Compensatory mutations improve general permissiveness to antibiotic resistance plasmids

Wesley Loftie-Eaton1,2, Kelsie Bashford1, Hannah Quinn3, Kieran Dong1, Jack Millstein1,2, Samuel Hunter2, Maureen K. Thomason4, Houra Merrikh4, Jose M. Ponciano5 and Eva M. Top1,2,*

Horizontal gene transfer mediated by broad-host-range plasmids is an important mechanism of antibiotic resistance spread. While not all bacteria maintain plasmids equally well, plasmid persistence can improve over time, yet no general evolutionary mechanisms have emerged. Our goal was to identify these mechanisms and to assess if adaptation to one plasmid affects the permissiveness to others. We experimentally evolved *Pseudomonas* sp. H2 containing multidrug resistance plasmid RP4, determined plasmid persistence and cost using a joint experimental–modelling approach, resequenced evolved clones, and reconstructed key mutations. Plasmid persistence improved in fewer than 600 generations because the fitness cost turned into a benefit. Improved retention of naive plasmids indicated that the host evolved towards increased plasmid permissiveness. Key chromosomal mutations affected two accessory helicases and the RNA polymerase β-subunit. Our and other findings suggest that poor plasmid persistence can be caused by a high cost involving helicase–plasmid interactions that can be rapidly ameliorated.

Analyses of the genome sequences of prokaryotes have clearly revealed that horizontal gene transfer (HGT) plays an important role in their evolution. It mediates wholesale acquisition of genes encoding traits that can be advantageous. A prime example of great concern to human health is the plasmid-mediated spread of antibiotic resistance and virulence genes, such as the recent worldwide dissemination of resistance to colistin, an antibiotic of last resort. Plasmids are very common in bacteria and those with a broad host range play a critical role in gene spread among phylogenetically distinct bacteria. The same plasmid types found before the antibiotic era now often contain multiple antibiotic resistance genes, and are present in both clinical and environmental isolates.

Not all bacteria maintain plasmids equally well, with great differences even between members of the same genus. To curb the spread of plasmid-mediated antibiotic resistance and virulence, it is vital that we understand the basis of these differences and how so-called ‘plasmid permissiveness’ can evolve over time. Persistence of a plasmid in a bacterial population hinges on correctly balancing efficient replication, accurate segregation, plasmid cost, conjugative transfer, post-segregational killing and selection on beneficial accessory traits. While these processes are entirely or partially encoded by the plasmid, they also require host housekeeping functions. Plasmids and chromosomes thus interact in ways we do not yet understand, resulting in host-dependent plasmid persistence.

Poor bacteria–plasmid relationships can improve due to evolution of the plasmid1,14,16, the host10,15,16, or both17,18. Plasmid-encoded mutations tend to affect replication, cost, or inheritance14,18–20, whereas those in the host often affect global gene expression and plasmid cost19,21. However, no general mechanism for plasmid–host stabilization has yet emerged. It is also unknown if a host that adapts to one plasmid can be pre-adapted to others, thereby facilitating persistence and spread of a range of multidrug resistance (MDR) plasmids.

To determine evolutionary patterns of plasmid stabilization in bacteria, several combinations of plasmids and hosts should be experimentally evolved and analysed. Here we report on the experimental evolution of an environmental strain, *Pseudomonas* sp. nov. H2, carrying plasmid RP4. Strain H2 poorly maintains this and other plasmids of the incompatibility (Inc) group IncP-1, in contrast to close relatives such as *P. putida*1,11. This makes it an attractive host to determine the molecular mechanisms of bacteria–plasmid evolution. RP4 is a highly promiscuous, MDR prototype IncP-1 plasmid that was found in *P. aeruginosa* isolated from a burn wound22,23. We show that its persistence evolved rapidly and resulted in plasmid addiction, due to chromosomal mutations in genes encoding accessory helicases and the β-subunit of the RNA polymerase, RpoB. More importantly, we demonstrate for the first time that a host adapted to one plasmid can become generally more permissive to plasmid carriage. Understanding how these bacterial genes affect MDR plasmid retention could aid alternative drug therapies needed to combat the spread of antibiotic resistance.

**Results**

An unstable plasmid–host pair stabilized due to host adaptation. To evolve a bacterium–plasmid association experimentally, three replicate populations of *Pseudomonas* sp. nov. H2 containing the promiscuous MDR plasmid RP4 were grown for ca. 600 generations (gen.) by serial batch transfer in tryptic soy broth (TSB) and in the presence of the host- and plasmid-selective antibiotics, rifampicin (Rif) and tetracycline (Tet), respectively. At 100 gen. intervals, plasmid persistence in the absence of Tet was assessed and shown to improve rapidly in all three replicate populations (Fig. 1; Supplementary Fig. 1).

To establish if plasmid persistence increased due to evolution of the plasmid, the host, or both, it was measured for three clones (numbered 1–3) from each replicate population, and for all their possible bacteria–plasmid permutations: the ancestral plasmid (P₀) and plasmids from nine evolved clones (P₁) in both ancestral (H₀) and evolved hosts (H₁). As host adaptation to its environment may affect plasmid persistence, we also determined the persistence of

---

1Department of Biological Sciences, University of Idaho, PO Box 443051, Moscow, ID 83844, USA. 2Institute for Bioinformatics and Evolutionary Studies (IBEST), University of Idaho, PO Box 443051, Moscow, ID 83844, USA. 3Point Loma Nazarene University, San Diego, CA 92106, USA. 4Department of Microbiology, Health Sciences Building, University of Washington, Seattle, WA 98195, USA. 5Department of Biology, University of Florida, PO Box 118525, Gainesville, FL 32611, USA. *e-mail: evatop@uidaho.edu
PA in one clone from each of three evolved plasmid-free populations (populations Ac–Cc, clones Ac1–Cc1, hereafter named ‘control’ populations or clones). The ‘plasmid persistence profiles’, representing the plasmid persistence dynamics over time, were analysed and compared using a recently described segregation and selection (SS) plasmid population dynamics model, combined with linkage cluster analysis18. The difference between two plasmid persistence profiles is reflected in the magnitude of the Bayesian information criterion (BIC) generated by the pairwise comparisons, which permits the complex time series data to be clustered objectively. The plasmid persistence profiles clustered into three distinct groups, A–C (Fig. 2a). The group that contained all the RP4-adapted HE hosts

![Plasmid persistence profiles measured at 100 gen. intervals during the 600 gen. evolution experiment for replicate populations A to C.](image)

**a**–c, Replicate populations A (a), B (b) and C (c). Each data point represents the mean fraction of plasmid-containing (P+) cells (n = 3).

**Fig. 2** | The mutations responsible for increased plasmid persistence are located in the host chromosome. **a**, Based on complete linkage distance analysis of the plasmid persistence profiles, permutations of all the RP4-adapted hosts, HrP_A or HrP_c, clustered together (Group A), whereas all permutations of evolved plasmids in the ancestral host (HrP_a) clustered as Group B, and all ancestral or control hosts with the ancestral plasmids clustered as Group C (HrP_a and HrP_c, respectively). **b**, Plasmid persistence profiles for the individual ancestral, evolved and control clones and their different permutations, grouped (A to C) according to their clustering. Each data point represents the mean fraction of plasmid-containing (P+) cells (n = 3). Height represents the maximum distance between elements of each cluster. The different host-plasmid permutations are: ancestral host–ancestral plasmid (HrP_a); evolved host–evolved plasmid (HrP_c); and evolved host–ancestral plasmid (HrP_a). See also Supplementary Fig. 1 and Supplementary Table 1 for the modelled predictions.
with either $P_a$ or $P_\beta$ (group A) was clearly separated from the group containing the ancestral hosts with evolved plasmids, $H_1P_\varepsilon$ (group B) (difference in the BIC ($\Delta$BIC) score: $-13,134$; the more negative the $\Delta$BIC, the larger the difference between groups). Figure 2b and Supplementary Fig. 1 show this clustering was due to improved persistence of both $P_\alpha$ and $P_\varepsilon$ in $H_1$ (group A). The mean predicted time for the plasmid-bearing fraction to reach 1% ($T_{1%}$) was 34 days for group A and only 18 days for group B (Supplementary Table 1). This improved persistence was not due to adaptation to the environment, as the control clones clustered separately (Fig. 2a, group C) owing to very poor retention of $P_\alpha$, much like $H_1P_\varepsilon$ clones (Fig. 2b; Supplementary Fig. 1; $\Delta$BIC: $-18,439$ and $-1,343$, respectively). These findings indicate that the populations evolved with RP4 improved their ability to retain this plasmid due to plasmid-adaptive chromosomal mutations.

Plasmid stabilization was due to plasmid addiction. To test if the improved persistence of RP4 was due to a change in conjugation rate ($\gamma$), segregation rate ($\lambda$), or plasmid cost ($\sigma$), we estimated model-based parameters and performed conjugation and competition experiments. First, to determine the best fitting model we fitted the plasmid persistence profiles to both the SS and the horizontal transfer (HT) models described previously\cite{23}. As the SS model provided a better fit, conjugation was considered not important for persistence of RP4 (Supplementary Table 2). This was supported by the low empirically measured transfer rate in liquid cultures (Supplementary Fig. 2). Estimates of $\lambda$ and $\sigma$ were obtained by jointly analysing data from plasmid persistence and competition experiments. For unstable bacteria–plasmid pairs, the traditional estimations of plasmid cost through competition assays are confounded by the high plasmid loss rate during the assay. Moreover, that loss rate is often not reported. For highly persistent bacteria–plasmid associations, the plasmid persistence data do not contain enough information about the dynamics, as plasmid-free hosts are rare and they can be confounded by evolutionary changes during the course of the persistence assay. Our joint analysis provides the benefit of more accurate estimation of both parameters while taking into account the growth dynamics in serial batch culture (See Methods and Supplementary Methods). The most likely estimates for $\lambda$ ranged from $6.16 \times 10^{-10}$ to $2.29 \times 10^{-5}$, thus negligibly low. However, $\sigma$ was drastically different between ancestral and evolved clones. While plasmid carriage imposed a fitness cost $\sigma$ of 6.5–7.9% in the $H_1P_\varepsilon$ clones, it conferred a clear fitness benefit of 2.4–8.5% in all $H_1P_\varepsilon$ clones (Fig. 3; Supplementary Table 3). Thus, in line with the first study on so-called addiction of a small non-conjugative plasmid\cite{15,16}, these results show that persistence of a self-transmissible MDR plasmid can improve solely due to chromosomal mutations that turned a plasmid cost into a benefit, resulting in plasmid addiction. So far as we know, no other study has shown the cost of a self-transmissible MDR plasmid to evolve to a benefit via only chromosomal mutations. It provides another explanation for how antibiotic resistance can spread and persist in bacterial populations.

![Fig. 3](image-url) **Fig. 3 | Evolution of plasmid cost into a benefit rather than a change in segregational loss frequency facilitated improved plasmid persistence.** a.b. The effect of the plasmid on host fitness, that is, plasmid cost (a) and segregational loss frequency (b) was jointly estimated using conjugated data from plasmid persistence assays and competition experiments using the SS model for the ancestral host-plasmid pairs $H_1P_\varepsilon$ and their respective evolved $H_1P_\varepsilon$ clones from replicate populations A to C. Evolved clones in each population are ordered sequentially from 1 to 3. For each conjugated dataset, $n = 3$. The vertical lines represent deviations in the model output. The large deviations indicate that a wide range of maximum likelihood estimates are possible for that parameter. The asymmetry of the lines is due to the data being plotted on a log scale.

Adaptation to one plasmid improved general plasmid permissiveness. To determine whether host adaptation to RP4 was plasmid-specific, or allowed for generally improved plasmid permissiveness, we tested the persistence of three naive plasmids in three plasmid-free segregants of RP4-adapted hosts (A2*, B2*, C2*), their corresponding ancestral hosts and three control hosts (Ac1–Cc1). These plasmids were one closely related IncP-1 β plasmid and two plasmids that belong to different incompatibility groups: an IncW plasmid and two plasmids were one closely related IncP-1 β plasmid and two plasmids that belong to different incompatibility groups: an IncW plasmid with a distantly related replication system and an IncQ plasmid with an unrelated replication machinery. Strikingly, all three plasmids showed higher persistence in the RP4-adapted clones than in the ancestral and control strains (Fig. 4; Supplementary Fig. 3; Supplementary Table 4). Thus, by adapting to one particular MDR plasmid, *Pseudomonas* sp. nov. H2 became more permissive towards distantly related antibiotic resistance plasmids. Bacterial communities and populations are known to be diverse in their plasmid permissiveness, but the underlying genetics are not understood\cite{27,28}. These findings also suggest that the plasmid addiction observed for RP4 was not due to the plasmid's Tet resistance (TetR) gene, as previously shown\cite{15}, because the TetR gene is not present on two of the three other plasmids. Addiction of our *Pseudomonas* strain to RP4 must thus be due to other genes shared by these three plasmids.

Genetic solutions towards improved plasmid cost and persistence. To identify mutations that could explain the improved
In the persistence of plasmid RP4, we determined the complete genome sequences of all nine evolved clones, three control and three ancestral clones. Fourteen mutated chromosomal loci were identified among the nine evolved clones after excluding mutations also found in the ancestral and control clones (Fig. 5; Supplementary Table 5). No mutations were found in any of the plasmid genomes, thus confirming that their improved persistence was due to host evolution.

Only two mutations were common among multiple clones from one or two plasmid-containing populations. One was a non-synonymous transversion in a gene encoding a Xpd/Rad3-like helicase protein (64% of the residues modelled at >90% accuracy on the Phyre protein fold recognition portal29). This single nucleotide polymorphism (SNP), hereafter named Xpd/Rad3D672A, was found in three sequenced clones from population A and two from population C.

**Fig. 4** | *Pseudomonas* sp. nov. H2 evolved to be more permissive towards both related and unrelated plasmids. a–c, The plasmid persistence profiles in three clones each of RP4-adapted, ancestral and control hosts are shown for three naive plasmids: the closely related IncP-1β plasmid pB10 (a), the distantly related IncW plasmid Rsa (b) and the unrelated IncQ-1α plasmid RSF1010 (c). They always clustered into two distinct groups: group A contained the three RP4-adapted hosts and group B the ancestral and control hosts, naive to the plasmid. Note that owing to the extremely low persistence of plasmids Rsa and RSF1010, the initial fractions of plasmid-containing (P+) ancestral cells were not close to 1, even though the pre-culture contained plasmid-selective antibiotics. This is owing to survival of plasmid-free cells in the pre-cultures and additional plasmid loss in the colonies during replica-plating. Moreover, the persistence of Rsa in the ancestral host was so poor that no temporal data could be collected to compare with the other Rsa persistence profiles. Each data point represents the mean fraction of P+ cells (n = 3). Height represents the maximum distance between elements of each cluster. See also Supplementary Fig. 3 and Supplementary Table 4 for the modelled predictions.

persistence of plasmid RP4, we determined the complete genome sequences of all nine evolved clones, three control and three ancestral clones. Fourteen mutated chromosomal loci were identified among the nine evolved clones after excluding mutations also found in the ancestral and control clones (Fig. 5; Supplementary Table 5). No mutations were found in any of the plasmid genomes, thus confirming that their improved persistence was due to host evolution.
(Supplementary Table 5). Moreover, the mutation was present in half of the sequenced target-specific amplicons obtained from these populations, but below the detection limit in population B, the ancestral and the control populations (Supplementary Fig. 4). In clones A2 and C3, it was the only mutation after excluding the control SNPs. Finally, based on analysis using SNAP2 (ProteinPredict) the altered residue may have a detrimental effect on protein structure and hence function of the accessory helicase (Supplementary Fig. 5). Together these findings suggest that the mutation was selected by the presence of the plasmid.

The only other common SNP was found in all three clones from population B. It was an A to C transversion 32 base pairs (bp) upstream of an operon encoding an rRNA-binding protein (40% of the residues modelled at >91% accuracy), a DNA repair protein (65% of the residues modelled at 100% accuracy) and an UvrD helicase (94% of the residues modelled at 100% accuracy). This atypical UvrD operon is located 2,007 bp downstream of the Xpd/Rad3-like open reading frame and both are within a low GC island that was probably acquired by HGT from an unknown source (Supplementary Fig. 6).

Fig. 5 SNPs identified in the chromosome of each sequenced clone, as compared with the Rif-sensitive reference strain Pseudomonas sp. nov. H2. White dots represent SNPs. The black line within the innermost green (low) and blue (high) ring indicates GC content. Each set of four coloured rings represents four clones from the following populations: control (red) and the three plasmid-containing populations A–C (orange, green and blue, respectively). Within these are shown, again from inside to outside: the ancestral (light-coloured) and three evolved (dark-coloured) clones. The latter correspond to control clones Ac1, Bc1 and Cc1, and RP4-adapted clones A1–3, B1–3 and C1–3. The outermost black line represents the reference genome. Only contigs with at least 1 SNP in comparison with the reference genome are displayed. Excluded are three small contigs (84, 86 and 132) possibly representing different regions of the same lapA gene and that contain multiple SNPs present in either the ancestral or various evolved strains. However, owing to highly variable sequencing coverage, most SNPs are likely to be sequencing errors. Numbered ticks indicate nucleotide position (×1,000) along the contig and the contig numbers are indicated in bold. When two or more SNPs cannot be resolved, the numbers in parentheses behind the gene or locus names indicate the number of unique variants followed by the number of nucleotides that separate those variants. The radius of contigs 24 and 67 is enlarged by 105%, and nucleotides 10,700–11,000 on contig 67 are zoomed 100-fold to highlight and resolve the helicase and rpoB mutations, respectively. Importantly, there were three rpoB mutations on contig 67: one in all strains including the RifR ancestors, causing Rif resistance, and two in all strains evolved with and without RP4. For details on the mutations and a full description of gene names, see Supplementary Table 5.
In contrast, these same two alleles, especially PuvrD_A-32C, improved for increased plasmid persistence, we separately reconstructed Xpd/Rad3D672A (denoted as XpdEvo or XpdAnc; Fig. 6). To identify mutations that may interact with the accessory helicase mutations, we screened for CDSs that had mutated in both the plasmid-adapted and plasmid-free control host Ac1. Only one such CDS was found and encodes the β-subunit of RNA polymerase (RpoB). Of the 12 evolved strains, 11 contained the same secondary SNP in rpoB, whereas the remaining strain, Cc1, had a different SNP in the same gene (Fig. 5; Supplementary Table 5). As the SNP in RpoB also arose in strains that did not carry RP4, they were probably adaptations to the growth environment. Thus, in control clone Ac1::PuvrD_A-32C, which retained the plasmid almost as well as the corresponding RP4-adapted host, the gene that most probably interacted with PuvrD_A-32C was rpoB.

Discussion

The role of helicases in plasmid stabilization. Strikingly, our and at least three other studies that evolved different host–plasmid pairs now suggest that maladapted interactions between plasmids and host-encoded helicases adversely affect plasmid cost and persistence. Moreover, these interactions can often be improved by single mutations, suggesting we are zooming in on a potential general mechanism of bacterial adaptation to plasmids. First, in P. aeruginosa PA01, loss-of-function mutations in a putative accessory helicase with a UvrD-like helicase C-terminal domain ameliorated the cost of a small non-mobilizable plasmid16. This initial cost was

**Fig. 6 | At least two mutations were required for full plasmid persistence, one plasmid-adaptive and one environment-adaptive mutation.** a, b. Complete linkage cluster analysis (left) of the persistence profiles (shown to the right) of plasmid-adapted clones A2 (A2.H2PA; a) and B2 (B2.H2PA; b), and control clone Ac1 and ancestral clone Anc, each containing either the evolved or ancestral allele of Xpd/Rad3D672A (denoted as XpdEvo or XpdAnc; a) and PuvrD_A-32C (UvrD Evo or UvrD anc; b). See also Supplementary Fig. 7 and Supplementary Tables 6–8 for the modelled predictions. Each data point represents the mean fraction of plasmid-containing cells (n = 3). Height represents the maximum distance between elements of each cluster.

This SNP was present in 30% of the sequenced amplicons from population B and below the detection limit in all other populations, again suggesting it was adaptive to the plasmid (Supplementary Fig. 4). Although no consensus promoter elements could be identified in the vicinity of the SNP, we still refer to it as PuvrD_A-32C. Real-time quantitative polymerase chain reaction (qPCR) transcription analyses on one ancestral and evolved strain, with and without plasmid, show that the presence of the plasmid downregulated transcription of uvrD in the ancestor but not in the evolved strain (Supplementary Fig. 9). This suggests that fine-tuning of UvrD levels was important for plasmid cost amelioration. The remaining SNPs were located in coding DNA sequences (CDSs) involved in metabolism, chemotaxis and response, or upstream of a gene encoding the 5S rRNA. As they were never found in more than two clones from one population, we assumed they were not under strong selection.

To determine whether the accessory helicase SNPs were required for increased plasmid persistence, we separately reconstructed Xpd/Rad3D672A and PuvrD_A-32C in an ancestral and control host. In the ancestral host, neither of the evolved alleles improved persistence of RP4. In contrast, these same two alleles, especially PuvrD_A-32C, improved persistence in control host Ac1, albeit not quite to the same level as the corresponding evolved clones (Fig. 6; Supplementary Fig. 7; Supplementary Tables 6–8). Moreover, the alternate plasmids pBl10, Rsa and RSF1010 were also much more persistent in control host Ac1 with evolved helicase alleles than in host Ac1 with ancestral allele. These results strongly suggest that epistasis between these accessory helicase mutations and at least one other mutation was responsible for general plasmid stabilization.

To identify mutations that may interact with the accessory helicase mutations, we screened for CDSs that had mutated in both the plasmid-adapted and plasmid-free control host Ac1. Only one such CDS was found and encodes the β-subunit of RNA polymerase (RpoB). Of the 12 evolved strains, 11 contained the same secondary SNP in rpoB, whereas the remaining strain, Cc1, had a different SNP in the same gene (Fig. 5; Supplementary Table 5). As the SNP in RpoB also arose in strains that did not carry RP4, they were probably adaptations to the growth environment. Thus, in control clone Ac1::PuvrD_A-32C, which retained the plasmid almost as well as the corresponding RP4-adapted host, the gene that most probably interacted with PuvrD_A-32C was rpoB.
due to the plasmid’s replication initiation (Rep) protein triggering an SOS response in the ancestral host, and the helicase knockout mutation reducing Rep expression. Second, experimental evolution of an IncP-1 mini-replicon in Shewanella oneidensis MR-1 improved plasmid cost and persistence through loss of a helicase (DnaB) binding domain in the plasmid’s Rep protein, reducing the protein’s affinity for DnaB. This probably avoided an SOS response that may explain the high cost of the ancestral plasmid. Third, when we evolved that same plasmid in another host, it stabilized in two clones due to a SNP in either the dnaB promoter or a uvrD gene. Finally, in the present study, SNPs affecting two accessory helicases again compensated for the cost of RP4, and improved the persistence of this and three other broad-host-range plasmids.

Helicases are involved in many aspects of DNA and RNA metabolism, such as replication (replicative helicases), and DNA repair, recombination, translocation, transcription, translation and resolution of replication–transcription conflicts (accessory helicases). Accessory helicases such as Xpd/Rad3 and UvrD generally have variable C- and N-terminal accessory domains that determine their physiological specificity. Interestingly, UvrD and the Xpd/Rad3-like helicase Dng have been shown to be upregulated as part of the SOS response induced by plasmid entrance and replication in a naive host. In Escherichia coli, the UvrD homologue Rep helicase has been shown to interact with DnaB, acting as a second motor that improves replication fork movement on the chromosome. UvrD helicases have also been shown to ‘backtrack’ the RNA polymerase complex to slow down transcription, thus preventing the complex from colliding with the replication fork and causing dsDNA breaks.

In our study, host adaptation to plasmid carriage was facilitated by two different mutations; one that probably changed the Xpd/Rad3-like structure, and one that probably affected UvrD abundance. It was interesting to see that our plasmid decreased uvrD transcript levels in the ancestor where it imposed a high cost, but not in the evolved strains, where it had become beneficial. This suggests that higher UvrD levels are needed for plasmid persistence. The mechanisms by which these mutations affected plasmid cost and persistence are currently not understood but the topic of future work. A simple explanation such as a change in plasmid copy number can probably be excluded based on very similar plasmid sequence coverage for the ancestral and evolved genomes (data not shown). We postulate three not necessarily mutually exclusive models: (1) the accessory helicases interact with DnaB or (2) the plasmid replication initiation protein TrfA to modulate plasmid replication efficiency, or (3) mutations in the helicases ameliorate the fitness cost of plasmid RP4 through their regulatory function.

Whatever the mechanism, further research should confirm that accessory helicases are involved in plasmid persistence across pathogens. The two accessory helicase genes in our Pseudomonas strain were probably acquired by HGT, consistent with previous findings for Pseudomonas PAO1, where it was proposed they caused genetic conflict with the plasmid. We intend to test whether these helicases hamper or improve persistence of various resident MDR plasmids in other strains, as this could aid the development of strategies aimed at slowing down the spread of antibiotic resistance in bacterial pathogens.

Potential epistasis between helicase and RpoB mutations. Stabilization of plasmid RP4 required not only mutations in loci selected in the presence of the plasmid, but also at least one mutation that seemed adaptive to the growth environment. The gene rpoB, which encodes the β-subunit of RNA polymerase, was mutated across all sequenced plasmid-containing and control clones. It is thus the most likely candidate for epistatic interactions with the accessory helicases. The RNA polymerase holoenzyme, consisting of five subunits, α, αII, β, β′ and ω, together with the σ factor, is responsible for transcription. The β-subunit specifically, in addition to DNA binding, is involved in the modulation of transcription through interaction with σ factors and DNA helicases. The rpoB mutation in question was secondary after the initial rpoB mutation that resulted in Rif resistance (RifR) in our ancestral strain. In the absence of the helicase mutations, these rpoB mutations did not improve plasmid persistence at all (Fig. 2). Thus, they most probably affect it not only through epistatic interaction with the helicase mutations.

We propose two possible mechanisms of epistasis between the rpoB and helicase mutations that are not mutually exclusive. First, based on their location, the SNPs in rpoB probably compensated for the cost imposed by the initial RifR rpoB mutation (Table S5) by ameliorating transcription efficiency. It is possible that without this compensatory mutation the helicase mutations were unable to significantly improve plasmid cost and persistence. Second, it is striking that accessory helicases can bind to the RNA polymerase complex, in particular RpoB, to slow down transcription and regulate backtracking. Was there a need for mutations in RpoB to modify this physical interaction? Our secondary RpoB mutations are closer to the active site than to the helicase-binding residues, suggesting they may not affect helicase binding. They are also not close to the rpoB mutations that were shown to rescue viability of strains without accessory helicases. Future studies are needed to determine the mechanism by which the helicase and possibly rpoB mutations can transform a plasmid cost into a benefit for its bacterial host.

In conclusion, to combat the spread and persistence of plasmid-mediated antibiotic resistance, new therapeutic approaches are needed that target mechanisms that affect stable retention of MDR plasmids. To do so, we need to understand which chromosomal gene products stabilize or destabilize MDR plasmids across bacterial species, and how. Our study led to at least three important conclusions that may impact the way we tackle MDR plasmid spread: (1) bacteria can adapt to conjugative MDR plasmids by changing plasmid cost into benefit, resulting in greatly improved plasmid persistence; (2) this can be due to mutations affecting helicases that initially impaired plasmid persistence, a recurring evolutionary pattern that may lead to new antimicrobial therapies; and (3) bacterial adaptation to one plasmid can lead to generally improved plasmid permisiveness, enabling future retention of MDR plasmids. So far as we know, this is the first time that antibiotic exposure has been shown to select for bacterial mutants with increased general permisiveness towards transmissible drug resistance plasmids. These mutations may threaten the efficacy of traditional antibiotic treatments even more than single drug resistance mutations, as adaptation of a pathogen to one plasmid may result in improved retention of other plasmid-encoded antibiotic resistance determinants.

Methods

Bacterial strains, plasmids and media. The bacterial strains and plasmids used in this study are listed in Supplementary Table 9. Unless otherwise stated, the RifR Pseudomonas sp. H2 was cultured in TSB or in tryptic soy agar (TSA) supplemented with Rif (50 µg ml⁻¹) and, when necessary for selection of plasmid RP4, Tet (10 µg ml⁻¹). The various strains of E. coli were cultured in Luria Bertani (LB) broth or on LB agar supplemented with diaminopimelic acid (10 µg ml⁻¹), ampicillin (100 µg ml⁻¹), kanamycin (50 µg ml⁻¹) and Tet (10 µg ml⁻¹) as required. All strains were grown at 30°C.

Plasmids RP4, pB10 and pRSA were introduced into H2 by conjugation from E. coli AT1036, whereas plasmid RSF1010.Km was introduced by mobilization from E. coli S17.1 pir+. Briefly, 1 ml of donor and recipient cultures were harvested by centrifugation and resuspended in 100 µl phosphate buffered saline (PBS, pH 7.4); 50 µl of each cell type was then mixed together and spotted onto a TSA surface and incubated at 30°C. After 1 h, the cell mixture was scraped off the agar surface using a sterile inoculation loop, suspended in 1 ml PBS and a dilution series spread onto the appropriate donor-, recipient- and transconjugant-selective TSA media. The remainder of the donor and recipient cultures were also spread onto transconjugant-selective TSA to verify media selectivity. Evolved RP4 plasmids were introduced into the ancestral H2 host by electroporation using standard methods.

Molecular methods. Conventional plasmid isolation and DNA manipulation techniques were used as described in ref. 41. Sanger sequencing was sourced to Eim.
Experimental evolution. Triplicate populations of the RifR strain *Pseudomonas* sp. H2 (RP4) were evolved for ca. 600 gen. in serial batch cultures in the presence of Tet to select for the plasmid and Rif to avoid contamination. In addition, triplicate plasmid-free populations were also evolved in the presence of only Rif. Each population was founded from a single colony that was inoculated into 5 ml TSB. During the first 25 days, 4.9 µl of culture was transferred to culture plates containing 30% glycerol to yield roughly 10³ gen. of cells, which was then resuspended in fresh medium to yield roughly 10⁴ gen. of cells. Each of the populations was maintained by serial dilution and spread onto TSA, and single colonies representing three unique clones from each population were randomly chosen every 100 gen. as follows: 4.9 µl of each culture was transferred into 5 ml TSB supplemented with Rif and maintained by serial dilution. Experimental evolution.

Comparison of plasmid persistence. Plasmid persistence in the evolving plasmid-containing populations was assayed every 100 gen. as follows: 4.9 µl of each culture was transferred into 5 ml TSB supplemented with Rif and maintained by serial batch transfer for 10 days. The growth conditions were the same as in the evolution experiment but without Tet selection for the plasmid. On a daily basis, the fraction of Rif-resistant *E. coli* cells was determined as previously described by spreading the TSA supplemented with Rif onto plates containing both Rif and Tet to differentiate between plasmid-containing and plasmid-free cells. Plasmid persistence for each of the nine evolved clones was measured as described above, but starting from their respective archived glycerol stocks.

Additionally, one of the first plasmid-free segregants that were obtained for each clone was archived at −70°C, similar to the original plasmid-containing clone. Persistence of the ancestral plasmid as well as other naïve plasmids in these evolved plasmid-free segregants (denoted by *) and in clones from the plasmid-free evolved populations was assayed in the same way after the plasmids were reintroduced into the respective hosts by means of conjugation or mobilization.

Plasmid persistence. The plasmid persistence profiles, defined as the time series obtained from the plasmid persistence assays, were compared and analyzed using the Hysplit software available online (http://www.ready.noaa.gov/Hysplit4/index.php). The two models, defined as the SS or HT model, describe the plasmid population dynamics models as described and implemented by us previously. The models employed the same parameters and were compared to find the best fitting model based on the lowest ΔBIC score, the more different the profiles (see ref. 18 for more detailed information). To visualize the differences in plasmid persistence profiles, a matrix of the ΔBIC scores was used to construct a complete linkage cluster dendrogram. Estimation of plasmid cost and segregational loss frequencies. Plasmid cost and segregational loss frequencies were estimated by jointly analysing data generated from competition and persistence experiments using the SS model. Competition experiments were done by mixing plasmid-carrying strains and their isogenic plasmid-free counterparts at a 1:1 ratio, and growing them under conditions resembling the plasmid persistence assays. Specifically, the cultures were inoculated from glycerol stock into 5 ml TSB containing the appropriate antibiotics and grown for 16 h. Thereafter, 1.5 ml of each culture was harvested by centrifugation to remove the spent media and antibiotic, and the DNA was resuspended in 1.5 ml PBS and its optical density measured at 600 nm (OD₆₀₀). The appropriate volumes that represented a similar cell count for each competitor were then mixed together in a total volume of 200 µl. Of each mixture, 4.9 µl was inoculated in triplicate into 5 ml TSB supplemented with Rif and grown for 24 h, after which the cultures were again diluted 1/10000-1/1000, grown for another 24 h, and dilution series of the original mixed culture and after each 24 h time period were spread onto TSA supplemented with Rif. From here, the fractions of plasmid-containing and -free cells within each population were determined by scoring the absence or presence of growth for 52 colonies replicated on TSA supplemented with Rif and with or without Tet. To be able to compare the information from the competition assay together with the information from the persistence assays, to better estimate the plasmid cost, the joint likelihood function coming from these two experiments was maximized, using our R package 'StabilityToolkit'. In an online appendix that is publicly available, we provide extensive instructions to download and install this package, and perform these analyses (see Code availability).

We note that although the SS and HT models are useful to jointly estimate cost, segregation and conjugation, they scale the cost parameter σ slightly differently than the traditional cost σ (refs 39,40). Despite the different scaling, in the Supplementary Methods we show that both formulations concur in what constitutes a plasmid cost and a plasmid benefit (positive and negative values of the cost parameters). A one-to-one transformation between σ and s is given by: s = σ (1 + s) (see Supplementary Information).

Conjugation assays to determine the conjugation rate constant. Conjugation between plasmid-containing donors (D) and plasmid-free recipients (R) was measured under growth conditions similar to those used for measuring plasmid persistence, with one difference. To distinguish D and R, the plasmid-free recipient strain was isogenic to the plasmid-free ancestor H2, but resistant to Rif. In addition, rather than Rif, SS and R cultures were grown for 16 h, their densities adjusted based on their OD₆₀₀ and 4.9 µl of each culture was inoculated together into tubes containing 5 ml TSB. To quantify the total initial (N₀) cell densities, a sub-sample of the mixture was immediately (t₀) diluted and spread onto D-, R- and transconjugant (T)-selective TSA and incubated at 30°C for 24 h. The cultures were incubated at 30°C with shaking for 24 h (tₜ) before being diluted and spread again onto the appropriate agar media to quantify total (Nₜ), D- and T cell densities (D, R and T). The cell densities were then calculated using the transfer rate constant γ (in ml (c.f.u. × t⁻¹)) as follows and as described by ref. 40:

$$\gamma = \frac{\ln(OD_{D}/OD_{R})}{t_t - t_0}$$

where γ is the maximum growth rate for the combined donor and recipient cultures, which is calculated from two optical density (OD₆₀₀) determinations, say a and b, taken during the exponential phase, as follows (see ref. 40):

$$\gamma = \frac{t_a - t_b}{\ln(a/b)}$$

DNA sequencing and analysis. Genomic DNA (gDNA) was extracted from populations using a GenElute Bacterial Genomic DNA Kit (Sigma-Aldrich) according to the manufacturer’s instructions. The quality and quantity of the gDNA was assessed on a 1% agarose gel and the concentrations were determined fluorometrically using Quant-iT PicoGreen dsDNA Reagent (ThermoFisher Scientific) on a SpectraMax Paradigm Multi-Mode Microtiter Plate Reader (Molecular Devices). To identify mutations in evolved clones compared with ancestral clones, the gDNA was sequenced using a whole-genome shotgun approach with paired 300-bp reads generated on MiSeq (Illumina) at the iBEST Genomics Resources Core (GRC) at the University of Idaho (USA). Sequencing adapters and low-quality bases were trimmed using custom scripts, after which the reads were mapped to the *Pseudomonas* sp. nov. H2 and RP4 genomes available in NCBI under accession numbers JRP000000000 and L277381, respectively, using BivSeq version 0.26.1 (ref. 18). The genome and sequencing data were visualized using Circos version 0.69. To determine SNP frequency within heterogeneous populations, amplicon libraries of the target regions were constructed using a dual barcoded two-step PCR amplification protocol for sequencing on an Illumina MiSeq sequencing platform. Sequencing adapters and low-quality bases were trimmed using custom scripts, containing universal CS1 and CS2 adapter sequences (Supplementary Table 10). The PCR reactions consisted of PCR buffer (New England Biolabs), 3 mM MgCl₂ (ThermoFisher Scientific), 0.24 mg ml⁻¹ bovine serum albumin (BSA; ThermoFisher Scientific), 200 µM deoxynucleotide triphosphates (dNTPs; ThermoFisher Scientific), 50 nM each of the appropriate forward and reverse common-sequence-tagged primers, 0.025 µg ml⁻¹ Taq DNA polymerase (New England Biolabs) and 2 nM g²⁻⁻⁻⁻²⁻⁻⁻ DNA in a final volume of 50 µl. The BioRad thermocycler parameters were as follows: 95°C for 2 min, 10 cycles of 95°C for 30 sec, 60°C for 30 sec and 68°C for 1 min, and finally 68°C for 10 min. The resulting PCR products were assessed on a 1% agarose gel, whereafter it was diluted 15-fold and used as template in a second PCR reaction to attach the appropriate barcodes. These PCR reactions consisted of PCR buffer (New England Biolabs), 4.5 mM MgCl₂ (ThermoFisher Scientific), 0.24 mg ml⁻¹ BSA (ThermoFisher Scientific), 200 µM dNTPs (ThermoFisher Scientific), 75 mM barcoded primer, 0.05 µM µ⁻¹ Taq DNA polymerase (New England Biolabs) and 1.300 PCR product 1 in a final volume of 50 µl. The BioRad thermocycler parameters were as follows: 95°C for 1 min, 10 cycles of 95°C for 30 sec, 60°C for 30 sec and 68°C for 1 min, and finally 68°C for 5 min. The quality and yield of the final amplicon libraries were assessed on an Agilent Bioanalyzer (Agilent Technologies) using the Agilent DNA 1000 Kit per manufacturer’s instructions. The resulting libraries were sequenced using an Illumina MiSeq sequencing platform using MiSeq Reagent kit v3 at the iBEST Genomics Resources Core. The same sequencing adapters and quality base were used as in the gDNA sequencing, while the unique primers were normalized using custom scripts and the SNP frequency determined using the Geneious R8 software package.
Gene expression analysis. Stationary phase bacterial cultures were diluted to an OD$_{600}$ of approximately 0.005 in fresh Terrific Broth, supplemented with Rif and Tet as required, and grown at 30 °C. When the cultures reached an optical density of approximately 0.5 and 1.0, 1 ml of cells were combined with 200 μl stop solution (95% ethanol, 5% acid phenol chloroform). The cells were then harvested by centrifugation at 4 °C (13,000 r.p.m., 2 min) and stored at −20 °C until RNA extraction. RNA was extracted using GeneJET RNA Purification Kit (ThermoFisher Scientific) according to manufacturer's instructions. Residual gDNA contamination was removed by DNAse I (ThermoFisher) treatment at 37 °C for 1 h. The amount of RNA was quantitated using a NanoDrop (NanoDrop2000). For complementary DNA (cDNA) construction, 1 μg total RNA was added to the iScript Reverse Transcription SuperMix Kit (BIO-RAD) following manufacturer's instructions. Real-time qPCR was performed on an 150 ng cDNA for $uvrD$ and 3 ng for 165 RNA, using the iQ™ Universal SYBR Green Supermix (BIO-RAD) and Bio-Rad CFX Connect qPCR machine. The $uvrD$ and 165 RNA cDNA concentrations were quantified using the UvrD-helicase-P2v:Rev 16 and 165 RNA primer sets. The samples were normalized using the 165 RNA levels and the $uvrD$ gene expression relative to the plasmid-free ancestor was determined using the ΔΔCt method. The statistical significance of differences in gene expression relative to the ancestor was determined using Tukey’s multiple comparisons test.

Allotypic exchange. Allotypic exchange experiments were performed based on a modified protocol from ref. 21. Briefly, the ancestral and evolved alleles were cloned into pPS04 and introduced into H2 by mobilization from E. coli Zhang (Biol R0417). Recombinants were selected on LB agar supplemented with kanamycin (50 μg ml$^{-1}$). For each allele, 46 merodiploid recombinants were replicated onto LB agar supplemented with kanamycin (50 μg ml$^{-1}$). Insertion of the vector into the H2 genome was confirmed by PCR using primers specific to the pPS04 $apf-1$ gene and by verifying vector absence using a plasmid miniprep kit (ThermoScientific, Waltham, MA). Thereafter, a single clone for each allele was inoculated into 5 ml of no salt LB (NSLB) supplemented with 15% (w/v) sucrose and incubated for ca. 16 h at 37 °C to select and enrich for double cross-over mutants. The stationary phase cultures were diluted and spread onto NSLB agar supplemented with 15% (w/v) sucrose, and ca. 1,000 colonies were screened for the absence of the $apf-1$ gene by replication onto NSLB agar supplemented with 15% (w/v) sucrose and with or without kanamycin. Absence of the pPS04 $apf-1$ gene was also confirmed by PCR. To verify the correct genotype for each allele, PCR products were generated in each case from 24 clones using the H2 $Xpd$/Rad3_Seq and H2 $UvrD$_Seq primer pairs, and sequenced by means of Sanger sequencing in both the forward and reverse directions (Elii Biopharm). In each case, a clone that retained the original allele was also saved and used to control for the effect of allotypic exchange on the host.

Data availability. All sequencing data are available under NCBI BioProject accession number PRJNA261945. All persistence and competition assay data are available from the Dryad Digital Repository: http://dx.doi.org/10.5066/dryad.v2svk.

Code availability. The StabilityToolkit package and instructions are available at https://github.com/mponciano/StabilityToolkit/blob/master/RunningStabToolsPack.zip.

Received: 2 November 2016; Accepted: 16 June 2017; Published online: 7 August 2017

References
1. Hogarten, J. P. & Townsend, J. P. Horizontal gene transfer, genome innovation and evolution. Nat. Rev. Microbiol. 3, 679–687 (2005).
2. Schwarz, S. & Johnson, A. P. Transferable resistance to colistin: a new but old threat. Nat. Rev. Microbiol. 12, 1413–1426 (2010).
3. De Gelder, L., Williams, J. J., Ponciano, J. M., Sota, M. & Top, E. M. Adaptive plasmid evolution results in host-range expansion of a broad-host-range plasmid. Genetica 178, 2179–2190 (2008).
4. Sota, M. et al. Shifts in the host range of a promiscuous plasmid through parallel evolution of its replication initiation system. ISME J. 4, 1568–1580 (2010).
5. Bouma, J. E. & Lenski, R. E. Evolution of a bacteria/plasmid association. Nature 335, 351–352 (1988).
6. Shintani, M. et al. Response of the Pseudomonas host chromosomal transcriptome to carriage of the IncP-7 plasmid pCAR1. Environ. Microbiol. 12, 1413–1426 (2010).
7. De Gelder, L., Williams, J. J., Ponciano, J. M., Sota, M. & Top, E. M. Adaptive plasmid evolution results in host-range expansion of a broad-host-range plasmid. Genetica 178, 2179–2190 (2008).
8. Loftie-Eaton, W. et al. Evolutionary paths that expand plasmid host-range: implications for spread of antibiotic resistance. Mol. Biol. Evol. 33, 885–897 (2016).
9. Maestro, B., Sanz, J. M., Diaz-Orejas, R. & Fernandez-Trehuagueros, E. Modulation of p5S01 host range by plasmid-encoded RepA initiator protein. J. Bacteriol. 185, 1367–1375 (2003).
10. Yano, H. et al. Evolved plasmid–host interactions reduce plasmid interference competition. Mol. Biol. Evol. 10, 751–766 (2016).
11. Heuer, H., Fox, R. E. & Top, E. M. Frequent conjunctive transfer accelerates adaptation of a broad-host-range plasmid to an unfavorable Pseudomonas putida host. FEMS Microbiol. Ecol. 59, 738–748 (2007).
12. Pansegrau, W. et al. Complete nucleotide sequence of Birmingham IncPox plasmid compilation and analysis. J. Mol. Biol. 239, 626–663 (1994).
13. Saunders, J. R. & Grinsted, J. Properties of RP4, an R factor which originated in Pseudomonas aeruginosa SB. J. Bacteriol. 112, 690–696 (1972).
14. Baquero, F., Coque, T. M. & de la Cruz, F. Ecology and evolution as targets: the need for novel eco-eco-drugs and strategies to fight antibiotic resistance. Antimicrob. Agents Chemother. 55, 3649–3660 (2001).
15. Ponciano, J. M., De Gelder, L., Top, E. M. & Joyce, P. The population biology of bacterial plasmids: a hidden Markov model approach. Genetica 175, 957–967 (2007).
16. Lenski, R. E., Simpson, S. C. & Nguyen, T. T. Genetic analysis of a plasmid-encoded, host genotype-specific enhancement of bacterial fitness. J. Bacteriol. 176, 3140–3147 (1994).
17. Heuer, H., Ebers, J., Weinert, N. & Smalla, K. Variance in permissiveness for broad-host-range plasmids among genetically indistinguishable isolates of Dickeya spp. from a small field plot. FEMS Microbiol. Ecol. 73, 199–206 (2010).
18. Klümper, U. et al. Broad host range plasmids can invade an unexpectedly diverse fraction of a soil bacterial community. ISME J. 9, 934–945 (2014).
19. Kelley, A. L., Mezulis, S., Tayes, M. C., Wass, N. M. & Sternberg, M. J. E. The Phyre2 web portal for protein modeling, prediction and analysis. Nat. Protoc. 10, 845–858 (2015).
20. Yachdav, G. et al. PredictProtein - an open resource for online prediction of protein structural and functional features. Nucleic Acids Res. 42, 337–343 (2014).
21. Byrd, A. K. & Raney, K. D. Superfamily 2 helicases. Front. Biosci. 17, 2070–2088 (2013).
22. Merrinck, C. N., Brewer, B. J. & Merrinck, H. The B. subtilis accessory helicase PcrA facilitates DNA replication through transcription units. PLoS Genet. 11, e1005289 (2015).
23. Fairfax-Williams, M. E., Guenther, U. & Jankowsky, E. SF1 and SF2 helicases: family matters. Nat. Rev. Microbiol. 13, 266–274 (2015).
24. Boubakri, H., de Septenville, A. L., Viguera, E. & Michel, B. The helicases UvrD, Rep and UvrD cooperate to promote replication across transcriptional units in vivo. EMBO J. 29, 145–157 (2010).
25. Ingmer, H., Miller, C. & Cohen, S. N. The RepA protein of plasmid pSC101 controls Escherichia coli cell division through the SOS response. Mol. Microbiol. 42, 519–526 (2001).
26. Reichard, H., de Septenville, A. L., Viguera, E. & Michel, B. The helicases DnR, Rep and UvrD cooperate to promote replication across transcriptional units in vivo. EMBO J. 29, 145–157 (2010).
27. Boubakri, H., de Septenville, A. L., Viguera, E. & Michel, B. The helicases DnR, Rep and UvrD cooperate to promote replication across transcriptional units in vivo. EMBO J. 29, 145–157 (2010).
28. Epstein, V. et al. UvrD facilitates DNA repair by pulling RNA polymerase backwaters. Nature 505, 372–377 (2014).
29. Sun, M. X et al. Rep provides a second motor at the replisome to promote duplication of protein-bound DNA. Mol. Cell 36, 654–666 (2009).
30. Rokhovu, S. & Nudler, E. RNA polymerase holoenzyme: structure, function and biological implications. Curr. Opin. Microbiol. 6, 93–100 (2003).
41. Kuznedelov, K. et al. A role for interaction of the RNA polymerase flap domain with σ subunit in promoter recognition. Science 295, 855–857 (2002).
42. Brandis, G., Wrande, M., Liljas, L. & Hughes, D. Fitness-compensatory mutations in rifampicin-resistant RNA polymerase. Mol. Microbiol. 85, 142–151 (2012).
43. Reynolds, M. G. Compensatory evolution in rifampin-resistant Escherichia coli. Genetics 156, 1471–1481 (2000).
44. Sambrook, J. & Russell, D. W. Molecular Cloning: A Laboratory Manual (Cold Spring Harbour Laboratory Press, New York, 2001).
45. Joyce, P. et al. Modeling the impact of periodic bottlenecks, unidirectional mutation, and observational error in experimental evolution. J. Math. Biol. 50, 645–662 (2005).
46. Simonsen, L., Gordon, D. M., Stewart, F. M. & Levin, B. R. Estimating the rate of plasmid transfer: an end-point method. J. Gen. Microbiol. 136, 2319–2323 (1990).
47. Deatherage, D. E. & Barrick, J. E. Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using bresq. Methods Mol. Biol. 1151, 165–188 (2014).
48. Krzywinski, M. I. et al. Circos: an information aesthetic for comparative genomics. Genome Res. 19, 1639–1645 (2009).
49. Hinello, I. R. et al. Precision-engineering the Pseudomonas aeruginosa genome with two-step allelic exchange. Nat. Protoc. 10, 1820–1841 (2015).

Acknowledgements
This work was supported by the National Institute of Allergy and Infectious Diseases grant R01 AI084918 of the National Institutes of Health (NIH). The genome resequencing was done by the IBEST Genomics Research Core and made possible thanks to an Institutional Development Award (IDeA) from the National Institute of General Medical Sciences (NIGMS) of the NIH under grant number P30 GM103324. K.B. was in part supported by a University of Idaho Department of Biological Sciences undergraduate research grant and by an NIGMS INBRE award, grant number P20 GM103408. H.Q. was supported by a National Science Foundation REU Site award, 1460696. We thank the laboratory of C. Marx for providing us with vector pPS04 and we thank H. Merrikh for useful suggestions.

Author contributions
W.L. and E.M.T conceived the project and wrote the manuscript. W.L., K.B., H.Q., K.D., M.K.T. and J.M. performed the experiments. W.L. and J.M.P. performed the statistical analyses. W.L. and S.H. performed the bioinformatic analysis. H.M. facilitated part of the work and helped with data interpretation.

Competing interests
The authors declare no competing financial interests.

Additional information
Supplementary information is available for this paper at doi:10.1038/s41559-017-0243-2.
Reprints and permissions information is available at www.nature.com/reprints.
Correspondence and requests for materials should be addressed to E.M.T.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.