Epistatic interactions between ancestral genotype and beneficial mutations shape evolvability in Pseudomonas aeruginosa

DOI:
10.1111/evo.12958

Epistatic interactions between ancestral genotype and beneficial mutations shape evolvability in *Pseudomonas aeruginosa*

Danna R. Gifford[1,2,*] Macarena Toll-Riera[1,3], and R. Craig MacLean[1]

Short title: Constraints on evolvability by magnitude epistasis

[*]**Corresponding author**: danna.gifford@manchester.ac.uk

Telephone: +44 161 2755 376

[1]**Address**

Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, United Kingdom

[2]**Present address**

Faculty of Life Sciences, University of Manchester, Michael Smith Building, Dover St, Manchester M13 9PL United Kingdom

[3]**Present address**

Institute of Evolutionary Biology and Environmental Sciences, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland

Data: Archived in Dryad (http://datadryad.org/) and NCBI Sequence Read Archive (SRA, http://www.ncbi.nlm.nih.gov/sra/) upon manuscript acceptance

Keywords: experimental evolution, evolvability, epistasis, antibiotic resistance

Abstract

The idea that interactions between mutations influence adaptation by driving populations to low and high fitness peaks on adaptive landscapes is deeply ingrained in evolutionary theory. Here we investigate the impact of epistasis on evolvability by challenging populations of two *Pseudomonas aeruginosa* clones bearing different initial mutations (in *rpoB* conferring rifampicin resistance, and the type IV pili gene network) to adaptation to a medium containing L-serine as the sole carbon source. Despite being initially indistinguishable in fitness, populations founded by the two ancestral genotypes reached different fitness following 300 generations of evolution. Genome sequencing revealed that the difference could not be explained by acquiring mutations in different targets of selection: the majority of clones from both ancestors converged on one of two strategies, 1) acquiring mutations in either PA2449 (*gcsR*, an L-serine-metabolism RpoN enhancer binding protein) or 2) protease genes. Additionally, populations from both ancestors converged on loss-of-function mutations in the type IV pili gene network, either due to ancestral or acquired mutations. No compensatory or reversion mutations were observed in RNA polymerase genes, in spite of the large fitness costs typically associated with mutations in *rpoB*. Although current theory points to sign epistasis as the dominant constraint on evolvability, these results suggest that the role of magnitude epistasis in constraining evolvability may be under-appreciated. The contribution of magnitude epistasis is likely to be greatest under the biologically-relevant mutation supply rates that make back-mutations probabilistically unlikely.

Introduction

Epistatic interactions between novel beneficial mutations and the ancestral genotype they arise in are thought to shape the course of adaptive evolution by restricting the availability of evolutionary trajectories (Phillips et al. 2000; de Visser et al. 2011). As a consequence, ancestral genotypes may differ in the rate of fitness increase in a new environment through the acquisition of

beneficial mutations, a trait known as 'evolvability'. Differences in evolvability can arise even if selection acts on the same adaptive targets (Colegrave and Collins 2008). A sizable number of studies have shown that ancestral fitness can constrain evolvability (e.g. Burch and Chao 2000; Silander et al. 2007; MacLean et al. 2010; Chou et al. 2011; Gifford et al. 2011; Khan et al 2011; Rokyta et al. 2011; Salverda et al. 2011; Kryazhimskiy et al. 2014; Wong and Seguin 2015), a phenomenon known as diminishing returns epistasis. Diminishing returns epistasis is the tendency for beneficial mutations to confer a smaller benefit in genetic backgrounds of higher fitness (de Visser et al. 2011). These studies match predictions from theoretical models of adaptive evolution predicting that the average fitness of beneficial mutations should be proportional to initial fitness in populations climbing the same adaptive peak (Martin and Lenormand 2006). It is a special case of a more general form of epistasis known as 'magnitude epistasis', defined as the interaction between two mutations that produces a non-additive effect on fitness, without changing the sign (i.e. beneficial/deleterious) of the effect of either mutation (de Visser et al. 2011). Magnitude epistasis is in turn contrasted with sign epistasis, which occurs when the sign of a mutation is conditional on genetic background. Theoretical models predict that sign epistasis will constrain evolvability by restricting populations to certain adaptive peaks (e.g. Weinreich et al. 2005;Kvitek and Sherlock 2011; Schenk et al. 2013; Szamecz et al. 2014), whereas magnitude epistasis (including diminishing returns) will restrict the height of adaptive peaks.

Although the roles of diminishing returns epistasis and sign epistasis in evolvability are well-supported by theoretical models and empirical tests, the role of magnitude epistasis, in the general sense, is less concrete. Determining the role of magnitude epistasis in evolvability is important because magnitude interactions are common among pairs of mutations (de Visser et al. 2011 and references therein; Chou et al. 2014). A small, but growing, number of studies that suggest that magnitude epistasis may also influence evolvability (Woods et al. 2011; Kryazhimskiy et al. 2014; Szamecz et al. 2014). Evolution experiments using ancestors with similar initial fitness can

demonstrate whether or not magnitude epistasis can constrain evolvability, under conditions that are not solely dictated by diminishing returns, e.g. when ancestral genotypes differ genetically, but not in fitness. This is particularly important in the context of antibiotic resistance, where different resistance mutations may differ in their ability to recover fitness through compensatory mutations, and reversions to sensitivity are rare (Andersson and Hughes 2010; MacLean et al. 2010; Hall et al. 2010; Barrick et al. 2010), which may trap populations on local fitness optima.

Here we show that evolvability can be constrained by magnitude epistasis that does not result from the diminishing returns-type pattern observed through acquiring multiple beneficial mutations. In our study, we used experimental evolution with two *Pseudomonas aeruginosa* strains having similar fitness in a simple growth environment (a defined medium with L-serine as the sole carbon source). The strains differed in two protein coding genes: *rpoB* (SNP versus indel) and *pilQ* (frame-shift versus wild-type). Mutations in *rpoB* that confer rifampicin resistance are known to impose large fitness costs through defects in transcription (Qi et al. 2014). Loss of PilQ function leads to the loss of external type IV pili (Martin et al. 1993), which is associated with the loss of twitching motility (Mattick 2002) and decreased virulence (Persat et al. 2015). Despite remarkably parallel evolution at the genetic pathway-level, the two strains differed significantly in the fitness benefit of acquired beneficial mutations. In addition, the genotype initially capable of expressing external pili tended to acquire loss of function mutations in pili genes. No reversions of the *rpoB* resistance alleles were observed, suggesting that the difference of evolvability was not due to a difference in reversability of SNPs versus small indels. Rather, the difference in evolvability was due to epistasis from one or both of the pre-existing mutations.

Methods

Strains and growth medium

The ancestral strains used in our experiment were isolated from independent populations of a

laboratory strain of *P. aeruginosa* PAO1 that were selected for resistance to a growth-inhibiting dose of rifampicin (64 mg/l) during a one-day fluctuation assay (MacLean and Buckling 2009). Sanger sequencing of *rpoB* initially revealed that each clone carried a single mutation in the rifampicin binding pocket of *rpoB* (alleles *rpoB*S536F and *rpoB*+P518, MacLean and Buckling 2009). Whole genome re-sequencing of the strains by Illumina HiSeq2000 indicated the presence of these mutations, and revealed one additional mutation in a protein coding gene in the same genetic background as the *rpoB*S536F mutation: a *pilQ* frame-shift (deletion of T at genomic position 5,677,812 producing a frame-shift 15 residues from the N-terminus) and two intergenic mutations (SNPs at genomic positions 1,060,840 G to A, and 5,036,891 A to C). We will identify these ancestral clones by their protein-coding changes: S536F/PilQfs and +P518. Strains were grown in minimal culture medium supplemented with L-serine as the carbon source [M9-serine: 3.75 g/l L-serine, 10.5 g/l M9 Broth, and 1 ml of 1 M $MgSO_4$].

Selection experiment

The selection experiment was conducted in conjunction with a previous study (see Gifford and MacLean 2013). Forty populations each were founded from single clones of S536F/PilQfs and +P518. Populations were grown statically (i.e. without constant shaking) at 37 °C in 96-well microtitre plates. Populations were propagated by serial transfer every 24 h for 45 transfers, diluting by 1/200 into 200 µl fresh M9-serine using a 96-pin replicator that transfers ~1µl of culture. Every 10 transfers, corresponding to approximately 76 generations, 50% of each population was stored at −80 °C in 50% glycerol.

Single-mutation fixation experiment

To determine the contribution of individual mutations, we isolated individual beneficial mutations using a mixed-marker experiment. 60 populations were initially founded by a 1:1 mix of YFP-

marked and unmarked cells. S536F/PilQfs and +P518 were tagged with a YFP genetic marker (accession no. DQ493879) using the mini-Tn7 insertion protocol (Choi and Schweizer 2006). Environment and transfer regime was as for the 270-generation experiment, except duration was reduced to 14 transfers. The YFP marker was used to detect selective sweeps. The proportion of YFP-marked cells was measured every second day from a 10 µl population sample, counting on average 10^3 cells by cytometry (BD Accuri™ C6 with Csampler™, BD Biosciences). When the YFP-marker was fixed (>90%) or lost (<10%) from a population, we stored 50% of that population at −80 °C in 50% glycerol. When populations were revived for sequencing and fitness assays, we checked that we recovered the correct marker.

Competitive fitness assays

Competitive fitnesses was assayed using flow cytometry. Evolved strains were competed against their ancestor, scaled by the initial fitness deficit of the ancestor relative to *P. aeruginosa* PAO1. To revive the frozen evolved and ancestral strains, we first pin replicated 1 µl of the frozen population into 200 µl M9-serine. Strains were grown in M9-serine for 48 h prior to competition experiments to physiologically acclimate to the medium (first for a 24 h growth period, then a 1 in 200 dilution, followed by a second 24 h growth period). We created mixtures of YFP-marked and unmarked cells and quantified fitness according to the change in frequency across a 24-hour growth period. The formula for the fitness of strain A relative to strain B over a fixed time period (24 h, approximately 7.64 generations) is $w_{AB} = \log_2(DA_f/A_i)/\log_2(DB_f/B_i)$, where subscripts i and f indicate initial and final frequencies, respectively, and D is the dilution factor used to initiate the competition experiment (1/200). Changes in fitness are expressed as selection coefficients (s) relative to ancestral fitness, $s = w_{evolved}/w_{ancestor} - 1$ (Lenski et al. 1991). Data were analyzed using R 3.0.1 (R Core Team 2013), and are available as online supplemental material.

Genomic DNA extraction and sequencing pipeline

We analyzed 22 S536F/PilQfs-derived and 18 +P518-derived lineages from the 270 generation selection experiment, and 18 S536F/PilQfs-derived 22 +P518-derived lineages from the single-mutation fixation experiment. Genomic DNA (gDNA) was extracted with the Promega Wizard Genomic DNA Purification Kit (Promega, UK). Complete details on the extraction method are available in the online supplementary methods. Sequencing was performed using Illumina HiSeq2000 with 100 bp paired-end reads, at the Wellcome Trust Centre for Human Genetics (Oxford, UK). Reads were filtered and analysed for mutations according to the pipeline described in San Millan et al. 2014 (full details also available in the online supplementary material). After filtering, median coverage depth for each clone ranged from 35X to 80X. Sequences are deposited in the NCBI Sequence Read Archive (SRP053291).

Results and discussion

Evolutionary trajectories

The ancestral strains used in our experiment were two clones isolated from a strain of *P. aeruginosa* PAO1, bearing different *rpoB* and *pilQ* alleles: S536F/PilQfs and +P518 (see methods and MacLean and Buckling, 2009). These clones had similar deficits in fitness (w) relative to their common PAO1 ancestor, and were chosen in an effort to remove fitness as a confounding effect on evolvability (S536F/PilQfs w=0.82±0.054 SE vs. +P518 w=0.784±0.031 SE, two-sample t-test: t_{10}=1.42, p=0.18).

We adapted 40 independent populations from each genetic background to growth on M9 salts with L-serine as the sole carbon source. Following 270 generations of adaptation, S536F/PilQfs lineages had higher fitness than +P518 lineages (S536F/PilQfs w=1.344±0.074 SE, +P518 w=1.044±0.065 SE, two-sample t-test: $t_{17.9}$=−2.58, p=0.02; fitness values scaled to initial

fitness deficit of ancestors). In both genetic backgrounds, fitness measured against the ancestor increased monotonically with the number of generations of evolution, until approximately generation 150. Clones that had evolved for more than 150 generations did not have higher fitness against the ancestor than those from generation 150 suggesting that either populations are close to reaching a local fitness optimum, or that mutations provide non-transitive fitness benefits (Figure 1).

Genetic basis of adaptation

We examined the underlying genetic basis for adaptation by whole-genome sequencing 40 independently evolved clones (22 of S536F/PilQfs and 18 of +P518) from the end of the experiment. Clones isolated from lineages descendant from both ancestors had between one and four mutations; although +P518-descendant clones had slightly more mutations on average, the difference was not significant (2.09±0.15 SE vs. 2.33±0.18 SE; Negative binomial GLM: intercept z=5.0, p<0.001; genotype z=0.514, p=0.61). This indicates that the difference in evolvability did not arise from a difference in the number of mutations fixed.

The genetic bases of adaptation in descendants of both ancestors was remarkably similar (Figure 2 and Tables S1 and S2), but the fitness associated with mutations was strongly contingent on ancestry. Nearly all evolved lineages acquired mutations in one of two selective targets: non-synonymous SNPs in PA2449, a TyrR-like RpoN enhancer binding protein (Lundgren et al. 2013, recently named *gcsR* for glycine cleavage system regulator, Sarwar et al. 2016), and small indels leading to putative loss-of-function mutations in two protease complexes that share similar protein targets and functions (*clpA* from the ClpAP serine protease complex, and *ftsH*, *hflC*, and *hflK* from the FtsH/HflKC complex, Lundgren et al. 2013; Zhou and Jin 1998). Mutations in pili formation genes, most often causing a loss of protein function, were a third target of selection in +P518 lineages, but not in S536F/PilQfs lineages,which already had a pili loss of function mutation (discussed below). We also observed four mutations in *cbrAB* (involved in amino acid metabolism,

196 Nishijyo et al. 2001), suggesting it may represent a fourth target of selection. We found no evidence
197 for large indels, IS transposition, or genome rearrangements.
198 The frequency of lineages with a mutation in either PA2449 or proteases (but not both) was
199 equal in both genetic backgrounds (Fisher's exact test: p=0.73). However, the fitness advantage
200 associated with mutations in these targets was greater in S536F/PilQfs lineages than +P518 lineages
201 (Figure 3; Type III ANOVA: ancestor genotype $F_{1,29}$=21.3, p<0.001, selective target $F_{1,29}$=12.0,
202 p=0.002, interaction $F_{1,29}$=4.75, p=0.038), although overlap at the level of individual mutations was
203 low (addressed below). The difference in selection coefficient was more pronounced for PA2449
204 mutations (two-sample t-test: $t_{16.21}$= 3.88, p=0.001) than protease mutations (two-sample t-test:
205 $t_{11.97}$= 1.97, p=0.07). This suggests that epistasis between the ancestral genotype and novel
206 beneficial mutations, but not the accessibility of those beneficial mutations, determined the
207 evolvability of the two ancestral genotypes.
208 Although the genetics of adaptation were highly similar in S536F/PilQfs and +P518
209 lineages, some differences are apparent. The first notable difference is that five S536F/PilQfs
210 descendant lineages had mutations in both PA2449 and a protease gene, but the fitness of these five
211 double mutants was indistinguishable from the 10 lineages with PA2449 mutations but no protease
212 mutations (Wilcoxon rank-sum test, W=37, p=0.16, Figure 3), indicating that PA2449/protease
213 double mutants were not responsible for S536F/PilQfs's greater evolvability. The second difference
214 was the acquisition of mutations type IV pili formation genes (*pil* genes, *tsaP* and *algR*) in 15 of 18
215 lineages in the +P518 background only. This difference likely arises due to the ancestral PilQfs
216 mutation in S536F/PilQfs, as mutations in *pilQ* or any of the genes mutated in evolved +P518
217 clones should result in a non-piliated phenotype (Mattick 2002; Ayers et al. 2009; Leighton et al.
218 2015; Hmelo et al. 2015). The convergence of both ancestral genotypes on pili gene defects
219 suggests that a non-piliated phenotype confers some benefit on its own. Notably, however, all but
220 one of the +P518 lineages carrying a PA2449 mutation also carried a pili gene mutation, suggesting

that the larger fitness benefit of PA2449 mutations in S536/PilQfs lineages over +P518 lineages does not arise due to the absence of pili generally. Third, although the same genes were mutated in both backgrounds, there was incomplete overlap in the specific mutations acquired in the genes targeted by selection. An alternate explanation for our results is that S536F/PilQfs by chance acquired larger-effect mutations than +P518. An approach to determine whether this is the case would be to use allelic exchange to place the acquired mutations into both ancestral backgrounds (Hmelo et al. 2015).

An assumption of this analysis is that 'accessory mutations' co-occurring with the primary adaptive mutations had negligible affects on fitness. To test this assumption, we sequenced 40 individual clones selected for single beneficial mutations in a mixed-marker experiment (see Methods; Table S3). For the 13 clones found to have only a PA2449 mutation, the fitness advantage equaled that of strains with both a PA2449 mutation and another mutation (Figure 3, Type III ANOVA: ancestor genotype $F_{1,26}$=12.7, p=0.001; presence of accessory mutation $F_{1,26}$=1.51, p=0.23, insignificant interaction term dropped). This suggests that PA2449 contributed the bulk of the fitness benefit and accessory mutations were of little consequence for evolvability, but we note that we had low power to detect small differences in the case of +P518 lineages. If we further restrict the analysis to evolved clones that possessed only a PA2449 mutation and a pili gene mutation (n=3 from each ancestral background), the difference in selection coefficient between S536F/PilQfs and +P518 remains significant (S536F/PilQfs s=1.08±0.038 SE vs. +P518 s=0.242±0.112 SE; two-sample t-test: $t_{2.42}$=6.84, p=0.01).

Environment-specific roles of the targets of selection

Of the targets of selection, PA2449 (*gscR*, Sarwar et al. 2016) has the most obvious role in L-serine metabolism. PA2449 is an RpoN enhancer binding protein that initiates RpoN-controlled transcription of genes for L-serine and L-glycine metabolism via regulation of the glycine cleavage system (Lundgren et al. 2013, Sarwar et al. 2016). The SNPs we observed in PA2449 are likely to

improve its function when L-serine is a carbon source, rather than cause loss of function, as PA2449 knockouts have severely limited growth on L-serine (Lundgren et al. 2013). PA2449 mutations are also not likely to compensate for the defect in RpoB. PA2449 mutations are only beneficial on L-serine and—to a lesser extent—L-glycine, which is part of the same metabolic pathway as L-serine (Lundgren et al. 2013, Sarwar et al. 2016). The same mutations are neutral or deleterious on other carbon sources (e.g. glucose and KB, see online supplemental methods and Figure S1). Why S536F/PilQfs should gain a bigger benefit than +P518 from PA2449 mutations is unclear. In *E. coli*, RNA polymerase (RNAP) with the S531F form of RpoB (equivalent to S536F in *P. aeruginosa*) behaves like 'stringent' RNAP, shifting gene expression toward genes regulated by the alternative sigma factors (Zhou and Jin 1998). S536F also has wild-type levels of transcriptional efficiency, but altered gene expression, relative to other *rpoB* alleles (Qi et al. 2014). However, Hall et al. 2010 found that +P518 had low evolvability among a set of *rpoB* mutants adapting to a complex environment, but did not determine the basis for the difference in evolvability (i.e. number of mutations or epistasis).Together, the data suggest that the +P518 mutation may have many pleiotropic effects on fitness. Given that PA2449 is a transcriptional activator, it is likely that the differences in evolvability stem from the mutations in *rpoB*, which is the master regulator of transcription. The convergence of +P518 and S536F/PilQfs lineages on a pili-less phenotype is indirect evidence that the *rpoB* mutations are responsible for the observed magnitude epistasis. However, we cannot strictly eliminate the possibility that the loss of PilQ in S536F/PilQfs clone specifically affects the fitness effect of PA2449 mutations, although we note that a +P518 lineage with both a SNP in PA2449 and a premature stop in PilQ did not have significantly higher fitness (w=0.948±0.0914 SE). Interestingly, mutations in DNA topoisomerase, which is also involved in bacterial gene expression, have also been shown to affect evolvability (Woods et al. 2011). It remains to be determined whether epistatic interactions occur more in core regulators of gene transcription.

In contrast, what benefit the loss of proteases and pili may provide in this environment is unclear. We observed loss-of-function mutations in four protease components from two complexes: *clpA* (from ClpAP), and *ftsH*, *hflC*, and *hflK* (from FtsH/HflKC), and several pili genes. The protease complexes are involved both in general proteolysis and disaggregation of mis-folded proteins, and specific regulation of transcription factors for fine-tuning cellular metabolism (Dougan 2013; Paget 2014). ClpA has been shown to regulate some of the same genes as PA2449 in *P. aeruginosa* PAO1, including genes involved in quorum-sensing (e.g. pyocyanin, elastase, and C4-homoserine lactone, Diggle et al. 2002). This suggests that PA2449 and proteases may both affect the expression of parts of the metabolic pathway involved in L-serine metabolism. One possibility is via RpoN, as PA2449 is an activator of RpoN-controlled transcription, and ClpA and RpoN are in the same protein-protein interaction network in *Escherichia coli*, (Rajagopala et al. 2014). Loss of pili is likely an adaptation to growth in a non-shaken environment, as the absence of pili is associated with an increase in swimming speed in *P. aeruginosa* PAO1 (Taylor and Buckling 2010). Alternatively, the loss of pili production may also reduce demand for RpoN, as pilus production is under RpoN control (Ishimoto and Lory 1989). Several of the accessory genes mutated are also transcriptional regulators or are involved in RpoN-controlled transcription (Table S2), suggesting that increasing RpoN availability may nominally increase fitness.

Among our populations, we found no evidence for reversion of resistance through back mutations in *rpoB*, nor any evidence for compensatory mutations through second-site mutations in any RNA polymerase genes. The absence of reversion is consistent with previous evolution studies with rifampicin resistant populations in antibiotic-free environments (e.g. Hall et al. 2010). However, the absence of any putative compensatory mutations in *rpoA*, *rpoB* or *rpoC* contrasts with previous studies (Reynolds 2000; Hall et al. 2011; Comas et al. 2012), suggesting improving L-serine metabolism confers larger fitness benefits than repairing RNA transcription. Previous results found that selecting rifampicin-sensitive PAO1 on L-serine also results in PA2449, *clpA*, and pili

296 gene mutations (Gifford et al. 2015), strongly suggesting that the observed mutations reflect general
297 adaptation to the culture medium or laboratory conditions, and not compensatory adaptation to
298 recover the cost of rifampicin resistance. The absence of reversions and compensatory mutations is
299 surprising, given the large initial cost of rifampicin resistance born by both mutants. The ability to
300 adapt to or compensate for the costs of antibiotic resistance is thought to influence which specific
301 resistance mutations are maintained at the population level over evolutionary time scales
302 (Andersson 2006), but we were not able to test this prediction here. This highlights the importance
303 of considering mutational supply in predicting adaptive trajectories; as PA2449 mutations are
304 beneficial in the rifampicin-sensitive PAO1 background (Gifford et al. 2015), reversion should still
305 confer a benefit. Although reversion is a valid mutational step that would allow further adaptation,
306 the probability of the mutations is low enough that the lineages were effectively trapped in the
307 region of a fitness peak by epistasis, and by the availability (or lack thereof) of reversion mutations.
308

309 **Conclusions**

310 Our results suggest that magnitude epistasis can constrain evolvability at a practical level, in
311 agreement with previous studies (Woods et al. 2011; Kryazhimskiy et al. 2014; Szamecz et al.
312 2014). However, further work is needed to determine which genes may constrain magnitude
313 epistasis at a global level. Our results are particularly relevant to the recovery of fitness costs
314 associated with antibiotic resistance. Chromosomal antibiotic resistance mutations often occur in
315 genes involved in gene expression and protein synthesis, and thus are likely to interact with many
316 other genes on a global level (e.g. *rpoB* for rifampicin, *gyrA* for fluoroquinolones, *rpsL* for
317 streptomycin, *fusA* for fusidic acid, Andersson and Hughes 2010). If resistance mutations have a
318 tendency to interact epistatically with generally beneficial mutations, this could have important
319 consequences for the long-term maintenance of resistance in antibiotic-free environments. Previous
320 studies of compensation in antibiotic resistant lineages have focused on the short-term

consequences of fitness costs. Our results are the first to show that resistance can also have long-term consequences for evolvability, which could have important implications in the migration and subsequent adaptation of resistant strains between agricultural and clinical hosts.

Acknowledgements

We thank Mila Kojadinovic for assistance in gDNA extractions. 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 to RCM. DRG was funded by the National Sciences and Engineering Research Council (NSERC, Canada), the Clarendon Fund and Somerville College. We thank the High-Throughput Genomics Group at the Wellcome Trust Centre for Human Genetics (funded by Wellcome Trust grant reference 090532/Z/09/Z and MRC Hub grant G090074791070) for the generation of the sequencing data.

Authors' contributions

Conceived of the experiment: DRG and RCM. Conducted laboratory work: DRG. Wrote and executed mutation calling pipeline: MT-R. Analysed data and wrote the manuscript: all authors.

Conflicts of interest

The authors declare no conflicts of interest.

References

Andersson DI, Hughes D. 2010. Antibiotic resistance and its cost: is it possible to reverse resistance? *Nature Reviews Microbiology* 8:260-271.

Andersson DI. 2006. The biological cost of mutational antibiotic resistance: any practical

346 conclusions? *Current Opinion in Microbiology* 9:461-465.

347 Ayers, M, Sampaleanu, L, Tammam, S, Koo, J, Harvey, H, Howell, P, and Burrows, L. 2009.
348 PilM/N/O/P proteins form an inner membrane complex that affects the stability of the
349 *Pseudomonas aeruginosa* type IV pilus secretin. *Journal of molecular biology*, 394(1):128-
350 142.

351 Bailey, SF, Rodrigue, N, and Kassen, R. 2015. The effect of selection environment on the
352 probability of parallel evolution. *Molecular Biology and Evolution* (online early).

353 Barrick, JE, Kauth, MR, Strelioff, CC, & Lenski, RE. 2010. *Escherichia coli rpoB* mutants have
354 increased evolvability in proportion to their fitness defects. *Molecular Biology and Evolution*,
355 27: 1338-1347.

356 Choi KH, Schweizer HP. 2006. mini-Tn7 insertion in bacteria with single *att*Tn7 sites: example
357 *Pseudomonas aeruginosa*. *Nature Protocols* 1:153-161.

358 Chou HH, Chiu HC, Delaney NF, Segrè D, Marx CJ. 2011. Diminishing returns epistasis among
359 beneficial mutations decelerates adaptation. *Science* 332:1190-1192.

360 Chou, HH, Delaney, NF, Draghi, JA, Marx, CJ. 2014. Mapping the fitness landscape of gene
361 expression uncovers the cause of antagonism and sign epistasis between adaptive mutations.
362 *PLoS Genetics*, *10*(2), e1004149.

363 Colegrave N, Collins S. 2008. Experimental evolution: experimental evolution and evolvability.
364 *Heredity* 100:464-470.

365 Comas I, Borrell S, Roetzer A, Rose G, Malla B, Kato-Maeda M, Galagan J, Niemann S, Gagneux
366 S. 2012. Whole-genome sequencing of rifampicin-resistant *Mycobacterium tuberculosis*
367 strains identifies compensatory mutations in RNA polymerase genes. *Nature Genetics* 44:106-
368 110.

369 Costanzo M, Baryshnikova A, Bellay J, Kim Y, Spear ED, Sevier CS, Ding H, Koh JL, Toufighi K,

Mostafavi S, et al. 2010. The genetic landscape of a cell. *Science* 327:425-431.

de Visser JAGM, Hoekstra RF, Van den Ende H. 1997. An experimental test for synergistic epistasis and its application in Chlamydomonas. *Genetics* 145:815-819.

de Visser JAGM, Hoekstra RF, Van den Ende H. 1997. Test of interaction between genetic markers that affect fitness in *Aspergillus niger*. *Evolution* 51:1499-1505.

de Visser JAGM, Cooper, TF, Elena SF. 2011. The cause of epistasis. *Proceedings of the Royal Society B* 278:3617-3624.

Diggle SP, Winzer K, Lazdunski A, Williams P, Cámara M. 2002. Advancing the quorum in *Pseudomonas aeruginosa*: MvaT and the regulation of N-acylhomoserine lactone production and virulence gene expression. *Journal of bacteriology* 184:2576-2586.

Dougan DA. 2013. *Regulated Proteolysis in Microorganisms*. volume 66. Springer.

Franke, J, Klözer, A, de Visser, JAGM, and Krug, J. 2011. Evolutionary accessibility of mutational pathways. *PLoS Computational Biology*, 7(8):e1002134.

Gifford DR, MacLean RC. 2013. Evolutionary reversals of antibiotic resistance in experimental -populations of *Pseudomonas aeruginosa*. *Evolution* 67:2973-2981.

Gifford DR, Schoustra SE, Kassen R. 2011. The length of adaptive walks is insensitive to starting fitness in *Aspergillus nidulans*. *Evolution* 65:3070-3078.

Gifford DR, Toll-Riera M, Kojadinovic M, MacLean RC. 2015. Here's to the losers. *The American Naturalist*, 186(1):41-49.

Hall AR, Griffiths VF, MacLean RC, Colegrave N. 2010. Mutational neighbourhood and mutation supply rate constrain adaptation in *Pseudomonas aeruginosa*. *Proceedings of the Royal Society B* 277:643-650.

Hall AR, Iles JC, MacLean RC. 2011. The fitness cost of rifampicin resistance in *Pseudomonas*

393 *aeruginosa* depends on demand for RNA polymerase. *Genetics* 187:817-822.

394 Hmelo LR, Borlee BR, Almblad H, Love ME, Randall TE, Tseng BS, Lin C, Irie Y, Storek KM,
395 Yang JJ, Siehnel RJ, Howell PL, Sing, PK, Tolker-Nielsen TT, Parsek MR, Schweizer HP,
396 Harrison JJ. 2015. Precision-engineering the Pseudomonas aeruginosa genome with two-step
397 allelic exchange. *Nature Protocols* 10(11):1820-1841.

398 Ishimoto KS, Lory S. 1989. Formation of pilin in Pseudomonas aeruginosa requires the alternative
399 sigma factor (RpoN) of RNA polymerase. *Proceedings of the National Academy of Sciences*
400 86:1954-1957.

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

403 Kryazhimskiy, S, Dushoff, J, Bazykin, GA, and Plotkin, JB. 2011. Prevalence of epistasis in the
404 evolution of influenza A surface proteins. *PLoS Genetics*, 7(2):e1001301.

405 Kryazhimskiy S, Rice DP, Jerison E, Desai MM. 2014. Global epistasis makes adaptation
406 predictable despite sequence-level stochasticity. *Science* 344:1519-1522.

407 Kvitek DJ, Sherlock G. 2011. Reciprocal sign epistasis between frequently experimentally evolved
408 adaptive mutations causes a rugged fitness landscape. *PLoS Genetics* 7:e1002056.

409 Lalić J, Elena SF. 2012. Magnitude and sign epistasis among deleterious mutations in a positive-
410 sense plant RNA virus. *Heredity* 109:71-77.

411 Leighton, TL, Buensuceso, R, Howell, PL, and Burrows, LL. (2015). Biogenesis of Pseudomonas
412 aeruginosa type IV pili and regulation of their function. *Environmental Microbiology* (online
413 early).

414 Lenski, RE, Rose, MR, Simpson, SC, and Tadler, SC. 1991. Long-term experimental evolution in
415 *Escherichia coli*. I. Adaptation and divergence during 2,000 generations. *The American*

Naturalist, 138(6):1315-1341.

Lundgren B, Thornton W, Dornan M, Villegas-Peñaranda L, Boddy C, Nomura C. 2013. The gene PA2449 is essential for glycine metabolism and pyocyanin biosynthesis in *Pseudomonas aeruginosa* PAO1. *Journal of Bacteriology* 195:2087-2100.

Lunzer M, Golding G, Dean A. 2010. Pervasive cryptic epistasis in molecular evolution. *PLoS Genetics* 6:e1001162.

MacLean R, Hall A, Perron G, Buckling A. 2010. The population genetics of antibiotic resistance: integrating molecular mechanisms and treatment contexts. *Nature Reviews Genetics* 11:405-414.

MacLean RC, Buckling A. 2009. The distribution of fitness effects of beneficial mutations in *Pseudomonas aeruginosa*. *PLoS Genetics* 5:e1000406, doi:10.1371/journal.pgen.1000406.

Martin PR, Hobbs M, Free PD, Jeske Y, & Mattick JS. 1993. Characterization of *pilQ*, a new gene required for the biogenesis of type 4 fimbriae in *Pseudomonas aeruginosa*. *Molecular Microbiology* 9(4):857-868.

Martin G and Lenormand T. 2006. A general multivariate extension of Fisher's geometrical model and the distribution of fitness effects across species. *Evolution*, 60(5):893-907.

Mattick, JS. 2002. Type IV pili and twitching motility. *Annual Reviews in Microbiology*, 56(1):289-314.

Nishijyo T, Haas D, Itoh Y. 2001. The CbrA-CbrB two-component regulatory system controls the utilization of multiple carbon and nitrogen sources in Pseudomonas aeruginosa. *Molecular microbiology* 40:917-931.

Paget M. 2014. Prokaryotic Gene Regulation by Sigma Factors and RNA Polymerase. In: *Molecular Life Sciences*. Springer.

439 Persat A, Inclan YF, Engel JN, Stone HA, and Gitai Z. 2015. Type IV pili mechanochemically
440 regulate virulence factors in *Pseudomonas aeruginosa*. *Proceedings of the National Academy*
441 *of Sciences*, 112(24):7563-7568.

442 Poelwijk, FJ, Kiviet, DJ, Weinreich, DM, and Tans, SJ. 2007. Empirical fitness landscapes reveal
443 accessible evolutionary paths. *Nature*, 445(7126):383-386.

444 Phillips PC, Otto SP, Whitlock MC. 2000. Beyond the average: the evolutionary importance of gene
445 interactions and variability of epistatic effects. In: JB Wolf, ED Brodie, MJ Wade, editors.
446 *Epistasis and the evolutionary process*. chapter 2. Oxford University Press.

447 Qi Q, Preston GM, MacLean RC. 2014. Linking System-Wide Impacts of RNA Polymerase
448 Mutations to the Fitness Cost of Rifampin Resistance in Pseudomonas aeruginosa. *mBio*
449 5:e01562-14.

450 Rajagopala SV, Sikorski P, Kumar A, Mosca R, Vlasblom J, Arnold R, Franca-Koh J, Pakala SB,
451 Phanse S, Ceol A, Häuser R, Siszler G, Wuchty S, Emili A, Babu M, Aloy P, Pieper R, Uetz P.
452 2014. The binary protein-protein interaction landscape of *Escherichia coli*. *Nature*
453 *Biotechnology* 32:285-290.

454 Reynolds MG. 2000. Compensatory evolution in rifampin-resistant *Escherichia coli*. *Genetics*
455 156:1471-1481.

456 Rokyta DR, Joyce P, Caudle SB, Miller C, Beisel CJ, Wichman HA. 2011. Epistasis between
457 beneficial mutations and the phenotype-to-fitness map for a ssDNA virus. *PLoS Genetics*
458 7:e1002075.

459 Romero PA, Arnold FH. 2009. Exploring protein fitness landscapes by directed evolution. *Nature*
460 *Reviews Molecular Cell Biology* 10(12):866-76.

461 R Core Team. 2013. R: a language and environment for statistical computing. R Foundation for

462 Statistical Computing, Vienna, Austria.

463 San Millan A, Peña-Miller R, Toll-Riera M, Halbert Z, McLean A, Cooper B, MacLean R. 2014.
464 Positive selection and compensatory adaptation interact to stabilize non-transmissible
465 plasmids. *Nature Communications* 5.

466 Sarwar Z, Lundgren BR, Grassa MT, Wang MX, Gribble M, Moffat JF, Nomura CT. 2016. GcsR, a
467 TyrR-like enhancer-binding protein, regulates expression of the glycine cleavage system in
468 *Pseudomonas aeruginosa* PAO1. *mSphere* 1(2):e00020-16.

469 Segre D, DeLuna A, Church GM, Kishony R. 2004. Modular epistasis in yeast metabolism. *Nature*
470 *Genetics* 37:77-83.

471 Schenk MF, Szendro IG, Salverda ML, Krug J, and de Visser JAGM. 2013. Patterns of epistasis
472 between beneficial mutations in an antibiotic resistance gene. *Molecular Biology and*
473 *Evolution*, 30(8):1779-1787.

474 Szamecz B, Boross G, Kalapis D, Kovács K, Fekete G, Farkas Z, Lázár V, Hrtyan M, Kemmeren P,
475 Groot Koerkamp MJA, Rutkai E, Holstege FCP, Papp B and Pál C. (2014). The genomic
476 landscape of compensatory evolution. *PLoS Biology*, 12(8):e1001935-e1001935.

477 Taylor, TB and Buckling, A. 2010. Competition and dispersal in *Pseudomonas aeruginosa*. *The*
478 *American Naturalist*, 176(1):83-89.

479 Taylor, TB and Buckling, A. 2011. Selection experiments reveal trade-offs between swimming and
480 twitching motilities in *Pseudomonas aerugionsa*. *Evolution*, 65(11):3060-3069.

481 Weinreich D, Watson R, Chao L. 2005. Perspective: sign epistasis and genetic constraint on
482 evolutionary trajectories. *Evolution* 59:1165-1174.

483 Weinreich, D, Delaney, N, DePristo, M, and Hartl, D. 2006. Darwinian evolution can follow only

very few mutational paths to fitter proteins. *Science*, 312(5770):111-114.

Whitlock MC, Phillips PC, Moore FBG, Tonsor SJ. 1995. Multiple fitness peaks and epistasis. *Annual Review of Ecology and Systematics* :601-629.

Wong, A, Seguin K. 2015. Effects of genotype on rates of substitution during experimental evolution. *Evolution* (Accepted Article, accessed 14/07/2015). doi: 10.1111/evo.12700.

Woods RJ, Barrick JE, Cooper TF, Shrestha U, Kauth MR, Lenski RE. 2011. Second-order selection for evolvability in a large *Escherichia coli* population. *Science* 331:1433-1436.

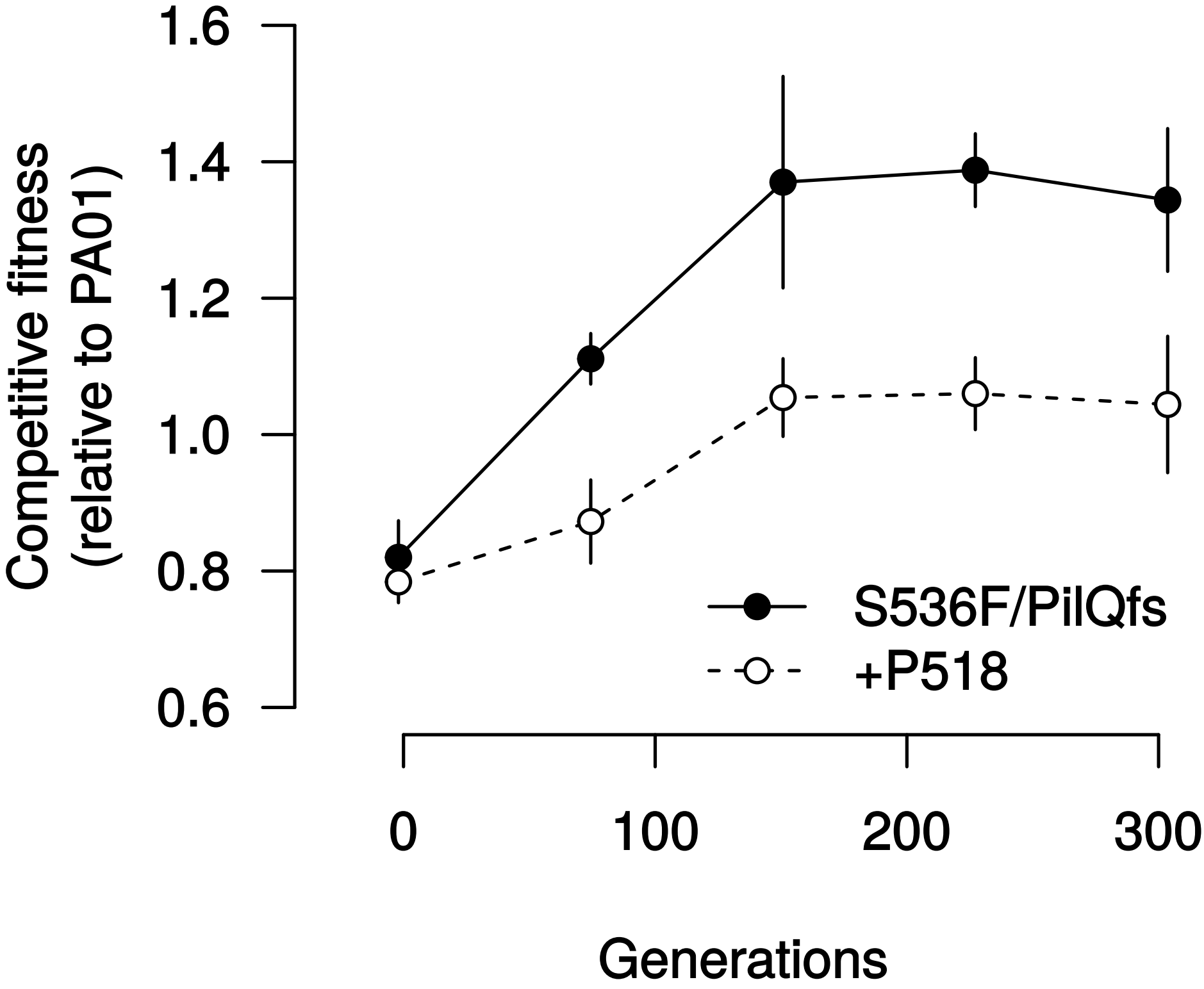
Zhou YN, Jin DJ. 1998. The *rpoB* mutants destabilizing initiation complexes at stringently controlled promoters behave like "stringent" RNA polymerases in *Escherichia coli*. *Proceedings of the National Academy of Science of the United Statesof America* 95:2908-2913.

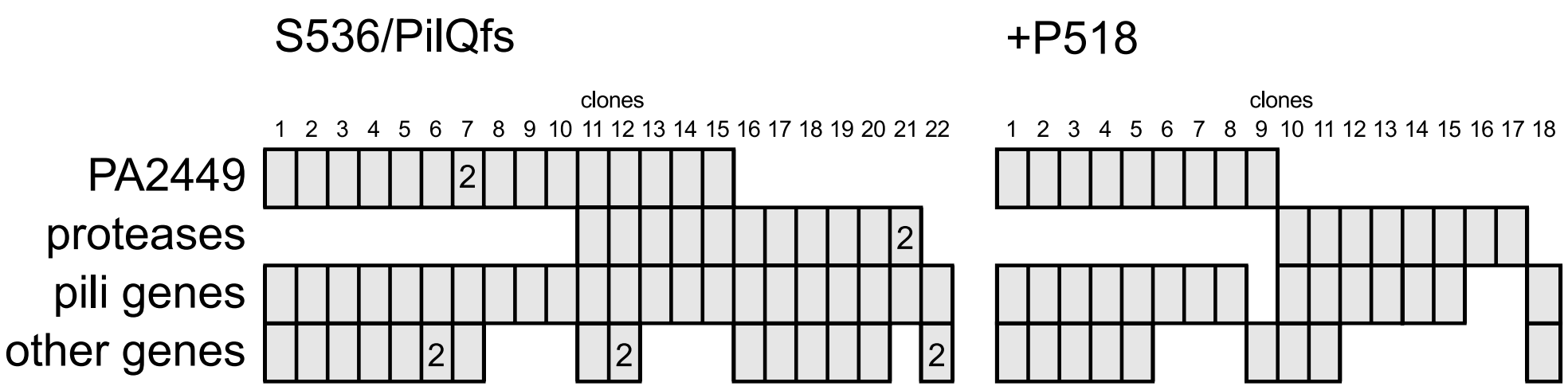
Figures and tables

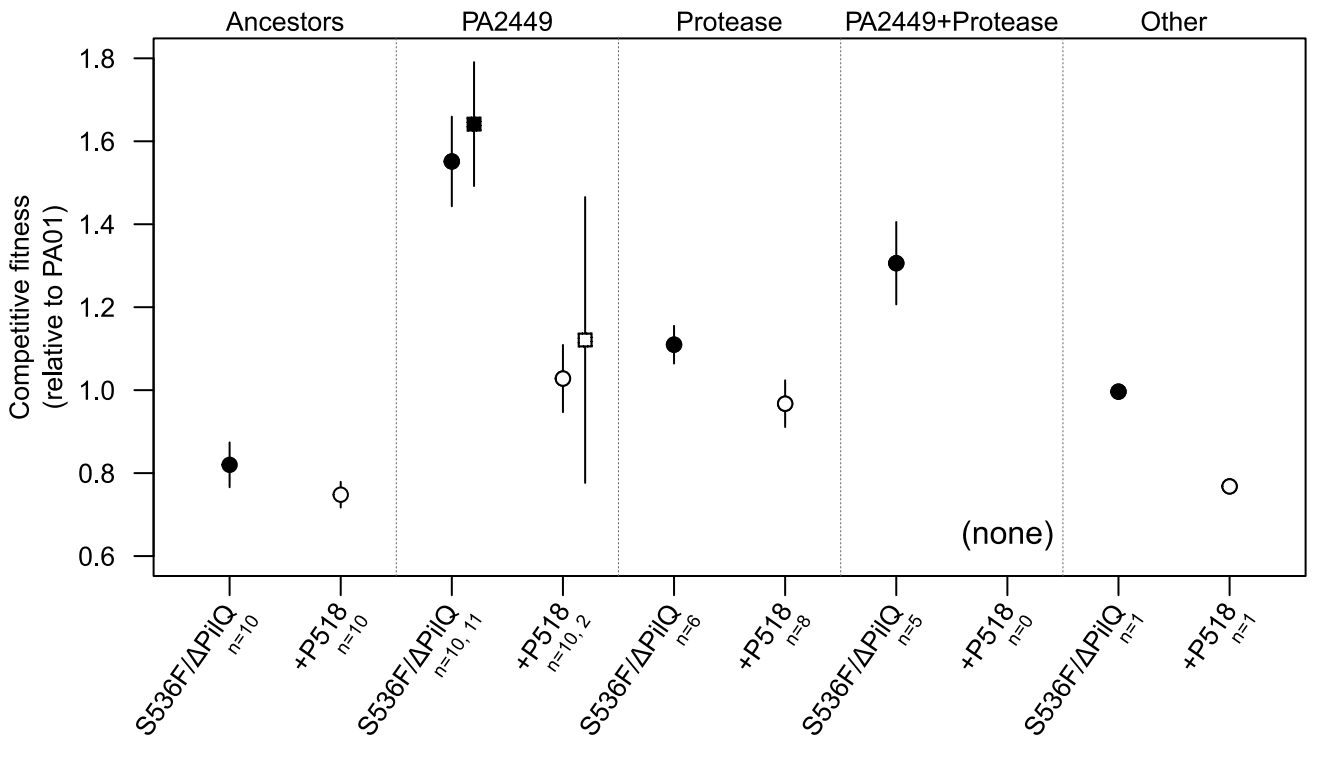
Figure 1: Average fitness (±SE) of *P. aeruginosa* lineages selected on L-serine (40 total populations, averaged over subsets of *n*=10 separate populations at each time point, normalized by initial fitness cost associated with *rpoB* mutations).

Figure 2: Mutations in evolved clones. Columns indicate individual clones and rows indicate genes that were mutated. Filled boxes indicate that a mutation was found in that gene (multiple filled boxes in a single column indicate the presence of mutations in multiple genes in the same lineage). Striped boxes indicate the presence of a mutation in the ancestral clone. Where indicated, a number indicates the presence of more than one mutation. Mutated protease genes include *clpA*, *ftsH*, *hflC* and *hflK*. Mutated pili-related genes include *pilB*, *pilC*, *pilE*, *pilI*, *pilM*, *pilQ*, *pilR*, *pilW*, *pilY1*, *tasP*, and *algR* (Mattick 2002; Ayers et al. 2009; Leighton et al. 2015). See Table S1 for specific mutations found in each gene.

Figure 3: Average fitness (±SE) of *P. aeruginosa* clones bearing mutations in either PA2449, protease gene(s), both PA2449 and a protease gene, or neither. Fitness was normalized by initial fitness cost of ancestor strain relative to PAO1. Circles denote clones from the 270 generation selection experiment (also bearing one or more accessory mutations), and squares denote clones from the single-mutation fixation experiment bearing only a single mutation in PA2449 (See Table S1-S3 for a list of all genes mutated in all strains).


Competitive fitness
(relative to PA01)
1.6
1.4
1.2
1.0
0.8
0.6
S536F/PilQfs
+P518
0
100
200
300
Generations


S536/PilQfs
+P518
clones
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22
clones
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
PA2449
proteases
pili genes
other genes
2
2
2
2
2


Ancestors
PA2449
Protease
PA2449+Protease
Other
Competitive fitness
(relative to PA01)
1.8
1.6
1.4
1.2
1.0
0.8
0.6
(none)
S536F/ΔPilQ n=10
+P518 n=10
S536F/ΔPilQ n=10, 11
+P518 n=10, 2
S536F/ΔPilQ n=6
+P518 n=8
S536F/ΔPilQ n=5
+P518 n=0
S536F/ΔPilQ n=1
+P518 n=1

Epistatic interactions between ancestral genotype and beneficial mutations shape evolvability in *Pseudomonas aeruginosa*

Danna R. Gifford[1,2,*] Macarena Toll-Riera[1,3], and R. Craig MacLean[1]

Supplementary methods

Genomic DNA extraction and sequencing pipeline

Whole genome sequencing analysis was performed as described in San Millan et al. (2014). Whole genome sequencing services were provided by the Wellcome Trust Centre for Human Genetics (Oxford, UK) using the Illumina HiSeq 2000 platform with 100 bp paired-end reads. Initial read filtering was done using NIH QC Toolkit (Patel and Jain, 2012). 5' or 3' ends were trimmed if the Phred quality score was less than 20. Reads were discarded if they were shorter than 50 bp after trimming, if more than 2% of bases were ambiguous, or if more than 20% of bases had a Phred score < 20. BWA was used to map filtered reads were mapped to the P. aeruginosa PA01 reference genome (NC_002516.2). Mapped reads were processed to increase the quality of the variant calling: 1) reads with multiple best hits were discarded; 2) duplicated reads were discarded using MarkDuplicates from the Picard package (http://picard.sourceforge.net); 3) reads around indels were locally realigned using RealignerTargetCreator and IndelRealigner from the GATK package to correct for misalignment; and 4) mate pairs were sorted using FixMateInformation in the Picard package. Variant calling was performed with GATK's Unified Genotyper (De Pristo et al., 2011) and Samtools's mpileup (Li et al., 2009). VCFtools (vcf-annotate, Danecek et al., 2011) and GATK toolkit (VariantFiltration, DePristo et al., 2011) were used to filter the raw variants for strand bias, end distance bias, base quality bias, SNPs around gaps, low coverage and erroneously high coverage. Variants were combined using GATK's CombineVariants (keeping any unfiltered). High quality variants not filtered were annotated using SnpEff (Cingolani et al., 2012).

Three approaches were used to detect structural variants. First, BreakDancer (Chen et al., 2009) was used to predict deletions, insertions, inversions, and translocations using deviations in the separation or orientation of mapped read pairs. Pindel (Ye et al., 2009) was used to infer deletions, short insertions, long insertions, inversions, tandem duplications, and breakpoints using a split-read approach (the output of BreakDancer was also fed to Pindel to improve its output). Finally, Control-FREEC (Boeva et al., 2011) was used to detect copy number variants (CNVs). Control-FREEC finds CNVs using depth-of-coverage (normalized by GC-content). Regions of low mappability were excluded by supplying Control-FREEC with mappability tracks generated by gem-mappability (GEM library, Marco-Sola et al., 2012).

Fitness assays in other environments

In addition to M9-serine, we assayed fitness of PA2449 mutants in three additional media. We replaced L-serine with other carbon sources, to test whether fitness gains were specific to adaptation to growth on L-serine. This included M9-glycine: 4.01 g/l L-glycine; M9-glucose: 12.5 g/l *α-D*-glucose; and M9KB: 10 g/l glycerol, 10 g/l proteose peptone no. 3; each with 10.5 g/l M9 Broth, and 1 ml of 1 M $MgSO_4$. Competition experiments were performed as described in the main text.

References

Chen, K., J. W. Wallis, M. D. McLellan, D. E. Larson, J. M. Kalicki, C. S. Pohl,S. D. McGrath, M. C. Wendl, Q. Zhang, D. P. Locke, et al. 2009. BreakDancer: an algorithm for high-resolution mapping of genomic structural variation. Nature Methods 6:677–681.

Cingolani, P., A. Platts, M. Coon, T. Nguyen, L. Wang, S. J. Land, X. Lu, D. M. Ruden, et al. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. Fly 6:80–92.

Danecek, P., A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, et al. 2011. The variant call format and VCFtools. Bioinformatics 27:2156–2158.

DePristo, M. A., E. Banks, R. Poplin, K. V. Garimella, J. R. Maguire, C. Hartl,A. A. Philippakis, G. del Angel, M. A. Rivas, M. Hanna, et al. 2011. A frame-work for variation discovery and genotyping using next-generation DNA sequencing data. Nature Genetics 43:491–498.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth,G. Abecasis, R. Durbin, et al. 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25:2078–2079.

Marco-Sola, S., M. Sammeth, R. Guigó, and P. Ribeca. 2012. The GEM map-per: fast, accurate and versatile alignment by filtration. Nature Methods 9:1185–1188.

Patel, R. K., and M. Jain. 2012. NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. PLoS One 7:e30619.

Ye, K., M. H. Schulz, Q. Long, R. Apweiler, and Z. Ning. 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.

Supplementary tables and figure captions

Table S1: Mutations in selection lineages (order does not imply order of substitution). Bold text indicates putative loss of function mutations (e.g. stop codons, frame-shifts, or indels in key functional domains). Mutations in intergenic regions are abbreviated as 'intergenic'. '+' and '–' indicate insertions and deletions of listed amino acids starting at given position.

	Gene 1	Gene 2	Gene 3	Gene 4	Mutation 1	Mutation 2	Mutation 3	Mutation 4
PA2449 (enhancer binding protein involved in regulation of L-glycine/L-serine metabolism)								
S536F/ΔPilQ	PA2449				L143P			
	PA2449				R326Q			
	PA2449				F428L			
	PA2449	PA1046			L154V	silent		
	PA2449	**PA1051**			L191V	**frame-shift**		
	PA2449	**PA1269**			A94V	**stop gained**		
	PA2449	PA2122			G185C	H81Y		
	PA2449	*morA*			P288L	M1035T		
	PA2449	PA2438	intergenic		R77C	+TI15	intergenic	
	PA2449	PA2449	*bauB*		S308P	A89D	T21I	
+P518	PA2449	***pilI***			E180G	**frame-shift**		
	PA2449	PA1269			V327L	A55T		
	PA2449	*pilR*			H4Q	T364P		
	PA2449	***pilQ***			L293Q	**stop gained**		
	PA2449	***pilC***	intergenic		P288L	**frame-shift**	intergenic	
	PA2449	***pilW***	intergenic		P288L	**frame-shift**	intergenic	
	PA2449	*pilW*	intergenic		S308P	C93R	intergenic	
	PA2449	***pilY1***	intergenic		S308P	**frame-shift**	intergenic	
	PA2449	***pilR***	***gacS***	intergenic	R77C	**VPPLRE303E**	**frame-shift**	intergenic
Protease mutants (*clpA, ftsH, hflK*)								
S536F/ΔPilQ	***clpA***	*czcS*			**frame-shift**	T175P		
	clpA	***hflK***			**frame-shift**	**stop gained**		
	ftsH	*gbdR*			R69C	C310R		
	ftsH	*cbrA*			L81P	F71L		
	hflK	*cbrA*			**frame-shift**	LA324–		
	hflK	PA4631			**frame-shift**	F143L		
+P518	***clpA***				**frame-shift**			
	clpA				**frame-shift**			
	clpA	**PA0020**			**GREH195–**	**frame-shift**		
	clpA	***pilB***			**frame-shift**	**frame-shift**		
	clpA	***pilE***			**stop gained**	**frame-shift**		
	clpA	*algR*			**frame-shift**	R235W		
	clpA	PA1365	***pilM***		**frame-shift**	NT794T	**frame-shift**	
	clpA	PA2063	*pilB*		**stop gained**	V265I	H421R	
PA2449 and protease (*clpA, ftsH, hflC*) double mutants								
S536F/ΔPilQ	PA2449	***clpA***			I184S	**frame-shift**		
	PA2449	*clpA*			S308P	V223G		
	PA2449	***hflC***			K314M	**stop gained**		
	PA2449	***ftsH***	*hutC*		R76L	**VFMF115V**	I34T	
	PA2449	***hflC***	*fruR*	PA0340	G185S	**frame-shift**	R57P	L88V
Others								
S536F/ΔPilQ	PA2897	*cbrA*			V109G	A408T		
+P518	*pilB*	*cbrB*			I129T	S216P		

Table S2: Description of genes mutated during long-term selection on L-serine.

Locus ID	Gene	Function	Relation to RpoN or L-serine metabolism?
PA0020		hypothetical protein	
PA0131	*bauB*	beta-alanine catabolism	
PA0340		hypothetical protein (putative permease)	
PA0410	*pilI*	twitching motility protein	Part of an RpoN-regulated network
PA0928	*gacS*	sensor/response regulator hybrid	RpoN represses GacA
PA1046		hypothetical protein (possible beta galactosidase)	
PA1051		putative gluconate symporter and related permeases	
PA1269		transcriptional regulator	GntR regulator family
PA1365		probable siderophore receptor	
PA2063		hypothetical protein	
PA2122		hypothetical protein	
PA2438		hypothetical protein 47% similarity with HflC	
PA2449		probable enhancer binding protein	RpoN bEBP
PA2524	*czcS*	phosphorelay sensor kinase activity	
PA2620	*clpA*	ATP-binding protease component	Protease component
PA2897		transcriptional regulator with aminotransferase domain	Amino acid transport and metabolism, upstream of *algU,* GntR regulator family
PA3563	*fruR*	fructose transport system repressor	Carbon utilization modulator
PA4526	*pilB*	type 4 fimbrial biogenesis protein	Part of an RpoN-regulated network
PA4527	*pilC*	type 4 fimbrial biogenesis protein	Part of an RpoN-regulated network
PA4547	*pilR*	two-component response regulator	RpoN bEBP
PA4552	*pilW*	type 4 fimbrial biogenesis protein	Part of an RpoN-regulated network
PA4554	*pilY1*	type 4 fimbrial biogenesis protein	Part of an RpoN-regulated network
PA4556	*pilE*	type 4 fimbrial biogenesis protein	Part of an RpoN-regulated network
PA4601	*morA*	motility regulator	Upstream of *glyA3*
PA4631		nucleoside-diphosphate-sugar epimerase	
PA4725	*cbrA*	two-component sensor (histidine kinase)	Part of two-component RpoN bEBP
PA4726	*cbrB*	two-component response regulator	Part of two-component RpoN bEBP
PA4751	*ftsH*	ATP-dependent zinc protease	Protease component
PA4941	*hflC*	protease subunit	Protease component
PA4942	*hflK*	protease subunit	Protease component
PA5040	*pilQ*	type 4 fimbrial biogenesis outer membrane protein precursor	Part of an RpoN-regulated network
PA5044	*pilM*	type 4 fimbrial biogenesis protein	Part of an RpoN-regulated network
PA5105	*hutC*	histidine utilization repressor	GntR regulator family
PA5261	*algR*	alginate biosynthesis regulatory protein	*algC* probably RpoN transcribed
PA5380	*gbdR*	glycine betaine- and dimethylglycine-responsive regulator	Betaine is a byproduct of L-serine catabolism via *pssA*

bEBP—bacterial enhancer binding protein

Table S3: PA2449 mutations fixed in single mutation selection experiment.

	Gene	Mutation
S536F	PA2449	R77C
S536F	PA2449	R77C
S536F	PA2449	L143V
S536F	PA2449	P181L
S536F	PA2449	P181S
S536F	PA2449	R183C
S536F	PA2449	A189V
S536F	PA2449	H193D
S536F	PA2449	P263L
S536F	PA2449	P263L
S536F	PA2449	S308P
+P518	PA2449	G22D
+P518	PA2449	I184V

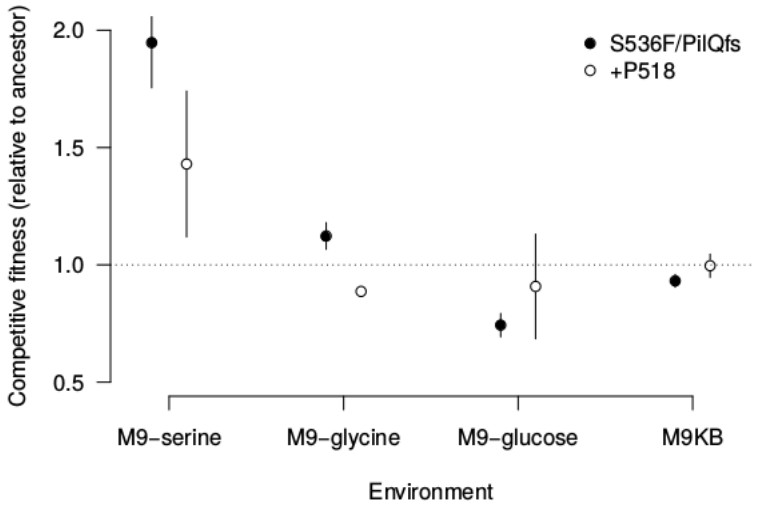
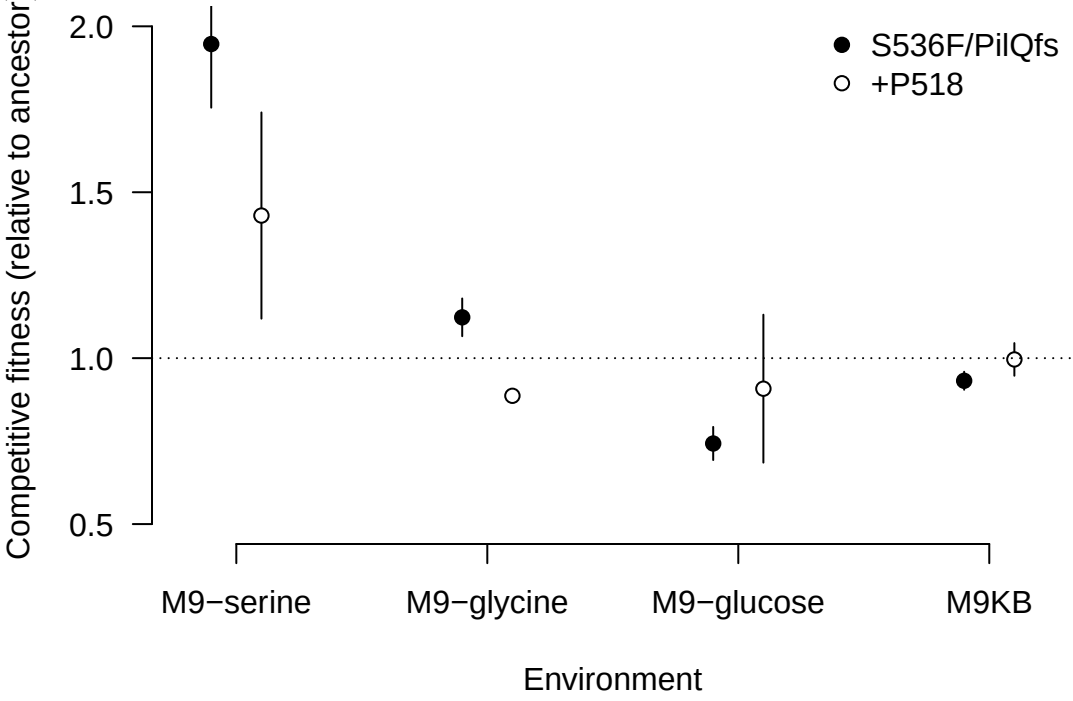


Figure S1: Fitness of PA2449 mutants assayed in M9-serine, M9-glycine, M9-glucose, and M9KB.


Competitive fitness (relative to ancestor)
2.0
1.5
1.0
0.5
S536F/PilQfs
+P518
M9−serine
M9−glycine
M9−glucose
M9KB
Environment