Testing the Role of Genetic Background in Parallel Evolution Using the Comparative Experimental Evolution of Antibiotic Resistance

Tom Vogwill,*,‡,[1] Mila Kojadinovic,‡,[2] Victoria Furió,[1] and R. Craig MacLean*,[1]

[1]Department of Zoology, University of Oxford, Oxford, United Kingdom

[2]CNRS, Aix-Marseille Université, Laboratoire de Bioénergétique et Ingénierie des Protéines, UMR 7281, IMM, Marseille, France

***Corresponding author:** E-mail: craig.maclean@zoo.ox.ac.uk; tom.vogwill@zoo.ox.ac.uk.

‡These authors contributed equally to this work.

Associate editor: Daniel Falush

Abstract

Parallel evolution is the independent evolution of the same phenotype or genotype in response to the same selection pressure. There are examples of parallel molecular evolution across divergent genetic backgrounds, suggesting that genetic background may not play an important role in determining the outcome of adaptation. Here, we measure the influence of genetic background on phenotypic and molecular adaptation by combining experimental evolution with comparative analysis. We selected for resistance to the antibiotic rifampicin in eight strains of bacteria from the genus *Pseudomonas* using a short term selection experiment. Adaptation occurred by 47 mutations at conserved sites in *rpoB*, the target of rifampicin, and due to the high diversity of possible mutations the probability of within-strain parallel evolution was low. The probability of between-strain parallel evolution was only marginally lower, because different strains substituted similar *rpoB* mutations. In contrast, we found that more than 30% of the phenotypic variation in the growth rate of evolved clones was attributable to among-strain differences. Parallel molecular evolution across strains resulted in divergent phenotypic evolution because *rpoB* mutations had different effects on growth rate in different strains. This study shows that genetic divergence between strains constrains parallel phenotypic evolution, but had little detectable impact on the molecular basis of adaptation in this system.

Key words: **antibiotic resistance, experimental evolution, comparative biology, adaptation, *Pseudomonas*, selection experiment.**

Article

Introduction

Although adaptation ultimately depends on the stochastic appearance of beneficial mutations, there are many examples of repeated adaptation from independent populations, lineages, or species. These are often referred to as homoplasies, and can be divided into whether they are examples of parallel, convergent, or reversal evolution (Rokas and Carroll 2008). Parallel evolution can be defined as repeated adaptation from independently evolved populations in response to similar selection pressures, where the ancestral state is initially conserved across populations (Zhang and Kumar 1997). There are examples of this from both the real world and from laboratory experiments, particular for antibiotic resistance (e.g., Jatsenko et al. 2010; Lieberman et al. 2011; Wong and Kassen 2011; Wong et al. 2012; Dettman et al. 2013) but also for many other traits (e.g., Woods et al. 2006; Ostrowski et al. 2008; Rokas and Carroll 2008; Lieberman et al. 2011; Dettman et al. 2013). Interestingly, many of these examples of parallel molecular evolution are occurring across genetic backgrounds—in other words, parallel evolution where the target of selection is conserved, but there are otherwise many differences between the genomes of these organisms. This suggests that genetic background plays a relatively unimportant role in determining the outcome of adaptation at the molecular level, but direct tests of this idea are lacking. The impact of parallel molecular evolution on phenotypic adaptation is also unclear. There is growing evidence that epistasis is pervasive (Costanzo et al. 2010; Khan et al. 2011; Breen et al. 2012), suggesting that parallel molecular evolution may lead to divergent phenotypic evolution across genetic backgrounds, but tests of this idea are also lacking.

The lack of estimates of the influence of genetic background on parallel evolution can be at least partially explained by the different methods which have used to study parallelism within versus between genetic backgrounds (Wood et al. 2005). Parallel evolution within a particular genetic background has often been experimentally tested using laboratory or greenhouse experimental evolution (reviewed in Wood et al. 2005). In contrast, parallel evolution between different genetic backgrounds has generally been explored using phylogenetic or other comparative methods (reviewed in Wood et al. 2005). There have been few attempts to incorporate comparative biology into experimental evolution (but see Wichman et al. 2000; Holder and Bull 2001; Bollback and Huelsenbeck 2009; Nguyen et al. 2012). By integrating these two differing methodologies, it is possible to directly test the role of genetic background in determining the trajectory of adaptation. If genetic differences between backgrounds play

an inconsequential role in adaptation, we would expect the same selection pressure to result in very similar outcomes in different backgrounds. In this scenario, the likelihood of parallel evolution within-backgrounds should be equal to the likelihood of parallel evolution between-backgrounds. In contrast, if genetic background is key a determinant of evolutionary potential, we would expect adaptation to be highly contingent on genetic background. In this scenario, the likelihood of within-background parallelism should be much greater than the likelihood of between-background parallelism.

Here, we use comparative experimental evolution to test whether genetic differences between strains of bacteria influence their adaptation to a common selection pressure. Specifically, we attempted to test whether parallel evolution was more likely between two populations of the same strain, than between two populations from different strains. To achieve this, we extend existing approaches by selecting for resistance to the antibiotic rifampicin in eight different strains of bacteria from seven different species within the genus *Pseudomonas*. These strains of *Pseudomonas* show substantial genetic divergence, ensuring that our selection experiment was initiated with meaningfully different genetic backgrounds. For example, the genomes of the *Pseudomonads* in our study system vary in size between 4.6 and 7.1 Mb and contain between approximately 4,100 genes and 6,100 genes (Özen and Ussery 2012). Crucially, in spite of this divergence, all of these strains can be cultured in the lab under the same set of conditions, thereby removing any potentially confounding effect of other environmental differences on adaptation.

We challenged bacteria with evolving resistance to the antibiotic rifampicin, which binds to a highly conserved pocket on the β-subunit of RNA polymerase (*rpoB*). Resistance to high doses of rifampicin typically evolves by point mutations in *rpoB*, which result in alterations to the structure of the rifampicin binding pocket. Previous work in *Escherichia coli* (e.g., Garibyan et al. 2003), *Mycobacterium tuberculosis* (e.g., Sandgren et al. 2009), *Pseudomonas aeruginosa* (e.g., Jatsenko et al. 2010), and *Staphylococcus aureus* (e.g., Wichelhaus et al. 2002) has shown that at least 75 different mutations in *rpoB* can drive the evolution of rifampicin resistance. Although some mutations have been reported for a range of species, there are also many strain-specific mutations. Even when the same mutation is reported in different strains, its frequency can be markedly variable. Parallel evolution is therefore possible but not universal, and consequently the evolution of rifampicin resistance is a convenient model for testing hypothesis about parallel molecular evolution. Previous work has also shown that different resistance mutations that are found within the same strain show substantial variation in rifampicin tolerance (MacLean and Buckling 2009), competitive ability (Gagneux et al. 2006; MacLean and Buckling 2009), and gene expression (Carata et al. 2009). This implies that each rifampicin resistance mutation represents a substantially different adaptation to rifampicin.

Because rigorous tests of the prevalence of parallel evolution require large sample sizes, we used an experimental design that involved selecting a large number of populations with the antibiotic rifampicin. After selection for rifampicin resistance, we isolated 75 independently evolved rifampicin-resistant clones from each strain. We then quantified the phenotypic outcome of selection, by measuring the growth rate of clones in the presence of rifampicin, and the genetic outcome of selection, by sequencing *rpoB* in our evolved clones.

Results

Molecular Basis of Adaptation

To determine the frequency of parallel evolution between different genetic backgrounds, we selected 75 rifampicin-resistant mutants in each of eight strains of *Pseudomonas* (fig. 1). Then, to determine the genetic basis of resistance evolution, we sequenced regions of *rpoB* that have been previously implicated in rifampicin resistance. Using this candidate gene approach, we discovered a single mutation per clone. Although we cannot exclude the possibility that our rifampicin-resistant clones carried second site mutations in other parts of their genome, several lines of evidence suggest that the importance of these mutations was minimal. First, the duration of our selection experiment was short, and the resistant clones that we isolated were separated from their common sensitive ancestors by a maximum of $\approx$30 generations. Given that the genomic mutation rate of bacteria is on the order of 0.003 mutations/genome/generation (Drake 1991), there was little opportunity for the clones that we isolated to acquire greater than one mutation. If second site mutations are important for resistance evolution, then different clones with the same *rpoB* mutation should vary phenotypically. However, we found no significant variation in growth rate among clones with the same resistance mutation (fully nested ANOVA (clone nested within mutation nested within strain): $F_{461,555} = 0.799$, $P = 0.994$).

In total, we found 47 different mutations (fig. 2). All but two of these were single nucleotide polymorphisms resulting in an amino acid substitution. The other two mutations were

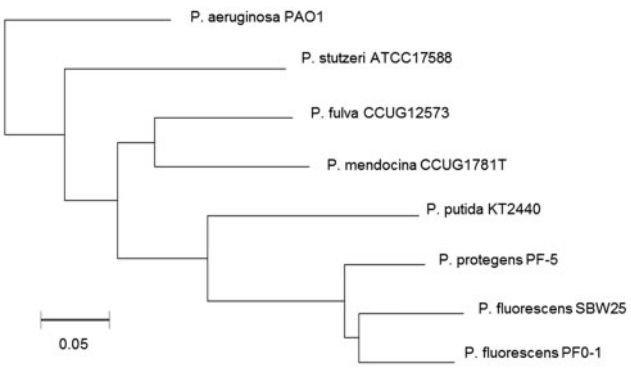

FIG. 1. Phylogeny of the *Pseudomonas* strains used in this study, as determined by the sequence of 55 highly conserved housekeeping genes. Branch lengths are proportional to genetic distance and all nodes on this phylogeny are strongly supported with at least 99% confidence, as determined by an approximate log-likelihood ratio test (Furio and Maclean, in preparation).

Codon affected	Amino acid change	Nucleotide change	P. fluorescens SBW25	P. fluorescens Pf0-1	P. protegens Pf-5	P. putida KT2440	P. fulva CCUG 12573	P. mendocina CCUG 1781 T	P. stutzeri ATCC 17588	P. aeruginosa PAO1
152	Q → L	A455T						0.160		0.147
152	Q → R	A455G						0.027		0.027
514	S → R	C1542G			0.013		0.013			
514	S → R	C1542A			0.013		0.013			
516	L → P	T1547C				0.013	0.013	0.013	0.013	
517	S → P	T1549C		0.040	0.013	0.013	0.013		0.013	
517	S → F	C1550T	0.013		0.147	0.013	0.040	0.027		0.013
518	Q → E	C1552G						0.013		
518	Q → K	C1552A							0.013	0.013
518	Q → P	A1553C						0.013	0.040	
518	Q → R	A1553G	0.027	0.013	0.040	0.067	0.040	0.027	0.160	0.053
518	Q → L	A1553T			0.040	0.040	0.133	0.027	0.053	
521	del 521	del 1561-1563						0.013		
521	D → N	G1561A	0.027			0.013	0.040	0.027	0.013	0.027
521	D → Y	G1561T	0.013		0.013		0.013			
521	D → A	A1562C	0.040							
521	D → G	A1562G	0.613	0.627	0.467	0.800	0.227	0.280	0.213	0.533
521	D → V	A1562T					0.027		0.027	
521	D → E	C1563G		0.013						
523	Insert N	1566 AAC 1567						0.013		
523	N → D	A1567G			0.013					
527	S → L	C1580T					0.053	0.040	0.027	
531	H → Y	C1591T	0.013	0.027	0.040	0.013	0.027	0.013	0.053	0.053
531	H → N	C1591A			0.013				0.027	
531	H → D	C1591G							0.013	
531	H → R	A1592G	0.067	0.093	0.053	0.013	0.080	0.040	0.133	0.027
531	H → L	A1592T	0.013				0.013	0.093	0.067	0.013
531	H → Q	C1593G	0.013							
534	R → H	G1601A					0.013	0.027	0.013	
534	R → P	G1601C							0.013	
536	S → F	C1607T	0.027		0.067		0.053	0.027	0.027	0.027
536	S → Y	C1607A	0.013			0.013			0.013	
538	L → V	C1612G		0.013						
538	L → P	T1613C		0.027	0.013		0.013	0.013	0.027	
538	L → H	T1613A			0.013				0.013	
539	G → D	G1616A					0.027			
563	C → Y	G1691A					0.013			
563	C → W	C1692G					0.027			
569	P → R	C1706G						0.013		
569	P → L	C1706T	0.027	0.013			0.027	0.053	0.013	0.040
575	G → C	G1723T						0.013		
575	G → D	G1724A						0.013		
577	I → L	A1729C	0.027	0.013						
577	I → F	A1729T						0.013	0.013	
577	I → S	T1730G	0.027				0.013			
577	I → T	T1730C	0.013							
579	S → F	C1736T	0.027	0.120	0.040		0.067			0.027

Fig. 2. Frequency of rifampicin resistance mutations across eight strains of *Pseudomonas*. Base pair changes are numbered according to the sequence of *rpoB* in *P. aeruginosa* PAO1, and rows that are grouped together represent mutations in the same codon. Colors on the heat map indicate the frequency of each mutation, with red the most frequent. Blacked-out squares indicate an amino acid substitution would not be possible in a particular strain as a result of single nucleotide substitution.

a deletion of a single amino acid and the insertion of a single amino acid, both occurring in *P. mendocina*. Due to the highly conserved nature of *rpoB*, all 47 of these mutations were theoretically possible for six of our strains, based on the ancestral nucleotide sequence. However, for *P. mendocina* and *P. stutzeri*, the mutation at nucleotide 1,736 would result in a different amino acid substitution than in the other strains, due to a preexisting mutation at another base within the same codon. The distribution of mutations across strains was notably uneven. Most mutations were found in only one ($n = 17$) or two strains ($n = 12$), whereas only four mutations were common to all strains (fig. 2). Mutations found in only a small number of strains tended to be rarer than mutations common across strains and were usually found in only a single clone (mean 1.6 clones per mutation found in only one strain). In contrast, mutations that were found across all strains tended to occur at a high frequency, with 47% of clones possessing the most common mutation, A1562G (fig. 2).

To determine the likelihood of parallel molecular evolution within strain, we calculated the probability that two rifampicin-resistant clones evolved from the same strain share the same *rpoB* mutation. We found that the probability of within-strain parallelism varied widely, with values ranging between 0.096 for *P. fulva* to 0.647 for *P. putida*. However, the phylogenetically corrected mean probability of parallel molecular evolution between populations of the same strain was 0.225, implying that within-strain parallelism is relatively rare. To calculate the probability of between-strain parallel evolution, we first calculated Gst', which measures the proportion of variance in *rpoB* mutations explained by differences between strain. We found that Gst' is equal to 0.147, which implies that differences between strain only account for 14.7% of the variance in *rpoB* mutation frequency. Therefore, the mean similarity of mutation spectra across strains is $1-0.147 = 0.853$, which implies that between-strain parallel evolution is almost as likely as within-strain parallel evolution. Therefore the probability of parallel evolution between populations of different strain can be estimated as the probability of within-strain parallel evolution (0.225) multiplied by the similarity of mutation spectra (0.853), which is equal to 0.19. In other words, we find that genetic divergence between strains has only a minor effect on the probability of parallel evolution in this system.

It is notable however that there are differences between strains with regards to their mutation spectrums. Clustering the mutation spectrums using squared Euclidian distances reveals that there are essentially two spectrums of resistance mutations across our eight strains (fig. 3*A*). The mutation spectra of *P. putida*, *P. protegens*, *P. fluorescens* spp., and *P. aeruginosa* are characterized by a relatively low diversity of adaptive mutations and the dominance of a single hotspot mutation (A1562G), whereas *P. fulva*, *P. stutzeri*, and *P. mendocina* are characterized by a greater diversity of mutations and a reduced prevalence of the A1562G mutation. It is notable that this clustering of mutation spectra into two discreet groups is not a reflection of phylogeny. Instead, the three strains with reduced frequencies of the hotspot

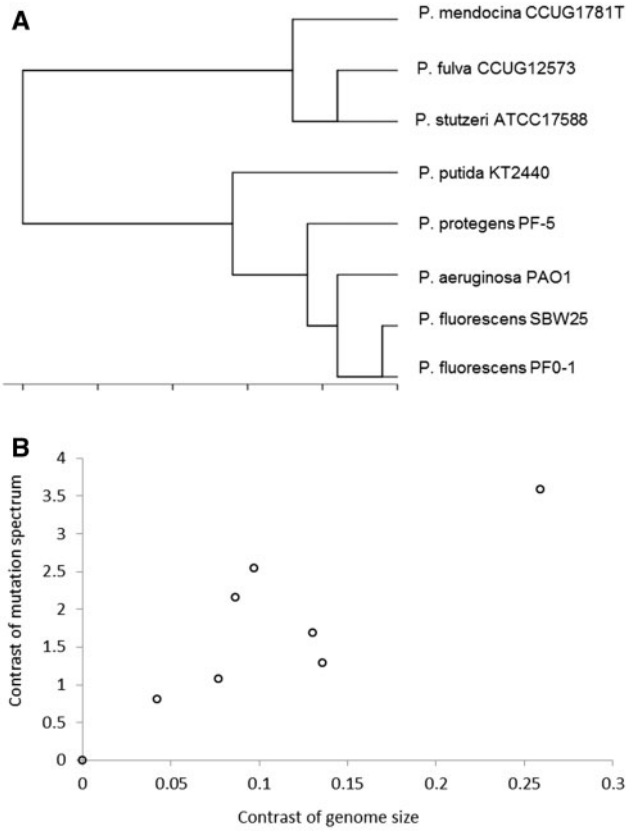

FIG. 3. (*A*) Rifampicin-resistant mutation spectrums cluster into two distinct groups. Dendrogram clustered according to squared Euclidian distances of mutation spectra. (*B*) Genome size significantly and phylogenetically independently correlates with mutation spectrum across *Pseudomonas*. Each point represents a separate PIC, and therefore signifies the absolute amount of change in a character between neighboring strains or nodes.

mutation also have notably smaller genomes than the other strains. Indeed, there is a significant and phylogenetically independent association between genome size and mutation spectrum (fig. 3*B*; correlation between phylogenetic independent contrasts (PICs): $r = 0.855$, $P < 0.01$, $df = 6$). Specifically, the genome sizes of *P. fulva* (4.8 Mb), *P. stutzeri* (4.6 Mb), and *P. mendocina* (5.3 Mb) are noticeably smaller than the genome sizes of our other isolates (>6.2 Mb in all cases).

Parallel Phenotypic Evolution

Having established a clear pattern of parallel evolution at the molecular level, we wished to see if this was reflected at the phenotypic level. To measure the phenotypic outcome of evolution, we assayed the growth rate of each evolved clone in culture medium supplemented with rifampicin at the relevant selective dose. To directly compare the growth rates of evolved clones from different strains, we standardize the growth rate for each clone relative to the growth rate of its rifampicin-sensitive ancestor in the absence of rifampicin. Therefore a growth rate equal to one implies that an evolved clone grows as quickly in the

presence of rifampicin as its ancestor grows in the absence of rifampicin.

We found that the growth rate of evolved clones varied significantly between strains (fig. 4A, one-way ANOVA: $F_{7,590} = 38.44$, $P < 0.001$), and differences between strains explained 31.3% of the variance in growth rate. This result implies that genetic background plays a bigger role in phenotypic evolution than in genetic evolution in this system. There are two potential causes of this variation. First, if resistance mutations have the same effect on growth rate irrespective of the genetic background in which they occur, then strains that have similar mutation spectra should show similar mean levels of adaptation. However, we did not find a significant correlation between mutation spectrum and mean growth rate across strains (fig. 4B; correlation of PICs for mutation spectrum and growth rate: $r = 0.084$, $P = 0.843$, $df = 6$). Alternatively, it is possible for resistance phenotypes to diverge because the same mutations have different effects on growth rate in different strains. According to this explanation, growth rate is determined by epistatic interactions between *rpoB* mutations and the rest of the genome. To directly test for epistasis, we compared the growth rate of the hotspot mutation A1562G across all strains. This mutation is the only mutation that was isolated at least twice in each strain, and 47% of the evolved clones have this mutation. The A1562G mutation had highly epistatic effects (fig. 4C; one-way ANOVA: $F_{7,274} = 32.56$, $P < 0.001$), with growth rates ranging from 0.97 in *P. aeruginosa* to 0.50 in *P. mendocina*. Although this analysis cannot tell us whether other mutations are epistatic, given that this mutation represents 47% of all isolates, we can conclude that epistasis has a major effect on adaptation in our system.

Discussion

We investigated how genetic background influences the evolution of rifampicin resistance in eight different strains of *Pseudomonas* in a common selective environment. We found that parallel evolution was more common within-strain than between-strain, implying that genetic background has a detectable impact on adaptation in this system. The magnitude of this effect varied, however, as genetic background accounted for approximately 14% of the variation in the genetic basis of resistance and approximately 30% of the variation in the phenotype of resistant mutants.

Overall, parallel molecular evolution between populations of the same strain was not very common, largely due to the high diversity of mutations that can result in rifampicin resistance. Fundamentally, the reason parallel evolution was detected is because mutations leading to high level rifampicin resistance are only possible in a single domain of *rpoB* (Campbell et al. 2001). However, despite this highly conserved and relatively narrow target of selection, the molecular basis of adaptation was highly diverse. We identified 47 different mutations leading to rifampicin resistance, a level of diversity common to other studies of rifampicin resistance (e.g., Garibyan et al. 2003; Jatsenko et al. 2010). Nevertheless, although 46 of these mutations were possible in all strains, only four mutations were actually isolated from all strains,

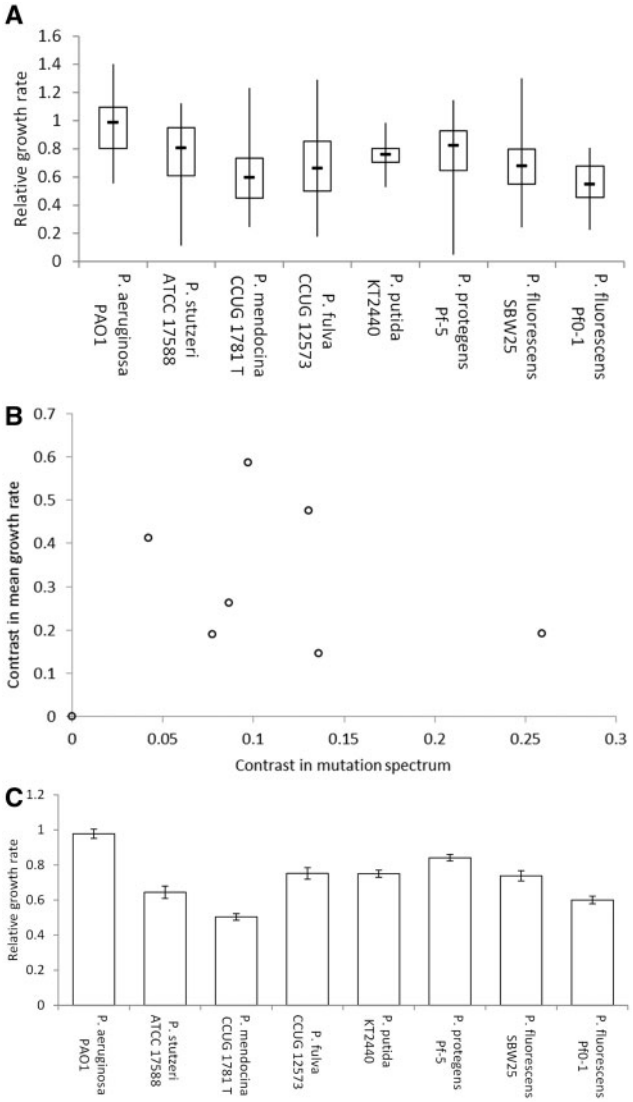

FIG. 4. (*A*) Boxplot of fitness effects of rifampicin resistance mutations across eight strains of *Pseudomonas*, based on relative growth rate in the presence of rifampicin. Each set of mutants is standardized to the growth rate of the relevant ancestral strain growing in the absence of rifampicin. (*B*) No significant correlation between mutation spectrum and mean growth rate in the presence of rifampicin. Each point represents a separate PIC, and therefore signifies the absolute amount of change in a character between neighboring strains or nodes. (*C*) Growth rate of the A1562G mutation growing in the presence of rifampicin. Bars show mean growth rate (± standard error) for all isolates carrying the A1562G mutation in each strain. Values are again standardized to the growth rate of the relevant ancestral strain growing in the absence of rifampicin.

and most mutations were only isolated once in each strain. The most plausible explanation for this pattern is simple chance; given the large number of possible mutations and only 75 mutants per strain, the probability of any individual mutation arising is low. The exceptions to this are mutations at nucleotide 455. These mutations were relatively common in *P. mendocina* and *P. aeruginosa* (mean frequency = 15.3%), but despite being possible, they were never isolated in any of the other strains. It is therefore conceivable that these mutations only confer resistance on these two genetic

backgrounds. This has been previously reported for clinically isolated rifampicin-resistant *M. tuberculosis*, where certain mutations only confer resistance on specific genetic backgrounds (Zaczek et al. 2009). However, it has not been previously identified for this particular locus within *rpoB*, so we cannot be certain that it is definitely occurring here.

We identified two alternative mutation spectra, which differ primarily in terms of the frequency of a hotspot mutation, A1562G. Specifically, in the strains with larger genomes the frequency of A1562G is high, whereas in strains with small genomes the frequency of A1562G is low. It has been previously demonstrated in *E. coli* that knocking-out genes involved in DNA mismatch-repair can result in a wide variety of alterations to mutation spectrum (Garibyan et al. 2003). Crucially, the frequency of this particular hotspot mutation has been shown to increase approximately 3-fold when *mutS* is knocked-out (Garibyan et al. 2003). However, *mutS* is found in all strains of *Pseudomonas* used in this study, and none of these strains have a high mutation rate as would be expected from a *mutS* knockout (Vogwill T, unpublished data). Of course gene presence might not be the major driver of mutation spectrum, and instead relative gene expression might be crucial. For example, it has been shown that alterations to temperature can also have pronounced effects on mutation spectrum (Jatsenko et al. 2010). This is again a particular relevant example, as altering the temperature altered the frequency of the A1562G mutation in *P. aeruginosa*. As environmental conditions were controlled in our study, this is arguing that the strains with smaller genomes show differential regulation of DNA-repair genes in response to the same environmental stimuli.

In line with previous work, we found evidence that the fitness effects of rifampicin resistance mutations are epistatic (MacLean et al. 2010; Hall and MacLean 2011; Angst and Hall 2013; Borrell et al. 2013). Although we only analyzed the fitness effects of the most common mutation, this mutation played a key role in phenotypic evolution because almost 50% of isolates carried this mutation. Hypothetically, the cause of this epistasis could be that in different strains the same mutation has differential effects of the binding of rifampicin to *rpoB*. If so, this epistasis is therefore reflecting the same mutation conferring higher levels of resistance in some strains than others. However, given that the rifampicin binding pocket is highly conserved across all strains (Campbell et al. 2001), it seems highly unlikely that this is the case. Moreover, these mutations have epistatic effects on fitness in the absence of rifampicin (Vogwill T, unpublished data). It is therefore more likely that the same mutation has different fitness costs in different strains. As the cost of rifampicin resistance is known to depend on demand for RNA-polymerase (Hall et al. 2011), a simple explanation is therefore that demand RNA-polymerase varies between strains. However, *rpoB* mutations are also known to be highly pleiotropic (Xu et al. 2002; Gao et al. 2013; Hall 2013; Rodriguez-Verdugo et al. 2013). Under laboratory conditions, *rpoB* mutations can be selected by a variety of selection pressures, such as temperature (Rodriguez-Verdugo et al. 2013) and aging colonies (Wrande et al. 2008). Although *rpoB* itself is highly conserved, these other cellular systems are far less so. It is therefore possible that the epistasis we detected is a result of complex interplay between *rpoB* mutations and the expression of genes involved in multiple other cellular processes.

One limitation of this study is that the duration of the selection experiment was short by microbial standards (<30 generations). On a purely practical level, the short duration of the experiment was necessitated by the number of populations used ($\approx 3{,}800$), but it is important to speculate about the extent to which time influenced the outcome of our study. It could be argued that the influence of genetic background would have been further erased by continued selection for adaptation. For example, Travisano et al. (1995) found that the influence of evolutionary history on fitness declined during adaptation. However, it could be argued that genetic background should become more important for longer adaptive walks, particular with regards to molecular evolution (Harms and Thornton 2013). In the case of rifampicin, studies in *M. tuberculosis* (Casali et al. 2012, Comas et al. 2012; de Vos et al. 2013) and *Salmonella typhimurium* (Brandis et al. 2012; Brandis and Hughes 2013) have shown that compensatory adaptation often occurs by mutations in *rpoC*, which encodes another subunit of RNA-polymerase. Intriguingly, approximately 25% of sites in *rpoC* which have been implicated in compensation for the cost of rifampicin resistance are not conserved between these two species (MacLean RC, unpublished data). This suggests that genetic background is likely to play a more important role in molecular evolution over longer time scales (see also Pelz et al. 2005; McCracken et al. 2009; Powles and Yu 2010).

Due to the intensity of selection in our experiment, the only possible mechanism of resistance was mutations in a highly conserved domain of RNA polymerase (Campbell et al. 2001). Speculatively, the impact of genetic background on adaptation would have been much stronger if we had used a weaker selective pressure, as this would have provided more potential routes for adaptation. For example, recent work has shown that mutations in many genes with diverse functional roles can result in small increases in resistance to antibiotics (Fajardo et al. 2008; Girgis et al. 2009). Hypothetically, selection for resistance to an inhibitory but sublethal dose could have resulted in a much broader diversity of outcomes. This would be due to different strains would have access to different routes to adaptation because of their differing genomic contents. If this argument is correct, then our estimate of the importance of genetic background should be seen as a conservative estimate of the contribution it can make to divergent responses to selection.

One key feature of our model system is that there are many different substitutions that can drive the evolution of resistance, making it possible for the probability of parallel evolution to vary over a wide range. Alternatively, if only a small number of substitutions could confer rifampicin resistance, very high levels of parallel molecular evolution would have been an almost guaranteed outcome of our experiment. A similar argument can be made with regards to the effect of horizontal gene transfer on parallel evolution. Resistance to many antibiotics evolves by the acquisition

Table 1. Oligonucleotides Used for PCR and Sequencing.

	Rifampicin-Resistance Determining Region	Other Region (First 500 bp of rpoB)
Pseudomonas Aeruginosa	5′-GTTCTTCAGCGCCGAGCG-3′ 5′-GCGATGACGTGGTCGGC-3′	5′–ATGGCTTACTCATACACTGAG-3′ 5′-CTCGATGCGCACGACCTG-3′
P. protegens, P. stutzeri, P. fluorescens Pf0-1, P. fulva, P. mendocina	5′-CTGTTCAACAACCTGTTCTTC-3′ 5′-CGCGCCTTGACGGTGAATTCG-3′	5′-AAAAAACGTATCCGCAAGGA-3′ 5′-TCGATGTCGTTGGTGTACA-3′
P. putida	5′-CGTATCGGTCGTACCGAGATC-3′ 5′-CGCGCCTTGACGGTGAATTCG-3′	5′-AAAAAACGTATCCGCAAGGA-3′ 5′-TCGATGTCGTTGGTGTACA-3′
P. fluorescens SBW25	5′-CTGTTCAACAACCTGTTCTTC-3′ 5′-GGCGCCTTGACGGTGAACTCG-3′	5′-AAAAAACGTATCCGCAAGGA-3′ 5′-CGGAGATGAACGGACCGCAG-3′

of mobile genetic elements which carry resistance genes, such as plasmids and lysogenic bacteriophages (reviewed in Andersson and Hughes 2010). Because horizontal gene transfer provides a vehicle for the movement of adaptive alleles between lineages, it should increase the probability of parallel molecular evolution. However, it is unclear if this will result in increased levels of phenotypic parallelism, because in keeping with the results present here, the fitness effect of acquiring a plasmid is host-strain dependent (e.g., Mongkolrattanothai et al. 2009; Humphrey et al. 2012; Sandegren et al. 2012).

Many of the most striking examples of parallel molecular evolution across genetic backgrounds come from the evolution of resistance to toxins, such as insecticides (ffrench-Constant et al. 2004), herbicides (Bernasconi et al. 1995), and antibiotics (e.g., Jatsenko et al. 2010). Our study explores the underlying mechanistic causes of adaption to rifampicin in multiple strains of *Pseudomonas*, and it suggests that parallel phenotypic evolution is likely to be less common than parallel molecular evolution. More generally, this study may stimulate future work aimed at integrating comparative approaches into experimental evolution.

Materials and Methods

Study System

The strains used, and their phylogenetic relatedness, is shown in figure 1. Before the selection experiment all strains were stored at −80°C in 20% glycerol.

Selection Experiment

For each strain, 480 replicate populations were inoculated with approximately 50 cells of a rifampicin sensitive ancestral strain and grown for 48 h in 200 µl of King's B (KB) media at 30°C with constant shaking at 200 rpm. Approximately 10^7 cells from each population were then plated onto KB-agar containing rifampicin at the appropriate MIC (minimum inhibitory concentration) for each strain (60 µg/ml for *P. aeruginosa*, 30 µg/ml for all other strains). MICs had been previously determined for each strain using standard methods as set by the European Committee for Antimicrobial susceptibility testing. For each strain, a single colony was isolated from 75 random populations that successfully evolved rifampicin resistance. If less than 75 of the initial 480 populations produced at least one colony, additional selection experiments were performed until 75 independent mutants were achieved for each strain. Specifically, this involved inoculating a further 480 populations for *P. fluorescens Pf0-1* and a further 768 populations for *P. fluorescens SBW25*. Clones were stored at −80°C in 20% glycerol.

Sequencing

The rifampicin resistance-determining region (RRDR) of *rpoB* was sequenced for each strain using the primers listed in table 1. Purification of polymerase chain reaction (PCR) products was performed using ExoStar enzymes, and sequencing was performed using BigDye terminator kit. Both were used according to the manufacturer's instructions. For populations lacking mutations in this region, we sequenced a second region of *rpoB* where rifampicin-resistant mutations are known to occur, specifically the first 1,000 nt of *rpoB*.

Growth Rate Assay

To assay the extent of adaptation to selective conditions, we measure the growth rate in the presence of rifampicin for each evolved clone. Overnight cultures of resistant clones were grown in KB supplemented with rifampicin (60 µg/ml for *P. aeruginosa*, 30 µg/ml for all other strains) were diluted 2,000-fold into fresh rifampicin-supplemented medium and incubated at 30°C with constant shaking at 200 rpm. Viable counts were taken every 2.5 h for 7.5 h. Viable cell counts were performed using BacTiter-Glo kits (Promega, UK), which quantifies the number of cells in a sample by measuring adenosine triphosphate (ATP) content. Compared with the more common optical density measures of cell density, these kits have a larger dynamic range and can be used for strains where optical density does not provide an accurate measure of cell density. For example, under our laboratory conditions some strains only achieve an optical density toward the very end of their growth curves, and therefore exponential growth rates could not be estimated from optical density data. This is despite reaching a higher cell density than other strain which do have an optical density. This is likely due to optical density confounding cell density with cell size, fluorescent protein content, biofilming, etc.

Growth rates were calculated as the Malthusian parameter, and standardized to the growth rate for each wild-type strains in the absence of rifampicin. The Malthusian parameter is the natural log of ratio of the final density to the initial density. As this is standardized to the ancestral value, this provides the relative growth rate of each mutant, which is a proxy for fitness. Each of the 600 rifampicin-resistant clones

were measured twice and we measured at least six replicate cultures of the rifampicin sensitive strain from each strain. For *P. aeruginosa*, this assay was repeated using both 30 and 60 μg/ml. There was a strong correlation for fitness measured in the two assays (df = 69, $r = 0.541$, $P < 0.001$) and no difference in mean fitness (paired *t*-test: $t = 0.245$, df = 70, $P = 0.807$), demonstrating that the dosage of rifampicin did not strongly affect the result of this assay.

Phylogenetic Reconstruction

To obtain an accurate phylogeny of the genus *Pseudomonas*, we retrieved the nucleotide sequences of 55 housekeeping genes (supplementary table S1, Supplementary Material online) for 28 *Pseudomonas* genomes (supplementary fig. S1, Supplementary Material online) plus *Azotobacter vinelandii* from the Integrated Microbial Genomes database (Markowitz et al. 2012) using reciprocal BLAST to make sure of their orthology. Multiple alignments for all genes were obtained using MUSCLE (Edgar 2004) and afterwards trimmed with GBlocks (Castresana 2000) and concatenated with SequenceMatrix (Vaidya et al. 2011). The resulting supermatrix was then used to infer the nucleotide substitution model using JModelTest 2 (Darriba et al. 2012). A maximum-likelihood (ML) phylogenetic tree was then inferred assuming a GTR+G+I model with PHYML (Guindon et al. 2010). Additional trees inferred using Bayesian inference and Neighbor Joining as well as a ML tree obtained with a supermatrix of the protein sequences yielded identical topologies (data not shown). The relevant phylogeny for the eight strains in this study system was then manually extracted. Branch lengths are proportional to genetic distance and all nodes on this phylogeny are strongly supported with at least 99% confidence, as determined by an approximate log-likelihood ratio test (Furio V, Maclean RC, unpublished manuscript).

Analysis of Molecular Data

The probability of within-strain parallel molecular evolution was calculated using Simpson's index of concentration (Simpson 1949), which measure the probability that two randomly chosen clones from the same strain possess the same mutation:

$$\lambda = \sum_{i=1}^{n} P_i^2$$

where P_i is the proportion of *i*-th mutation. This was calculated for each of our eight strains, from which the phylogenetically corrected mean was then calculated using phylogenetically independent contrasts (Felsenstein 1985; Garland et al. 1999).

We estimate the level of between strain parallel evolution by first calculating Nei's Gst (Nei 1973) and then standardizing it to Hedrick's Gst' (Hedrick and Goodnight 2005). Gst is a measure of the genetic dissimilarity across multiple samples, calculated by estimating the proportion of genetic diversity due to differences in allele frequencies in different subpopulations. Here, we use it in an analogous manner to measure the proportion of genetic diversity due to different mutation frequencies in different strains. Specifically, as we are interested in the similarity of different mutation spectra, here we use 1−Gst'. A value of 0 would therefore indicate no parallel evolution, whereas a value of 1 would indicate between-strain parallel evolution is as likely as within-strain parallel evolution.

Gst is calculated using the formula (Nei 1973):

$$\mathrm{Gst} = \frac{\mathrm{Ht} - \mathrm{Hs}}{\mathrm{Ht}}$$

where Hs is mean within-strain diversity, which is equal to 1−λ, and Ht is total diversity, which is calculated as:

$$\mathrm{Ht} = \sum_{i=1}^{n} \overline{P}_i^2$$

Note that here we again use PICs to calculate the phylogenetically corrected mean proportion for each mutation. This Gst was then standardized to Gst' (Hedrick and Goodnight 2005), using the formula:

$$\mathrm{Gst}' = \frac{\mathrm{Gst}(K - 1 + \mathrm{Hs})}{(K-1)(1-\mathrm{Hs})}$$

where *K* is the number of samples.

Associations between mutation spectrum, genome size, and fitness, were performed using PICs using the method of Felsenstein (Felsenstein 1985). PICs for genome size and fitness were calculated as standard. Contrasts for mutation spectrum between tips/nodes were calculated as squared differences (as opposed to subtraction), with the rest of the PIC procedure performed using standard methods.

The dendrogram of mutation spectrums was created using the hierarchical clustering function in SPSS, with clusters created based on the squared Euclidian distance between groups.

Analyzing Phenotypic Diversity

Growth rate data was analyzed using a one-way ANOVA. To control for phylogeny, the phylogenetically corrected mean was used when calculating the amongst-group variance.

Supplementary Material

Supplementary material is available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

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. V.F. was supported by an MEC Postdoctoral Fellowship from the Spanish Government (EX-2010-0958). T.V. and M.K. carried out the experimental work and analyzed experimental data with R.C.M.; V.F. constructed the

phylogeny; V.F. and T.V. carried out comparative analyses; T.V. and R.C.M. prepared the manuscript and all authors contributed to designing the study.

References

Andersson DI, Hughes D. 2010. Antibiotic resistance and its cost: is it possible to reverse resistance? *Nat Rev Microbiol.* 8: 260–271.

Angst D, Hall A. 2013. The cost of antibiotic resistance depends on evolutionary history in *Escherichia coli*. *BMC Evol Biol.* 13:163.

Bernasconi P, Woodworth AR, Rosen BA, Subramanian MV, Siehl DL. 1995. A naturally occurring point mutation confers broad range tolerance to herbicides that target acetolactate synthase. *J Biol Chem.* 270:17381–17385.

Bollback JP, Huelsenbeck JP. 2009. Parallel genetic evolution within and between bacteriophage species of varying degrees of divergence. *Genetics* 181:225–234.

Borrell S, Teo Y, Giardina F, Streicher EM, Klopper M, Feldmann J, Müller., Victor TC, Gagneux S. 2013. Epistasis between antibiotic resistance mutations drives the evolution of extensively drug-resistant tuberculosis. *Evol Med Public Health.* 2013:65–74.

Brandis G, Hughes D. 2013. Genetic characterization of compensatory evolution in strains carrying *rpoB* Ser531Leu, the rifampicin resistance mutation most frequently found in clinical isolates. *J Antimicrob Chemother.* 68:2493–2497.

Brandis G, Wrande M, Liljas L, Hughes D. 2012. Fitness-compensatory mutations in rifampicin-resistant RNA polymerase. *Mol Microbiol.* 85:142–151.

Breen MS, Kemena C, Vlasov PK, Notredame C, Kondrashov FA. 2012. Epistasis as the primary factor in molecular evolution. *Nature* 490: 535–538.

Campbell EA, Korzheva N, Mustaev A, Murakami K, Nair S, Goldfarb A, Darst SA. 2001. Structural mechanism for rifampicin inhibition of bacterial RNA polymerase. *Cell* 104:901–912.

Carata E, Peano C, Tredici S, Ferrari F, Tala A, Corti G, Bicciato S, De Bellis G, Alifano P. 2009. Phenotypes and gene expression profiles of *Saccharopolyspora erythraea* rifampicin-resistant (rif) mutants affected in erythromycin production. *Microb Cell Fact.* 8:18.

Casali N, Nikolayevskyy V, Balabanova Y, Ignatyeva O, Kontsevaya I, Harris SR, Bentley SD, Parkhill J, Nejentsev S, Hoffner SE, et al. 2012. Microevolution of extensively drug-resistant tuberculosis in Russia. *Genome Res.* 22:735–745.

Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol.* 17: 540–552.

Comas I, Borrell S, Roetzer A, Rose G, Malla B, Kato-Maeda M, Galagan J, Niemann S, Gagneux S. 2012. Whole-genome sequencing of rifampicin-resistant *Mycobacterium tuberculosis* strains identifies compensatory mutations in RNA polymerase genes. *Nat Genet.* 44: 106–110.

Costanzo M, Baryshnikova A, Bellay J, Kim Y, Spear ED, Sevier CS, Ding H, Koh JLY, Toufighi K, Mostafavi S, et al. 2010. The genetic landscape of a cell. *Science* 327:425–431.

Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nat Methods.* 9: 772–772.

de Vos M, Müller B, Borrell S, Black PA, van Helden PD, Warren RM, Gagneux S, Victor TC. 2013. Putative compensatory mutations in the *rpoC* gene of rifampin-resistant *Mycobacterium tuberculosis* are associated with ongoing transmission. *Antimicrob Agents Chemother.* 57:827–832.

Dettman JR, Rodrigue N, Aaron SD, Kassen R. 2013. Evolutionary genomics of epidemic and nonepidemic strains of *Pseudomonas aeruginosa*. *Proc Natl Acad Sci U S A.* 110:21065–21070.

Drake JW. 1991. A constant rate of spontaneous mutation in DNA-based microbes. *Proc Natl Acad Sci U S A.* 88:7160–7164.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32:1792–1797.

Fajardo A, Martínez-Martín N, Mercadillo M, Galán JC, Ghysels B, Matthijs S, Cornelis P, Wiehlmann L, Tümmler B, Baquero F, et al. 2008. The neglected intrinsic resistome of bacterial pathogens. *PLoS One* 3:e1619.

Felsenstein J. 1985. Phylogenies and the comparative method. *Am Nat.* 125:1–15.

ffrench-Constant RH, Daborn PJ, Goff GL. 2004. The genetics and genomics of insecticide resistance. *Trends Genet.* 20:163–170.

Gagneux S, Long CD, Small PM, Van T, Schoolnik GK, Bohannan BJM. 2006. The competitive cost of antibiotic resistance in *Mycobacterium tuberculosis*. *Science* 312:1944–1946.

Gao W, Cameron DR, Davies JK, Kostoulias X, Stepnell J, Tuck KL, Yeaman MR, Peleg AY, Stinear TP, Howden BP. 2013. The *RpoB* H481Y rifampicin resistance mutation and an active stringent response reduce virulence and increase resistance to innate immune responses in *Staphylococcus aureus*. *J Infect Dis.* 207: 929–939.

Garibyan L, Huang T, Kim M, Wolff E, Nguyen A, Nguyen T, Diep A, Hu K, Iverson A, Yang H, et al. 2003. Use of the *rpoB* gene to determine the specificity of base substitution mutations on the *Escherichia coli* chromosome. *DNA Repair.* 2:593–608.

Garland T, Midford PE, Ives AR. 1999. An introduction to phylogenetically based statistical methods, with a new method for confidence intervals on ancestral values. *Am Zool.* 39:374–388.

Girgis HS, Hottes AK, Tavazoie S. 2009. Genetic architecture of intrinsic antibiotic susceptibility. *PLoS One* 4:e5629.

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol.* 59:307–321.

Hall AR. 2013. Genotype-by-environment interactions due to antibiotic resistance and adaptation in *Escherichia coli*. *J Evol Biol.* 26: 1655–1664.

Hall AR, Iles JC, MacLean RC. 2011. The fitness cost of rifampicin resistance in *Pseudomonas aeruginosa* depends on demand for RNA polymerase. *Genetics* 187:817–822.

Hall AR, MacLean RC. 2011. Epistasis buffers the fitness effects of rifampicin- resistance mutations in *Pseudomonas aeruginosa*. *Evolution* 65: 2370–2379.

Harms MJ, Thornton JW. 2013. Evolutionary biochemistry: revealing the historical and physical causes of protein properties. *Nat Rev Genet.* 14:559–571.

Hedrick PW, Goodnight C. 2005. A standardized genetic differentiation measure. *Evolution* 59:1633–1638.

Holder KK, Bull JJ. 2001. Profiles of adaptation in two similar viruses. *Genetics* 159:1393–1404.

Humphrey B, Thomson N, Thomas C, Brooks K, Sanders M, Delsol A, Roe J, Bennett P, Enne V. 2012. Fitness of *Escherichia coli* strains carrying expressed and partially silent IncN and IncP1 plasmids. *BMC Microbiol.* 12:53.

Jatsenko T, Tover A, Tegova R, Kivisaar M. 2010. Molecular characterization of Rifr mutations in *Pseudomonas aeruginosa* and *Pseudomonas putida*. *Mutat Res.* 683:106–114.

Khan AI, Dinh DM, Schneider D, Lenski RE, Cooper TF. 2011. Negative epistasis between beneficial mutations in an evolving bacterial population. *Science* 332:1193–1196.

Lieberman TD, Michel JB, Aingaran M, Potter-Bynoe G, Roux D, Davis MR, Skurnik D, Leiby N, LiPuma JJ, Goldberg JB, et al. 2011. Parallel bacterial evolution within multiple patients identifies candidate pathogenicity genes. *Nat Genet.* 43:1275–1280.

MacLean RC, Buckling A. 2009. The Distribution of fitness effects of beneficial mutations in *Pseudomonas aeruginosa*. *PLoS Genet.* 5: e1000406.

MacLean RC, Perron GG, Gardner A. 2010. Diminishing returns from beneficial mutations and pervasive epistasis shape the fitness landscape for rifampicin resistance in *Pseudomonas aeruginosa*. *Genetics* 186:1345–1354.

Markowitz VM, Chen IMA, Palaniappan K, Chu K, Szeto E, Grechkin Y, Ratner A, Jacob B, Huang J, Williams P, et al. 2012. IMG: the integrated microbial genomes database and comparative analysis system. *Nucleic Acids Res.* 40:D115–D122.

McCracken KG, Barger CP, Bulgarella M, Johnson KP, Sonsthagen SA, Trucco J, Valqui TH, Wilson RE, Winker K, Sorenson MD. 2009. Parallel evolution in the major haemoglobin genes of eight species of Andean waterfowl. *Mol Ecol.* 18:3992–4005.

Mongkolrattanothai K, Pumfrey L, Mankin P, Gray BM. 2009. Acquisition of high-level mupirocin resistance and its fitness cost among methicillin-resistant *Staphylococcus aureus* strains with low-level mupirocin resistance. *J Clin Microbiol.* 47:4158–4160.

Nei M. 1973. Analysis of gene diversity in subdivided populations. *Proc Natl Acad Sci U S A.* 70:3321–3323.

Nguyen AH, Molineux IJ, Springman R, Bull J. 2012. Multiple genetic pathways to similar fitness limits during viral adaptation to a new host. *Evolution* 66:363–374.

Ostrowski EA, Woods RJ, Lenski RE. 2008. The genetic basis of parallel and divergent phenotypic responses in evolving populations of *Escherichia coli*. *Proc R Soc Lond B Biol Sci.* 275:277.

Özen A, Ussery D. 2012. Defining the *Pseudomonas* genus: where do we draw the line with *Azotobacter*? *Microb Ecol.* 63:239–248.

Pelz HJ, Rost S, Hünerberg M, Fregin A, Heiberg AC, Baert K, MacNicoll AD, Prescott CV, Walker AS, Oldenburg J, et al. 2005. The genetic basis of resistance to anticoagulants in rodents. *Genetics* 170: 1839–1847.

Powles SB, Yu Q. 2010. Evolution in action: plants resistant to herbicides. *Ann Rev Plant Biol.* 61:317–347.

Rodriguez-Verdugo A, Gaut B, Tenaillon O. 2013. Evolution of *Escherichia coli* rifampicin resistance in an antibiotic-free environment during thermal stress. *BMC Evol Biol.* 13:50.

Rokas A, Carroll SB. 2008. Frequent and widespread parallel evolution of protein sequences. *Mol Biol Evol.* 25:1943–1953.

Sandegren L, Linkevicius M, Lytsy B, Melhus A, Andersson DI. 2012. Transfer of an *Escherichia coli* ST131 multiresistance cassette has created a Klebsiella pneumoniae-specific plasmid associated with a major nosocomial outbreak. *J Antimicrob Chemother* 67: 74–83.

Sandgren A, Strong M, Muthukrishnan P, Weiner BK, Church GM, Murray MB. 2009. Tuberculosis drug resistance mutation database. *PLoS Med.* 6:e1000002.

Simpson EH. 1949. Measurement of diversity. *Nature* 163:688–688.

Travisano M, Mongold JA, Bennett AF, Lenski RE. 1995. Experimental tests of the roles of adaptation, chance, and history in evolution. *Science* 267:87–90.

Vaidya G, Lohman DJ, Meier R. 2011. SequenceMatrix: concatenation software for the fast assembly of multi-gene datasets with character set and codon information. *Cladistics* 27:171–180.

Wichelhaus TA, Böddinghaus B, Besier S, Schäfer V, Brade V, Ludwig A. 2002. Biological cost of rifampin resistance from the perspective of *Staphylococcus aureus*. *Antimicrob Agents Chemother.* 46: 3381–3385.

Wichman HA, Scott LA, Yarber CD, Bull JJ. 2000. Experimental evolution recapitulates natural evolution. *Philos Trans R Soc Lond B Biol Sci.* 355:1677–1684.

Wong A, Kassen R. 2011. Parallel evolution and local differentiation in quinolone resistance in *Pseudomonas aeruginosa*. *Microbiology* 157: 937–944.

Wong A, Rodrigue N, Kassen R. 2012. Genomics of adaptation during experimental evolution of the opportunistic pathogen *Pseudomonas aeruginosa*. *PLoS Genet.* 8:e1002928.

Wood T, Burke J, Rieseberg L. 2005. Parallel genotypic adaptation: when evolution repeats itself. In: Mauricio R, editor. genetics of adaptation. Netherlands: Springer. p. 157–170.

Woods R, Schneider D, Winkworth CL, Riley MA, Lenski RE. 2006. Tests of parallel molecular evolution in a long-term experiment with *Escherichia coli*. *Proc Natl Acad Sci U S A.* 103:9107.

Wrande M, Roth JR, Hughes D. 2008. Accumulation of mutants in "aging" bacterial colonies is due to growth under selection, not stress-induced mutagenesis. *Proc Natl Acad Sci U S A.* 105: 11863–11868.

Xu J, Tozawa Y, Lai C, Hayashi H, Ochi K. 2002. A rifampicin resistance mutation in the *rpoB* gene confers ppGpp-independent antibiotic production in *Streptomyces coelicolor* A3(2). *Mol Genet Genomics.* 268:179–189.

Zaczek A, Brzostek A, Augustynowicz-Kopec E, Zwolska Z, Dziadek J. 2009. Genetic evaluation of relationship between mutations in *rpoB* and resistance of *Mycobacterium tuberculosis* to rifampin. *BMC Microbiol.* 9:1–8.

Zhang J, Kumar S. 1997. Detection of convergent and parallel evolution at the amino acid sequence level. *Mol Biol Evol.* 14:527–536.