silwood Park Campus Phillips, Robyn; University of Oxford, Zoology Gifford, Danna; University of Oxford, Zoology; University of Manchester, Faculty of Life Sciences MacLean, Craig; University of Oxford, Zoology
Subject:	Evolution < BIOLOGY
Keywords:	experimental evolution, evolutionary rescue, population bottlenecks, genomics
Proceedings B category:	Evolutionary Biology

SCHOLARONE™
Manuscripts

1 **Divergent evolution peaks under intermediate Population Bottlenecks during bacterial**
2 **Experimental Evolution**

3

4 Tom Vogwill^{1, 2*}, Robyn L. Phillips^{1*}, Danna R. Gifford^{1,3*}, and R. Craig MacLean¹

5

6 *joint first authors.

7

8 **Affiliations:**

9 ¹University of Oxford, Department of Zoology, South Parks Road, Oxford OX1 3PS, UK

10 ²Present address: Imperial College London, Department of Life Sciences, Silwood Park Campus.

11 ³Present address: University of Manchester, Faculty of Life Sciences.

12

13 Correspondence to: Craig MacLean (craig.maclean@zoo.ox.ac.uk) or Tom Vogwill
14 (tom.vogwill@zoo.ox.ac.uk)

15

16

Abstract

There is growing evidence that parallel molecular evolution is common, but its causes remain poorly understood. Demographic parameters such as population bottlenecks are predicted to be major determinants of parallelism. Here we test the hypothesis that bottleneck intensity shapes parallel evolution by elucidating the genomic basis of adaptation to antibiotic-supplemented medium in hundreds of populations of the bacterium *Pseudomonas fluorescens* PF0-1. As expected, bottlenecks decreased the rate of phenotypic and molecular adaptation. Surprisingly, bottlenecks had no impact on the likelihood of parallel adaptive molecular evolution at a genome-wide scale. However, bottlenecks had a profound impact on the genes involved in antibiotic resistance. Specifically, under either intense or weak bottlenecks, resistance predominantly evolved by strongly beneficial mutations which provide high-levels of antibiotic resistance. In contrast with intermediate bottlenecks regimes resistance evolved by a greater diversity of genetic mechanisms, significantly reducing the observed levels of parallel genetic evolution. Our results demonstrate that the population bottleneck can be a major predictor of parallel evolution, but precisely how may be more complex than many simple theoretical predictions.

32

Keywords: experimental evolution; population bottlenecks; evolutionary rescue; genome sequencing; parallel evolution.

35 **Introduction**

36 Parallel evolution, where the same beneficial mutations are fixed in independent populations or
 37 lineages, has now been documented in a wide range of organisms and in response to a range of
 38 selection pressures (Conte et al., 2012; Stern, 2013; Wood et al., 2005). However parallelism seems
 39 to be particularly common in bacteria, although it is far from universal. For example, some degree of
 40 parallel genetic evolution is commonly observed during host specialization in pathogens (Lieberman
 41 et al., 2011; Marvig et al., 2015a; Marvig et al., 2015b) and in endosymbionts (Dahan et al., 2015;
 42 McCutcheon et al., 2009; Wernegreen, 2002), and parallel evolution in antibiotic resistance genes
 43 occurs across highly divergent bacteria (Blair et al., 2015; Lieberman et al., 2011; Marvig et al.,
 44 2015a; Vogwill et al., 2014). It is unclear, however, what determines the precise level of observed
 45 parallel evolution in bacteria. It can partly be explained by bacteria having small compact genomes,
 46 orders of magnitude smaller than higher eukaryotes. It is also clear that, in some cases, genetic
 47 constraints promote parallel evolution (Conte et al., 2012; Losos, 2011; Ord and Summers, 2015). For
 48 example, there are very few genes in bacterial genomes that can be mutated to produce a high level
 49 of resistance to many antibiotics (Breidenstein et al., 2008; Dötsch et al., 2009; Fajardo et al., 2008;
 50 Liu et al., 2010), and unsurprisingly parallel evolution of resistance by mutations in these genes is
 51 common.

52 In addition to genetic constraints, demographic factors such as population bottlenecks are likely to
 53 be a major determinant of the repeatability of adaptation (de Visser and Krug, 2014; De Visser and
 54 Rozen, 2005; Szendro et al., 2013). Population bottlenecks are a common and unavoidable aspect of
 55 the demography of most organisms, but are practically unavoidable for pathogenic bacteria, due to
 56 transmission between hosts as well as strong selection from immune systems and antibiotics.
 57 Population bottlenecks can affect adaptation in a variety of ways, but these can be broadly grouped
 58 into genetic effects, which can be either stochastic or deterministic, and demographic effects, which
 59 are generally more deterministic in that bottlenecks increase mortality. For example, bottlenecks
 60 reduce genetic variation by stochastically eliminating rare alleles from populations, and the simplest
 61 consequence of bottlenecking is a reduction in the rate of adaptation (Colegrave, 2002; De Visser
 62 and Rozen, 2005). However, population bottlenecking is also predicted to have important
 63 consequences for the genetic mechanisms of adaptation. In large populations that experience weak
 64 bottlenecking, independently derived beneficial mutations can compete with each other, which has
 65 the potential to eliminate weakly beneficial mutations. The consequence of this effect, known as
 66 Hill-Robertson (Hill and Robertson, 1966) or clonal interference (Gerrish and Lenski, 1998), is that
 67 adaptation in large populations will be driven by strongly beneficial mutations in a sub-set of genes

68 that are under strong selection, resulting in a high probability of parallel evolution. This argument is
69 based on classical concepts from population genetics, and is solely based on differences in relative
70 fitness between competing genotypes. Put simply, this argument predicts that increasing the
71 severity of population bottlenecks should decrease the probability of parallel evolution.

72 In certain circumstances, however, the relationship between bottlenecking and parallelism should
73 not be quite so straight forward, as the increased mortality from bottlenecks can also affect
74 adaptation. If the intensity of bottlenecking is greater than the population growth rate, population
75 size will begin to decline, which will eventually result in extinction unless selection acts to increase
76 the population growth rate. In this scenario, which is often known as evolutionary rescue, the fate of
77 beneficial mutations depends on how they alter absolute fitness as opposed to relative fitness (Bell,
78 2013; Gomulkiewicz and Holt, 1995). For example weakly beneficial mutations, which only lead to a
79 small increases in fitness, may not be able to fix in response to stringent bottlenecking (Wahl and
80 Gerrish, 2001; Wahl et al., 2002). This is because these mutations will have a net reproductive rate
81 that is effectively smaller than zero after the additional mortality associated with bottlenecking is
82 taken into account. Therefore despite increasing relative fitness, they would not increase absolute
83 fitness. Put simply, strong bottlenecking is expected to also lead to the disproportionate loss of
84 weakly beneficial mutations (Wahl and Gerrish, 2001; Wahl et al., 2002), and therefore bias selection
85 to just a sub-set of genes that have large phenotypic effects, again leading to high levels of parallel
86 evolution.

87 Taking these two arguments together, it could be predicted that either intermediate bottlenecking
88 should lead to the lowest levels of parallel evolution, or alternatively that bottlenecking does not
89 affect the probability of parallel evolution. In this paper, we test the role of population bottlenecking
90 on the rate and mechanisms of adaptation using an experimental model system. We propagated
91 hundreds of populations of the bacterium *Pseudomonas fluorescens* PFO-1 in standard laboratory
92 culture medium supplemented with the antibiotic rifampicin. We manipulated the strength of daily
93 population bottlenecking over one order of magnitude (200-fold to 2000-fold reduction in
94 population density) by changing the fraction of each population that was transferred to fresh culture
95 medium on a daily basis. Crucially, the combination of a potent dose of antibiotic and population
96 bottlenecking used in our experiment ensured that populations from all bottleneck treatments could
97 only persist until the end of the experiment by evolving an increased growth rate. Thus, our
98 experiment challenged bacterial populations with 'evolutionary rescue'.

99 Previous experiments that have investigated the impact of population bottlenecking on parallel
100 evolution have focused on testing for parallelism at a phenotypic level (Lachapelle et al., 2015;

Rozen et al., 2008; Schoustra et al., 2009), by measuring divergence between populations in phenotypic traits that are closely linked to fitness. Tests on the role of bottlenecks, or even more generally population size, are rare for molecular evolution (but see Miller et al., 2011). We therefore tested our hypothesis using both phenotypic assays but also whole-genome sequencing, which would allow us to take a hierarchical approach to testing parallel evolution (Bailey et al., 2015; Dettman et al., 2012; Stern, 2013).

Materials and Methods

Strains, Culture conditions, and Antibiotics

Pseudomonas fluorescens PFO-1 was obtained from Gail Preston (Department of Plant Sciences, University of Oxford) in January 2012. Prior to experimentation it was stored at -80°C in 25% glycerol. All culturing was performed in King's B (KB) media, at 30°C with constant shaking at 250rpm. Rifampicin is an inhibitor of RNA polymerase and was stored according to the manufacturer's instructions.

Selection experiment

We used a sub-lethal dose of rifampicin as the main selection pressure. The minimum inhibitory concentration of rifampicin had been previously determined for PFO-1 under our experimental conditions (8 µg/ml), and we use 80% of this concentration in our experiment (6.4 µg/ml). This limits the ancestral growth rate to 10% of its maximum, and effectively creates a declining population, akin to evolutionary rescue conditions.

We used three different bottleneck sizes to manipulate population size. Specifically, we diluted selection lines either 1/200, 1/600, or 1/2000 into 200µl of fresh KB media containing rifampicin on a daily basis, henceforth referred to as weak, intermediate, and strong bottleneck treatments, respectively. To begin the experiment, a single colony of PFO-1 was isolated by streaking on agar. It was inoculated in 1ml of KB media, grown overnight at 30°C, and then used to found 96 replicates of the weak bottleneck treatment, 96 replicates of the intermediate bottleneck treatment, and 192 replicates of the strong bottleneck treatment. This is approximately 2×10^6 , 6×10^5 , and 2×10^5 cells, respectively, transferred at the start of the experiment. However after the initial transfer the number of cells transferred will decrease with each transfer, unless populations adapt, in which case it may be greater than these values.

As variable bottlenecks also varies the maximum number of generations per day (if all populations were to return to the same density after each transfer), we ran the selection experiment for 14, 12, and 10 days, respectively. This results in approximately 110 total generations, assuming the ancestral growth. During the experiment, every two days samples of all populations were transferred to 25% glycerol and stored at -80°C.

Sequencing

Following the selection experiment a single colony was isolated for genomic sequencing from 34 randomly chosen weak bottleneck populations, 33 randomly chosen intermediate bottleneck populations, and all 26 surviving strongly bottlenecked populations. Although sequencing a single clone ignores any within population diversity, given current coverage levels/sequencing technology it provides the most convenient way to measure between population parallelisms. Genomic DNA was extracted from these 93 clones using the Promega Genomic Wizard kits, and the protocol was otherwise performed as per the manufacturer's instructions. DNA was then quantified using the Quantifluor dsDNA system from Promega. Sequencing was conducted by the Wellcome Trust Centre for Human Genetics using HiSeq2500 and 100-bp-paired-end reads. We also sequenced three clones of the ancestor used to initiate the experiment, to detect differences from the published reference.

Bioinformatics

We analysed the sequencing data using the method first described in (San Millan et al., 2014) (see also online supplementary information). Briefly, quality-checked reads were aligned to the *PFO-1* reference sequence (NC_007492.2) using BWA. We called variants (e.g. single nucleotide polymorphisms (SNPs), large and small indels, copy number variants, inversions, translocations) using multiple tools: GATK Unified Genotyper (DePristo et al., 2011), samtools pileup (Li et al., 2009), BreakDancer (Chen et al., 2009), Pindel (Ye et al., 2009), Control-FREEC (Boeva et al., 2012). Variants were annotated using SnpEff (Cingolani et al., 2012). Information about gene function was obtained from the Pseudomonas Genome Database (Winsor et al., 2011).

Fitness assays

We use growth rate in the presence of rifampicin as a proxy for fitness in the evolved clones. Specifically, we measured the rate of exponential growth of each clone growing in the presence of the experimental dose of rifampicin. For each assay, each clone was grown overnight in KB media, diluted 1000-fold in KB-media containing rifampicin and grown overnight at 30°C with constant shaking at 250rpm. OD600 readings were taken every 20 minutes using a BioTek synergy plate reader (Winooski, VT). Each assay was replicated four times under these conditions. Assays were

performed in blocks, with each clone assayed in at least three different blocks, with two replicates per block. Blocks were standardised by subtracting the mean growth rate of six ancestral controls included within each block. We define exponential growth rate as the maximum rate of growth over six consecutive readings. Subsets of assays were repeated with 200x, 600x, and 2000x dilution during inoculation, to mimic the experimental transfer sizes. The inoculum levels of these assays did not affect the maximum rate of growth, at least at the bottleneck sizes used here.

Results

Bottleneck intensity and phenotypic evolution

Population genetics theory predicts that population bottlenecking should constrain adaptation by reducing genetic diversity and increasing the rate of population decline, which we tested in two ways. First, we assayed the proportion of populations which had gone extinct during the selection experiment, by plating samples of all populations from the end of the experiment onto agar plates lacking rifampicin, the dominant selective pressure in the experiment. These plates would therefore be permissive to any non-adapted cells still present in any populations at the end of the experiment. As predicted by theory, the strong bottleneck treatment led to high probability of extinction (86%) while more relaxed bottlenecks resulted in far less extinction (48% and 8%, respectively). Secondly, we assayed the fitness of clones from a subset of surviving populations from each treatment, by measuring the absolute fitness (i.e. growth rate) of independently evolved clones sampled at the end of the experiment (figure 1). Specifically, we assayed a single clone from each of 34 weakly bottlenecked populations, each of 33 medium bottlenecked populations, and from all 26 surviving strongly bottlenecked populations. As expected, the clones from the weak bottleneck treatment show significantly higher fitness than either of the other treatments (bonferroni-corrected t-tests: weak versus intermediate: $t = 3.78$, $DF = 65$, $P < 0.001$; weak versus strong: $t = 3.82$, $DF = 58$, $P < 0.001$; strong versus intermediate: $t = 0.16$, $DF = 57$, $P = 0.873$). Therefore even when adaptation was able to prevent population extinction, population bottlenecking constrained the efficacy of natural selection.

Evolutionary theory also predicts that bottlenecking should lead to increased divergence between populations. To test this hypothesis, we estimate the variance component independently for each treatment, fitting a model taking into both account variances between clones as well as experimental error. We find that in contrast to the theoretical expectation, variance between clones decreases with increasing bottleneck intensity (variance components; weak=0.864, intermediate = 0.208, strong = 0.017; pairwise F-tests on variance: weak vs intermediate: $F_{33,32} = 4.15$, $P < 0.001$;

weak vs strong: $F_{33,25} = 50.8$, $P < 0.001$; intermediate vs strong: $F_{32,25} = 12.2$, $P < 0.001$;). Therefore at the phenotypic level, evolution was most parallel with the strongest bottlenecks.

Bottlenecking and genome-wide divergent molecular evolution

To determine the molecular basis of adaptation, we sequenced the genome of each of the 93 clones used for the fitness assays. We also sequenced three clones from the ancestral stock, to identify differences between the starting point of our experiment and the published reference sequence. In total we identified 259 mutations across these clones (mean 2.78 mutations per clone, range 1-6), spread across 91 loci including 66 genes. Specifically we identified 174 intragenic non-synonymous SNPs, 7 synonymous intragenic SNPs, 51 intragenic indels, 22 intergenic mutations, and 5 large deletions (>50 base pairs). Parallel evolution was reasonably common, with 19 genes mutated in at least two independent clones, accounting for 71.4% of all detected mutations. Given the high levels of repeated gene use, and low proportion of synonymous mutations, it is reasonable to assume that the majority of detected mutations are at least weakly beneficial.

Interestingly, we found that the number of mutations per clone decreased with the increasing intensity of bottlenecking (generalized linear model with Poisson distribution and log-linked function: Wald Chi-square = 10.2, DF= 2, $P < 0.01$; figure 2a). Specifically, we detected significantly more mutations per clone in the weak bottleneck treatment than either of the other treatments (bonferroni-corrected t-tests: weak versus intermediate: $t = 4.65$, DF = 65, $P < 0.001$; weak versus strong: $t = 4.86$, DF = 58, $P < 0.001$; strong versus intermediate: $t = 0.23$, DF = 57, $P = 0.82$). Given that fitness evolves most rapidly in weakly bottlenecked populations, the link between population bottlenecking and the rate of molecular evolution provides further evidence to support the idea that the majority of detected mutations were beneficial. Indeed, we find that the number of mutations per clone significantly correlates with fitness, even after correcting for the effect of bottleneck size (general linear model on fitness with bottleneck intensity as a fixed factor and mutations-fixed as a covariate; bottleneck: $F_{2,89} = 4.7$, $P < 0.05$; mutations : $F_{1,89} = 7.89$, $P < 0.01$).

To test the hypothesis that population bottlenecking alters the probability of parallel evolution, we first calculate a distance matrix using Jaccard's index (Jaccard, 1908). This index is commonly used to assay parallel evolution, and measures the proportion of genetic changes in common between a pair of clones. When calculated in a pairwise manner for all clones within the same group, it provides a measure of mean within group-parallel evolution. Interestingly, bottleneck intensity did not affect the mean proportion of shared mutated genes (permutational analysis of multivariate homogeneity

of group dispersion (Anderson, 2006): $F_{2,90} = 0.988$, $P = 0.386$; figure 3a) or shared SNPS (permutational analysis of multivariate homogeneity of group dispersion: $F_{2,90} = 0.091$, $P = 0.919$; figure 3b). However, even if the level of parallel evolution is the same in all groups, it does not mean all groups are fixing the same mutations. Using permutational multivariate analysis (Anderson, 2001) of the Jaccard distance matrix (equivalent to a one-way ANOVA on univariate data) reveals that is indeed the case for both genes ($F_{2,90} = 3.67$, $P < 0.001$) and SNPs ($F_{2,90} = 1.86$, $P < 0.01$). In other words, the mean number of shared mutations is higher within groups than between them.

Evolution of major genes

To identify which genes were more likely to be mutated in particular bottleneck treatments, we decided to focus our analysis on genes which were likely to be having a large fitness effect. The dose of rifampicin used in our experiment reduced the growth rate of the ancestral clone by 90%, implying that antibiotic resistance mutations are a reasonable candidate to be strongly beneficial. In support of this idea, the two most commonly mutated genes (representing more than a third of total mutations) are both known to confer antibiotic resistance. The most commonly mutated gene was the beta-subunit of RNA polymerase (*rpoB*; 64 mutations across 59 clones), which is unsurprising since *rpoB* mutations are the major mechanism of clinical rifampicin resistance (Goldstein, 2014). The second most common target of selection was *cpxA* (39 mutations across 39 clones), a regulator of the cell envelope stress response (Dorel et al., 2006). *CpxA* regulates several efflux pumps known to be associated with antibiotic resistance (Dorel et al., 2006; Mahoney and Silhavy, 2013; Suzuki et al., 2014; Weatherspoon-Griffin et al., 2014), although not previously involved in resistance to rifampicin. Only 8 of the 93 clones lacked mutations in either of these genes, suggesting these two genes are the two major mechanisms of resistance.

Intriguingly, bottleneck intensity had a significant impact on the molecular mechanisms of antibiotic resistance (chi-squared test on proportion of clones with mutations in *rpoB* and *cpxA*: chi-squared = 18.75, $P < 0.005$; figure 4). Under both weak and strong bottlenecking, *rpoB* mutations predominated, while in contrast *cpxA* mutations were more common at intermediate bottlenecks. However, in general, the intermediate bottleneck treatment shows less bias towards any one particular mechanism, and consequently shows the highest diversity of resistance mechanisms and the lowest probability of parallel evolution (Simpson's Index of diversity of resistance mechanisms: strong bottlenecks = 0.524, intermediate bottlenecks = 0.681, weak bottlenecks = 0.569). To explain why the diversity of mechanisms was highest at intermediate bottlenecks, we reanalysed the fitness data with respect to resistance mechanisms. As shown, clones with mutations in *rpoB* are significantly fitter than clones with mutations in *cpxA* (one-way ANOVA on "clones with mutations in

rpoB but not *cpxA*” versus “clones with mutations in *cpxA* but not *rpoB*”: $F_{1,70} = 21.90$, $P < 0.001$; figure 5). This is still true if the analysis is limited to clones possessing only a single mutation in *rpoB* (mean fitness = 0.798 ± 0.078 , $n=3$) or *cpxA* (mean fitness = 0.499 ± 0.104 , $n=4$) and no other mutations anywhere else in their genome (one-way ANOVA on clones with only a mutation in *rpoB* versus clones with only a mutation in *cpxA*: $F_{1,5} = 6.73$, $P < 0.05$). Therefore both strong and weak bottlenecks lead to a bias towards the more strongly beneficial mutations, and consequently divergence is not greatest at the strongest bottleneck. However for intermediate bottlenecks there is less bias towards strongly beneficial mutations, and consequently greater diversity.

Epistasis between major genes

From figure 5, it can be seen that clones possessing mutations in both *rpoB* and *cpxA* have would appear to have intermediate fitness compared to clones only possessing mutations in one of these two genes. However, the fitness of clones with mutations in *rpoB* does not significantly differ if the clone also has a mutation in *cpxA* (independent sample t-test: clones with *rpoB* versus clones with both *rpoB* and *cpxA*: $t = 0.867$ $DF = 57$, $P = 0.365$). As both *rpoB* and *cpxA* mutations increase growth rate, this demonstrates negative epistatic fitness effects for these two genes. In other words, the fitness benefit of having mutations in both genes is less than expected from the fitness effects of mutations in either of the two genes alone.

Evolution within major genes

Given that most of the dynamics of fitness were being driven by mutations in just two genes, we tested whether bottleneck intensity was affecting which nucleotides were being selected within these genes. Previous work has shown that different SNPs in *rpoB* can have different effects on bacterial fitness by altering both the level of rifampicin resistance, as well as competitive ability and growth rate. Given this diversity of fitness effects, as well as the large number of possible mutations in *rpoB*, different spectra of *rpoB* substitutions might be expected to evolve in response to varying bottleneck intensity. However we find no evidence that this was occurring, suggesting selection was weaker within genes than between genes (chi-squared test on distribution of *rpoB* SNPs: chi-squared = 30.40, $P=0.2$; figure 6a). A similar argument can intuitively be expected to apply to mutations within *cpxA*, the second most common target of selection. However again we find no evidence of selection favouring differing SNPs within different bottleneck treatments (chi-squared test on distribution of *cpxA* SNPs: chi-squared = 6.53, $P=0.3$; figure 6b).

Discussion

Bottlenecks are a common and unavoidable aspect of the demography of most organisms. In this experiment, we studied the phenotypic and genetic consequences of population bottlenecking during adaptation. In our experiment, bottlenecking had a profound impact on the likelihood of adaptation, on the rate of fitness evolution, and on the rate of substitution of mutations. These effects are simple to understand using conventional population genetics reasoning: bottlenecking reduces effective population size, resulting in a greater rate of loss of beneficial mutations to genetic drift.

However the impact of population bottlenecking on patterns of molecular evolution is more subtle. Our experimental design imposed strong selection for antibiotic resistance and *P. fluorescens* evolved resistance predominantly using two different genes. One of these appears to be strongly beneficial and is associated with relatively high fitness (*rpoB*), while the other is only weakly beneficial and results in more modest fitness gains (*cpxA*). *rpoB* mutations preventing rifampicin from binding to its target domain (Campbell et al., 2001) and this is associated with elevated rifampicin resistance and decreased competitive ability, due to the pleiotropic effects of resistance mutations (MacLean and Buckling, 2009; Qi et al., 2014). The role of *cpxA* mutations in rifampicin resistance is less well understood, but regulates many efflux pumps known to confer antibiotic resistance (Mahoney and Silhavy, 2013; Suzuki et al., 2014; Weatherspoon-Griffin et al., 2014). Under weak bottlenecking, the higher population size meant an increased probability of two beneficial mutations being present in each population, and competition between independent beneficial mutations therefore favoured *rpoB* over *cpxA*. In contrast, under stronger bottlenecking the higher absolute fitness of *rpoB* mutations reduced the likelihood of stochastic loss at each bottleneck event, again resulting in a disproportionate loss of weakly beneficial *cpxA* alleles. Collectively, these biases resulted in a high likelihood of parallel evolution under either intense or weak bottlenecking. Interestingly, this is an effect that can only be understood by considering the impact of beneficial mutations on absolute fitness, and not relative fitness alone, a distinction often emphasized by evolutionary rescue theory (reviewed in Bell, 2013).

However, more broadly we found the level of parallel molecular evolution at a genome wide scale was insensitive to population bottlenecking. A major contributor to this is the sheer diversity of evolution at a genome wide scale, even under the relatively simple laboratory conditions we employed. Although we found considerable evidence for parallel evolution, nearly 30% of all mutations were in genes only mutated in a single clone. Coupled with this, much of the variation in fitness can be attributed to a small subset of genes directly involved in adaptation to rifampicin.

Although the dynamics of these major genes were significantly affected by bottlenecking, statistically this effect is hidden by the diffuse nature of evolution across the rest of the genome. In other words, our data suggests that most mutations were only weakly beneficial at best, and consequently their dynamics were not influenced by the intensities of bottlenecking imposed by our experiment.

Similarly, we also failed to detect any effect of population bottlenecking on the frequency of particular SNPs within major genes. This is likely because in most cases there will be greater variation in the fitness effects of mutations in different genes rather than between different mutations within the same gene. Consequently selection will be more evident between genes than within them. This is not to say that there cannot be considerable diversity in the phenotypic effects of different nucleotide substitutions within a particular gene (for example Bataillon et al., 2011; Kassen and Bataillon, 2006; MacLean and Buckling, 2009). However these within-gene differences are likely to be most important when adaptation is only possible, or at least most likely, through a single gene, such as the strong selection imposed by clinical doses of antibiotics.

Adaptation is often predicted to be most repeatable at large population sizes, because clonal interference is most prevalent in large populations. Indeed, this is supported by several publications which measured repeatability at the phenotypic level (Lachapelle et al., 2015; Rozen et al., 2008; Schoustra et al., 2009). Parts of our results are in agreement with this line of reasoning, such as the high levels of parallel evolution in the weak bottlenecking treatment. However, in our experiment, extinction was just as effective as clonal interference in limiting which mutations could fix, and consequently we did find limited evidence of repeatability increasing with population size. Therefore our results suggest caution should be taken in assuming large population size will always lead to the most repeatable adaptation, particularly in contexts where clonal interference is unlikely to be the only factor influencing adaptation.

In this manuscript, we only utilised a single relatively low concentration of rifampicin. If a stronger concentration had been used, it is likely that parallel evolution would have been more common across all bottleneck treatments, as mutations in few genes can result in high level antibiotic resistance (Blair et al., 2015). This is particularly true for rifampicin, as almost all clinical rifampicin resistance mutations are within *rpoB* (Goldstein, 2014). Similarly, we only used three bottleneck intensities, which only capture a fraction of the bottleneck sizes which are likely to occur in clinical pathogens. Hopefully future work will measure the intensity of bottlenecking experienced by bacterial pathogens *in vivo* due to transmission and host immune responses, and thereby provide a guide for future *in-vitro* investigation.

Parallel evolution is common in bacteria in both natural (for example Lieberman et al., 2011; Marvig et al., 2015a; Marvig et al., 2015b) and laboratory environments (for example Vogwill et al., 2014; Webber et al., 2015). Given the difficulties associated with applying many classical tests for positive selection to bacterial populations (Kryazhimskiy and Plotkin, 2008), it has been suggested that parallel evolution should be used to test for positive selection (Farhat et al., 2013). Our results suggest both optimism and caution towards this approach. Theoretical reasoning and previous experiments suggests that this may be a dangerous approach to use, as population demography might play an important role in shaping the likelihood of parallel evolution (Lachapelle et al., 2015; Rozen et al., 2008; Schoustra et al., 2009). However, our results suggest that genome-wide patterns of parallelism may be relatively independent from population bottlenecking, which is likely to be a key feature of the demography of many bacteria, especially bacterial pathogens. An important caveat is that population bottlenecking might have a strong effect on patterns of parallel evolution in genes that are likely to be under strong selection, such as antibiotic resistance genes. More generally, the causes of parallel evolution in bacteria remain unclear. Our data suggests that competition between beneficial mutations could be a predominant factor, but equally our data suggests that high levels of parallel molecular evolution can still happen in the absence of this competition.

Author contributions

Obtained funding: RCM; conceived the experiment: TV, DRG, RCM; designed the experiment: TV, RLP, RCM; performed the experimental work: TV, RLP; performed the bioinformatics: DRG; analysed the data: TV; wrote the manuscript: TV, RCM; commented on the manuscript: TV, RLP, DRG, RCM.

Competing interests

We have no competing interests.

Acknowledgements

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement no. 281591 and from the Royal Society.

References

388

- 389 Anderson, M.J., 2001. A new method for non-parametric multivariate analysis of variance. *Austral*
 390 *Ecology* 26, 32-46.
- 391 Anderson, M.J., 2006. Distance-Based Tests for Homogeneity of Multivariate Dispersions. *Biometrics*
 392 62, 245-253.
- 393 Bailey, S.F., Rodrigue, N., Kassen, R., 2015. The Effect of Selection Environment on the Probability of
 394 Parallel Evolution. *Molecular Biology and Evolution*.
- 395 Bataillon, T., Zhang, T., Kassen, R., 2011. Cost of Adaptation and Fitness Effects of Beneficial
 396 Mutations in *Pseudomonas fluorescens*. *Genetics* 189, 939-949.
- 397 Bell, G., 2013. Evolutionary rescue and the limits of adaptation. *Philosophical Transactions of the*
 398 *Royal Society of London B: Biological Sciences* 368.
- 399 Blair, J.M.A., Webber, M.A., Baylay, A.J., Ogbolu, D.O., Piddock, L.J.V., 2015. Molecular mechanisms
 400 of antibiotic resistance. *Nat Rev Micro* 13, 42-51.
- 401 Boeva, V., Popova, T., Bleakley, K., Chiche, P., Cappel, J., Schleiermacher, G., Janoueix-Lerosey, I.,
 402 Delattre, O., Barillot, E., 2012. Control-FREEC: a tool for assessing copy number and allelic
 403 content using next-generation sequencing data. *Bioinformatics* 28, 423-425.
- 404 Breidenstein, E.B.M., Khaira, B.K., Wiegand, I., Overhage, J., Hancock, R.E.W., 2008. Complex
 405 Ciprofloxacin Resistome Revealed by Screening a *Pseudomonas aeruginosa* Mutant Library
 406 for Altered Susceptibility. *Antimicrobial Agents and Chemotherapy* 52, 4486-4491.
- 407 Campbell, E.A., Korzheva, N., Mustaev, A., Murakami, K., Nair, S., Goldfarb, A., Darst, S.A., 2001.
 408 Structural Mechanism for Rifampicin Inhibition of Bacterial RNA Polymerase. *Cell* 104, 901-
 409 912.
- 410 Chen, K., Wallis, J.W., McLellan, M.D., Larson, D.E., Kalicki, J.M., Pohl, C.S., McGrath, S.D., Wendl,
 411 M.C., Zhang, Q., Locke, D.P., Shi, X., Fulton, R.S., Ley, T.J., Wilson, R.K., Ding, L., Mardis, E.R.,
 412 2009. BreakDancer: an algorithm for high-resolution mapping of genomic structural
 413 variation. *Nat Meth* 6, 677-681.
- 414 Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., Ruden, D.M.,
 415 2012. A program for annotating and predicting the effects of single nucleotide
 416 polymorphisms, SnpEff. *Fly* 6, 80-92.
- 417 Colegrave, N., 2002. Sex releases the speed limit on evolution. *Nature* 420, 664-666.
- 418 Conte, G.L., Arnegard, M.E., Peichel, C.L., Schluter, D., 2012. The probability of genetic parallelism
 419 and convergence in natural populations. *Proceedings of the Royal Society of London B:*
 420 *Biological Sciences*.
- 421 Dahan, R.A., Duncan, R.P., Wilson, A.C., Dávalos, L.M., 2015. Amino acid transporter expansions
 422 associated with the evolution of obligate endosymbiosis in sap-feeding insects (Hemiptera:
 423 Sternorrhyncha). *BMC Evolutionary Biology* 15, 1-11.
- 424 de Visser, J.A.G.M., Krug, J., 2014. Empirical fitness landscapes and the predictability of evolution.
 425 *Nat Rev Genet* 15, 480-490.
- 426 De Visser, J.A.G.M., Rozen, D.E., 2005. Limits to adaptation in asexual populations. *Journal of*
 427 *Evolutionary Biology* 18, 779-788.
- 428 DePristo, M.A., Banks, E., Poplin, R., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis, A.A., del
 429 Angel, G., Rivas, M.A., Hanna, M., McKenna, A., Fennell, T.J., Kernysky, A.M., Sivachenko,
 430 A.Y., Cibulskis, K., Gabriel, S.B., Altshuler, D., Daly, M.J., 2011. A framework for variation
 431 discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* 43, 491-
 432 498.
- 433 Dettman, J.R., Rodrigue, N., Melnyk, A.H., Wong, A., Bailey, S.F., Kassen, R., 2012. Evolutionary
 434 insight from whole-genome sequencing of experimentally evolved microbes. *Molecular*
 435 *Ecology* 21, 2058-2077.
- 436 Dorel, C., Lejeune, P., Rodrigue, A., 2006. The Cpx system of *Escherichia coli*, a strategic signaling
 437 pathway for confronting adverse conditions and for settling biofilm communities? *Research*
 438 *in Microbiology* 157, 306-314.

- 439 Dötsch, A., Becker, T., Pommerenke, C., Magnowska, Z., Jänsch, L., Häussler, S., 2009. Genomewide
440 Identification of Genetic Determinants of Antimicrobial Drug Resistance in *Pseudomonas*
441 *aeruginosa*. *Antimicrobial Agents and Chemotherapy* 53, 2522-2531.
- 442 Fajardo, A., Martínez-Martín, N., Mercadillo, M., Galán, J.C., Ghysels, B., Matthijs, S., Cornelis, P.,
443 Wiehlmann, L., Tümmeler, B., Baquero, F., Martínez, J.L., 2008. The Neglected Intrinsic
444 Resistome of Bacterial Pathogens. *PLoS ONE* 3, e1619.
- 445 Farhat, M.R., Shapiro, B.J., Kieser, K.J., Sultana, R., Jacobson, K.R., Victor, T.C., Warren, R.M.,
446 Streicher, E.M., Calver, A., Sloutsky, A., Kaur, D., Posey, J.E., Plikaytis, B., Oggioni, M.R.,
447 Gardy, J.L., Johnston, J.C., Rodrigues, M., Tang, P.K.C., Kato-Maeda, M., Borowsky, M.L.,
448 Muddukrishna, B., Kreiswirth, B.N., Kurepina, N., Galagan, J., Gagneux, S., Birren, B., Rubin,
449 E.J., Lander, E.S., Sabeti, P.C., Murray, M., 2013. Genomic analysis identifies targets of
450 convergent positive selection in drug-resistant *Mycobacterium tuberculosis*. *Nat Genet* 45,
451 1183-1189.
- 452 Gerrish, P.J., Lenski, R.E., 1998. The fate of competing beneficial mutations in an asexual population.
453 *Genetica* 102, 127-144.
- 454 Goldstein, B.P., 2014. Resistance to rifampicin: a review. *J Antibiot* 67, 625-630.
- 455 Gomulkiewicz, R., Holt, R.D., 1995. When does Evolution by Natural Selection Prevent Extinction?
456 *Evolution* 49, 201-207.
- 457 Hill, W.G., Robertson, A., 1966. The effect of linkage on limits to artificial selection. *Genetics*
458 *Research* 8, 269-294.
- 459 Jaccard, P., 1908. Nouvelles recherches sur la distribution florale. *Bulletin de la Société Vaudense des*
460 *Sciences Naturelles* 44, 223-270.
- 461 Kassen, R., Bataillon, T., 2006. Distribution of fitness effects among beneficial mutations before
462 selection in experimental populations of bacteria. *Nat Genet* 38, 484-488.
- 463 Kryazhimskiy, S., Plotkin, J.B., 2008. The Population Genetics of dN/dS. *PLoS Genet* 4, e1000304.
- 464 Lachapelle, J., Reid, J., Colegrave, N., 2015. Repeatability of adaptation in experimental populations
465 of different sizes. *Proceedings of the Royal Society of London B: Biological Sciences* 282.
- 466 Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R.,
467 Subgroup, G.P.D.P., 2009. The Sequence Alignment/Map format and SAMtools.
468 *Bioinformatics* 25, 2078-2079.
- 469 Lieberman, T.D., Michel, J.-B., Aingaran, M., Potter-Bynoe, G., Roux, D., Davis, M.R., Skurnik, D.,
470 Leiby, N., LiPuma, J.J., Goldberg, J.B., McAdam, A.J., Priebe, G.P., Kishony, R., 2011. Parallel
471 bacterial evolution within multiple patients identifies candidate pathogenicity genes. *Nat*
472 *Genet* 43, 1275-1280.
- 473 Liu, A., Tran, L., Becket, E., Lee, K., Chinn, L., Park, E., Tran, K., Miller, J.H., 2010. Antibiotic Sensitivity
474 Profiles Determined with an *Escherichia coli* Gene Knockout Collection: Generating an
475 Antibiotic Bar Code. *Antimicrobial Agents and Chemotherapy* 54, 1393-1403.
- 476 Losos, J.B., 2011. CONVERGENCE, ADAPTATION, AND CONSTRAINT. *Evolution* 65, 1827-1840.
- 477 MacLean, R.C., Buckling, A., 2009. The Distribution of Fitness Effects of Beneficial Mutations in
478 *Pseudomonas aeruginosa*. *PLoS Genet* 5, e1000406.
- 479 Mahoney, T.F., Silhavy, T.J., 2013. The Cpx Stress Response Confers Resistance to Some, but Not All,
480 Bactericidal Antibiotics. *Journal of Bacteriology* 195, 1869-1874.
- 481 Marvig, R.L., Dolce, D., Sommer, L.M., Petersen, B., Ciofu, O., Campana, S., Molin, S., Taccetti, G.,
482 Johansen, H.K., 2015a. Within-host microevolution of *Pseudomonas aeruginosa* in Italian
483 cystic fibrosis patients. *BMC Microbiology* 15, 1-13.
- 484 Marvig, R.L., Sommer, L.M., Molin, S., Johansen, H.K., 2015b. Convergent evolution and adaptation
485 of *Pseudomonas aeruginosa* within patients with cystic fibrosis. *Nat Genet* 47, 57-64.
- 486 McCutcheon, J.P., McDonald, B.R., Moran, N.A., 2009. Convergent evolution of metabolic roles in
487 bacterial co-symbionts of insects. *Proceedings of the National Academy of Sciences* 106,
488 15394-15399.

- Miller, C.R., Joyce, P., Wichman, H.A., 2011. Mutational Effects and Population Dynamics During Viral Adaptation Challenge Current Models. *Genetics* 187, 185-202.
- Ord, T.J., Summers, T.C., 2015. Repeated evolution and the impact of evolutionary history on adaptation. *BMC Evolutionary Biology* 15, 1-12.
- Qi, Q., Preston, G.M., MacLean, R.C., 2014. Linking System-Wide Impacts of RNA Polymerase Mutations to the Fitness Cost of Rifampin Resistance in *Pseudomonas aeruginosa*. *mBio* 5.
- Rozen, D.E., Habets, M.G.J.L., Handel, A., de Visser, J.A.G.M., 2008. Heterogeneous Adaptive Trajectories of Small Populations on Complex Fitness Landscapes. *PLoS ONE* 3, e1715.
- San Millan, A., Peña-Miller, R., Toll-Riera, M., Halbert, Z.V., McLean, A.R., Cooper, B.S., MacLean, R.C., 2014. Positive selection and compensatory adaptation interact to stabilize non-transmissible plasmids. *Nat Commun* 5.
- Schoustra, S.E., Bataillon, T., Gifford, D.R., Kassen, R., 2009. The Properties of Adaptive Walks in Evolving Populations of Fungus. *PLoS Biol* 7, e1000250.
- Stern, D.L., 2013. The genetic causes of convergent evolution. *Nat Rev Genet* 14, 751-764.
- Suzuki, S., Horinouchi, T., Furusawa, C., 2014. Prediction of antibiotic resistance by gene expression profiles. *Nature Communications* 5, 5792.
- Szendro, I.G., Franke, J., de Visser, J.A.G.M., Krug, J., 2013. Predictability of evolution depends nonmonotonically on population size. *Proceedings of the National Academy of Sciences* 110, 571-576.
- Vogwill, T., Kojadinovic, M., Furió, V., MacLean, R.C., 2014. Testing the Role of Genetic Background in Parallel Evolution Using the Comparative Experimental Evolution of Antibiotic Resistance. *Molecular Biology and Evolution*.
- Wahl, L.M., Gerrish, P.J., 2001. THE PROBABILITY THAT BENEFICIAL MUTATIONS ARE LOST IN POPULATIONS WITH PERIODIC BOTTLENECKS. *Evolution* 55, 2606-2610.
- Wahl, L.M., Gerrish, P.J., Saika-Voivod, I., 2002. Evaluating the Impact of Population Bottlenecks in Experimental Evolution. *Genetics* 162, 961-971.
- Weatherspoon-Griffin, N., Yang, D., Kong, W., Hua, Z., Shi, Y., 2014. The CpxR/CpxA Two-component Regulatory System Up-regulates the Multidrug Resistance Cascade to Facilitate *Escherichia coli* Resistance to a Model Antimicrobial Peptide. *The Journal of Biological Chemistry* 289, 32571-32582.
- Webber, M.A., Whitehead, R.N., Mount, M., Loman, N.J., Pallen, M.J., Piddock, L.J.V., 2015. Parallel evolutionary pathways to antibiotic resistance selected by biocide exposure. *Journal of Antimicrobial Chemotherapy*.
- Wernegreen, J.J., 2002. Genome evolution in bacterial endosymbionts of insects. *Nat Rev Genet* 3, 850-861.
- Winsor, G.L., Lam, D.K.W., Fleming, L., Lo, R., Whiteside, M.D., Yu, N.Y., Hancock, R.E.W., Brinkman, F.S.L., 2011. *Pseudomonas* Genome Database: improved comparative analysis and population genomics capability for *Pseudomonas* genomes. *Nucleic Acids Research* 39, D596-D600.
- Wood, T.E., Burke, J.M., Rieseberg, L.H., 2005. Parallel genotypic adaptation: when evolution repeats itself. *Genetica* 123, 157-170.
- Ye, K., Schulz, M.H., Long, Q., Apweiler, R., Ning, Z., 2009. Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. *Bioinformatics* 25, 2865-2871.

Figure legends

536 Figure 1: The impact of population bottlenecking on adaptation. Each symbol indicates the fitness of
537 independently evolved clones, as measured by growth rate in the presence of rifampicin relative to
538 the ancestor growing in the presence of rifampicin.

539 Figure 2: The impact of bottlenecking on the rate of molecular adaptation. Panel (a) shows the
540 number of mutations detected per clone as a function of bottlenecking intensity (weak bottleneck:
541 mean (\pm st. err.) = 3.65 ± 0.17 ; intermediate: mean (\pm st. err.) = 2.48 ± 0.19 ; strong: mean (\pm st. err.) =
542 2.42 ± 0.19). Plotted points in panel (b) show the fitness of independently evolved clones as a
543 function of the number of mutations acquired during the experiment. Fitness was measured as
544 relative growth rate in the presence of rifampicin, and fitness increases with mutation number.

545 Figure 3: The impact of population bottlenecking on parallel evolution. Parallelism was measured as
546 the mean proportion of shared mutations between pairs of clones that evolved under the same
547 bottlenecking treatment using the Jaccard index. Panel (a) shows parallel evolution at the level of
548 genes, and panel (b) shows parallelism at the level of individual SNPs.

549 Figure 4: The impact of bottlenecking on resistance mechanisms. This figure is a heat map showing
550 the frequency of mutations in *rpoB* and *cpxA* across bottlenecking treatments. *RpoB* is more
551 common with either weak or strong bottlenecks, while *cpxA* is most common at intermediate
552 bottlenecks. The intermediate treatment shows the least bias to any one mechanism, and
553 consequently the highest diversity of mechanisms.

554 Fig 5: Fitness effects of alternative resistance mechanisms. Bars show the mean (\pm s.e.m) fitness of
555 clones according to resistance mechanism and bottlenecking intensity.

556 Figure 6: The impact of bottlenecking on evolution within resistance genes. This figure shows a heat
557 map of the frequency of SNPs within (a) *rpoB* and (b) *cpxA* across bottlenecking treatments.

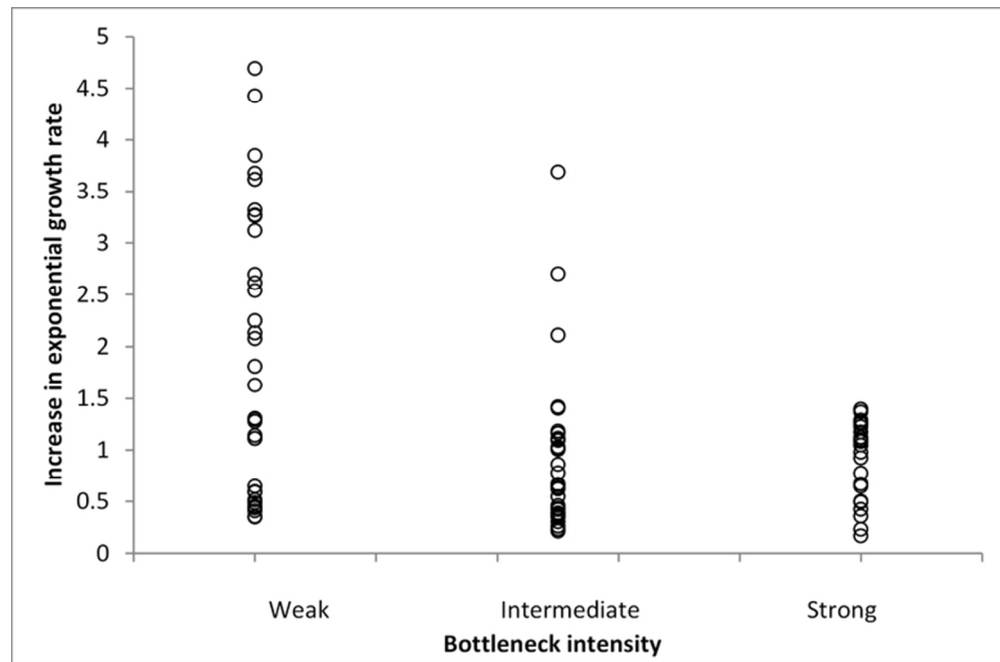


Figure 1: The impact of population bottlenecks on adaptation. Each symbol indicates the fitness of independently evolved clones, as measured by growth rate in the presence of rifampicin relative to the ancestor growing in the presence of rifampicin.
65x43mm (300 x 300 DPI)

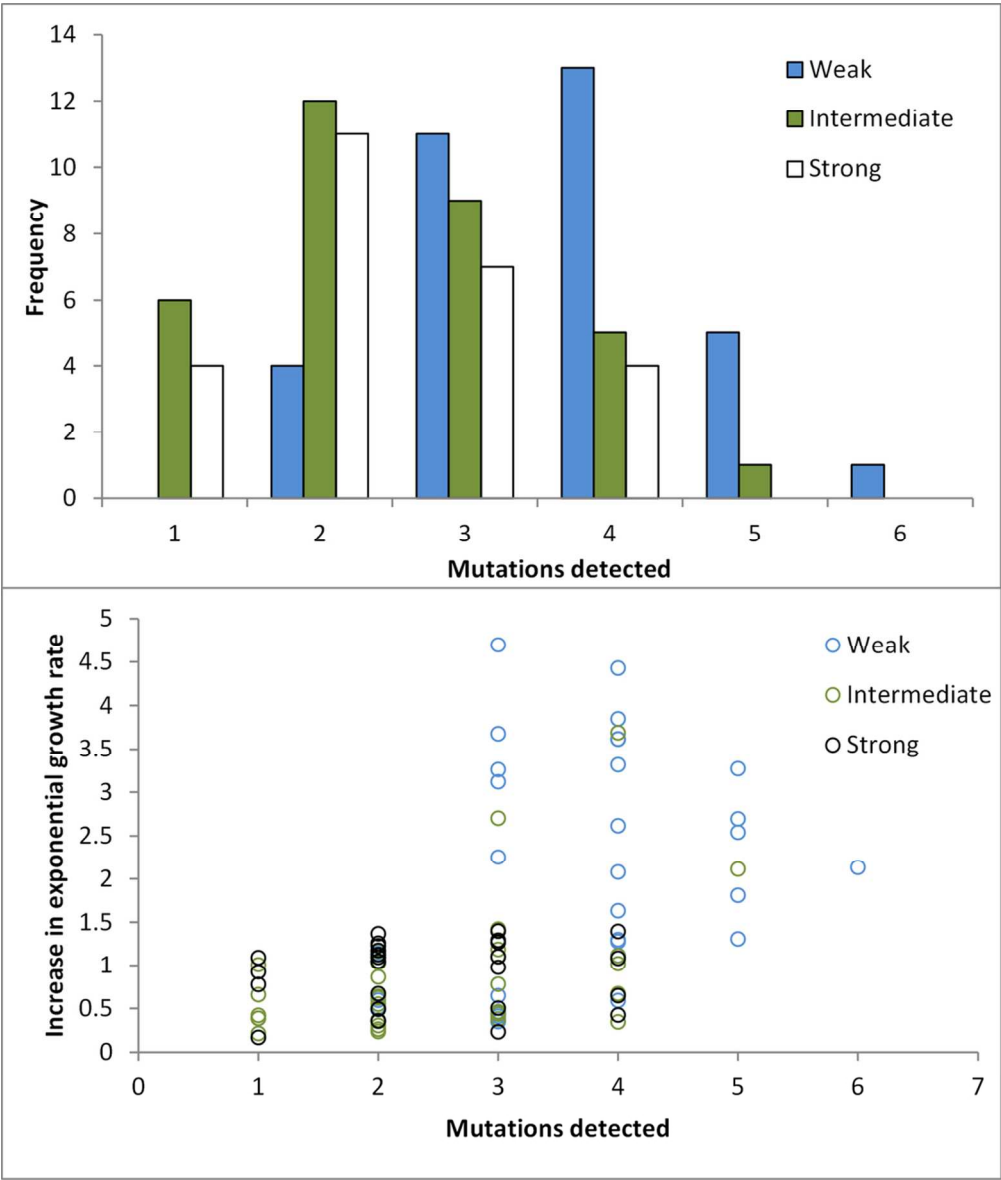


Figure 2: The impact of bottlenecking on the rate of molecular adaptation. Panel (a) shows the number of mutations detected per clone as a function of bottlenecking intensity (weak bottleneck: mean (\pm st. err.) = 3.65 ± 0.17 ; intermediate: mean (\pm st. err.) = 2.48 ± 0.19 ; strong: mean (\pm st. err.) = 2.42 ± 0.19). Plotted points in panel (b) show the fitness of independently evolved clones as a function of the number of mutations acquired during the experiment. Fitness was measured as relative growth rate in the presence of rifampicin, and fitness increases with mutation number.

93x110mm (300 x 300 DPI)

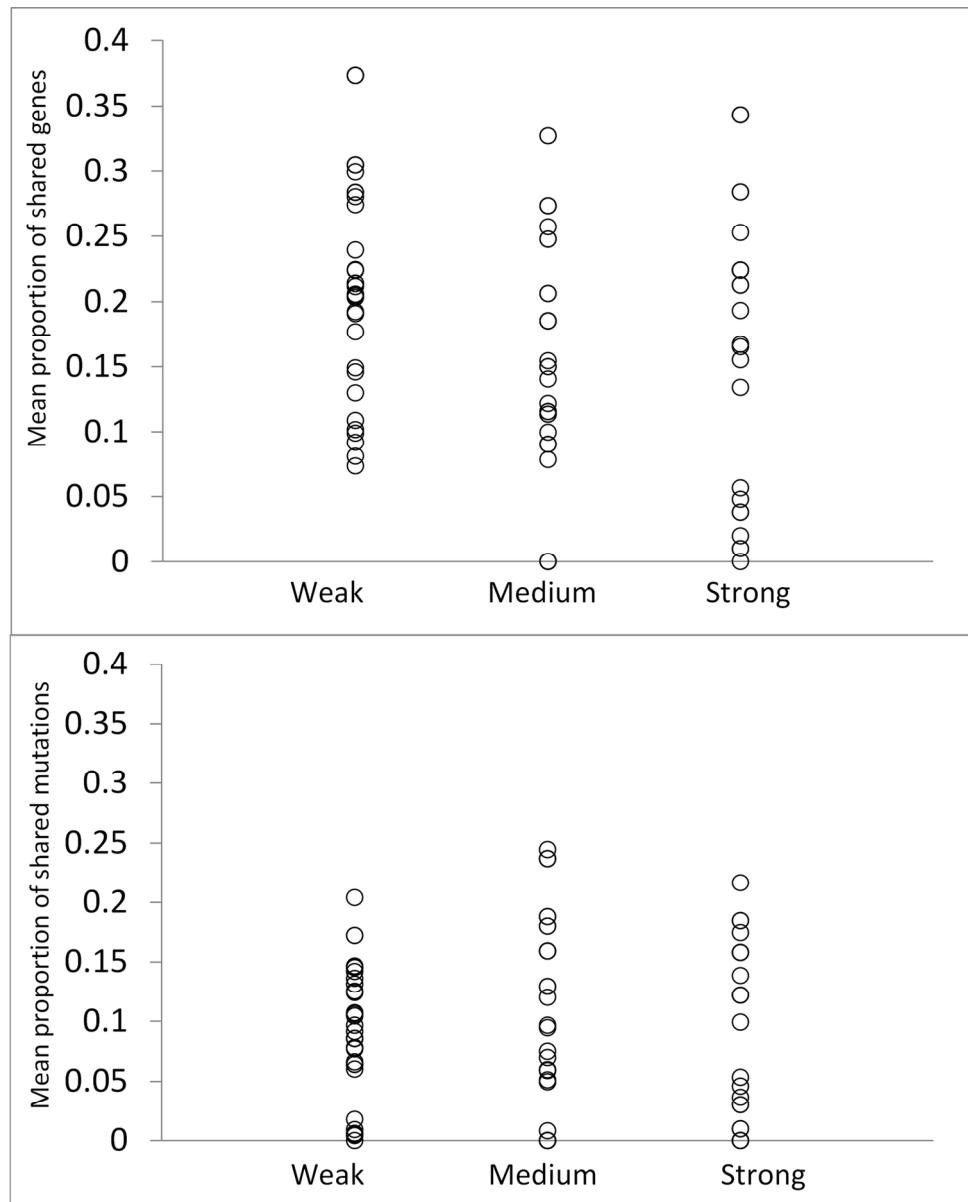


Figure 3: The impact of population bottlenecks on parallel evolution. Parallelism was measured as the mean proportion of shared mutations between pairs of clones that evolved under the same bottlenecking treatment using the Jaccard index. Panel (a) shows parallel evolution at the level of genes, and panel (b) shows parallelism at the level of individual SNPs.

135x168mm (300 x 300 DPI)

	rpoB	cpxA	Both rpoB and cpxA	Neither rpoB or cpxA
Weak	0.588	0.206	0.206	0.000
Intermediate	0.273	0.455	0.152	0.121
Strong	0.654	0.154	0.038	0.154

Figure 4: The impact of bottlenecking on resistance mechanisms. This figure is a heat map showing the frequency of mutations in rpoB and cpxA across bottlenecking treatments. RpoB is more common with either weak or strong bottlenecks, while cpxA is most common at intermediate bottlenecks. The intermediate treatment shows the least bias to any one mechanism, and consequently the highest diversity of mechanisms.

96x31mm (300 x 300 DPI)

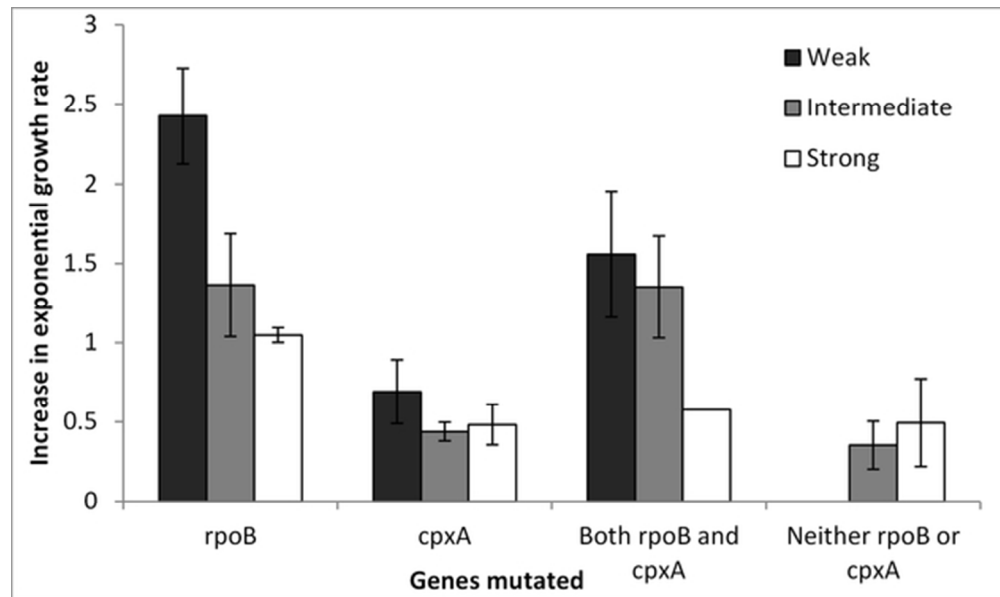


Fig 5: Fitness effects of alternative resistance mechanisms. Bars show the mean (\pm s.e.m) fitness of clones according to resistance mechanism and bottlenecking intensity.
53x31mm (300 x 300 DPI)

	Bottleneck Intensity		
Amino acid change	Weak	Int.	Strong
Gln766Leu	0.04	0.00	0.00
Asn765Ser	0.00	0.00	0.05
Ile577Leu	0.04	0.00	0.00
Thr558Ala	0.00	0.06	0.00
Gly542Cys	0.00	0.00	0.05
Gly539Ala	0.00	0.06	0.00
Leu538Val	0.00	0.06	0.05
Ser536Tyr	0.07	0.00	0.00
Asp521Val	0.00	0.06	0.05
Asp521Gly	0.74	0.63	0.67
Ser517Pro	0.00	0.06	0.00
Ser514Arg	0.00	0.00	0.05
His452Arg	0.04	0.00	0.00
Gln152Leu	0.00	0.06	0.10
Gln152Arg	0.07	0.00	0.00

	Bottleneck Intensity		
Amino acid change	Weak	Int.	Strong
Ala407Glu	0.00	0.05	0.00
Arg239His	0.86	0.85	1.00
Leu230Ser	0.00	0.05	0.00
Gln229Leu	0.07	0.00	0.00
Thr173Pro	0.00	0.05	0.00
Leu80Pro	0.07	0.00	0.00

Figure 6: The impact of bottlenecking on evolution within resistance genes. This figure shows a heat map of the frequency of SNPs within (a) *rpoB* and (b) *cpxA* across bottlenecking treatments.
95x137mm (300 x 300 DPI)