The Role of Clonal Interference in the Evolutionary Dynamics of Plasmid-Host Adaptation

Julie M. Hughes,[a,b] Brian K. Lohman,[a,b] Gail E. Deckert,[a,b] Eric P. Nichols,[c] Matt Settles,[b] Zaid Abdo,[d,b] and Eva M. Top[a,b]

Department of Biological Sciences, University of Idaho, Moscow, Idaho, USA[a]; Institute for Bioinformatics and Evolutionary Studies, University of Idaho, Moscow, Idaho, USA[b]; Department of Computer Science, Indiana University, Bloomington, Indiana, USA[c]; and Department of Mathematics and Statistics, University of Idaho, Moscow, Idaho, USA[d]

ABSTRACT **Promiscuous plasmids replicate in a wide range of bacteria and therefore play a key role in the dissemination of various host-beneficial traits, including antibiotic resistance. Despite the medical relevance, little is known about the evolutionary dynamics through which drug resistance plasmids adapt to new hosts and thereby persist in the absence of antibiotics. We previously showed that the incompatibility group P-1 (IncP-1) minireplicon pMS0506 drastically improved its stability in novel host *Shewanella oneidensis* MR-1 after 1,000 generations under antibiotic selection for the plasmid. The only mutations found were those affecting the N terminus of the plasmid replication initiation protein TrfA1. Our aim in this study was to gain insight into the dynamics of plasmid evolution. Changes in stability and genotype frequencies of pMS0506 were monitored in evolving populations of MR-1 (pMS0506). Genotypes were determined by sequencing *trfA1* amplicons from individual clones and by 454 pyrosequencing of whole plasmids from entire populations. Stability of pMS0506 drastically improved by generation 200. Many evolved plasmid genotypes with point mutations as well as in-frame and frameshift deletions and duplications in *trfA1* were observed in all lineages with both sequencing methods. Strikingly, multiple genotypes were simultaneously present at high frequencies (>10%) in each population. Their relative abundances changed over time, but after 1,000 generations only one or two genotypes dominated the populations. This suggests that hosts with different plasmid genotypes were competing with each other, thus affecting the evolutionary trajectory. Plasmids can thus rapidly improve their stability, and clonal interference plays a significant role in plasmid-host adaptation dynamics.**

IMPORTANCE **Promiscuous plasmids play an important role in the spread of antibiotic resistance and many other traits between closely and distantly related bacteria. However, little is known about the dynamics by which these broad-host-range antibiotic resistance plasmids adapt to novel bacteria and thereby become more persistent, even in the absence of antibiotics. In this study, we show that after no more than 200 generations of growth in the presence of antibiotics, a plasmid that was initially poorly maintained in a novel bacterial host evolved to become drastically more persistent in the absence of antibiotics. In each of the evolving populations, an unexpectedly large number of bacterial variants arose with distinct mutations in the plasmid's replication initiation protein. Our results suggest that clonal interference, characterized by competition between variant clones in a population, plays a major role in the evolution of the persistence of drug resistance.**

Received 14 March 2012 **Accepted** 30 May 2012 **Published** 3 July 2012

Citation Hughes JM, et al. 2012. The role of clonal interference in the evolutionary dynamics of plasmid-host adaptation. mBio 3(4):e00077-12. doi:10.1128/mBio.00077-12.

Editor Nancy Moran, Yale University

Address correspondence to Eva M. Top, evatop@uidaho.edu.

Horizontal gene transfer plays an important role in bacterial adaptation and evolution, and plasmids are among the major agents of gene exchange (1–3). Plasmids are extrachromosomal DNA molecules that frequently code for traits such as antibiotic resistance, virulence factors, and the ability to degrade organic compounds. Some of these plasmids, especially those of the incompatibility group P-1 (IncP-1), are able to transfer to and replicate in a wide range of phylogenetically distinct bacteria and are termed broad-host-range (BHR) or promiscuous plasmids (4, 5). They act as gene shuttles for a variety of functional traits, which they can spread to diverse bacteria, including pathogens (6).

While BHR plasmids replicate in a wide range of bacteria, they are not always well adapted to new hosts. Some plasmids do not segregate perfectly between daughter cells upon cell division, and most confer a cost to their hosts. Therefore, the proportion of plasmid-bearing cells in a bacterial population can slowly or rapidly decline in the absence of selection for the plasmid (7, 8). This is referred to as poor plasmid stability or poor persistence. Several studies have shown adaptation of plasmids to their hosts through cost amelioration (9–14), but only a few have used plasmid stability improvement as a measure of plasmid-host adaptation (8–10). Only in some of these studies have the mutations responsible for plasmid-host adaptation been identified (8, 12, 15). Despite the obvious importance to human health, the rate at which plasmids evolve to higher stability in novel hosts and the underlying evolutionary dynamics have not yet been assessed.

While much research has focused on the evolution of bacteria rather than plasmids, very few studies have been able to monitor

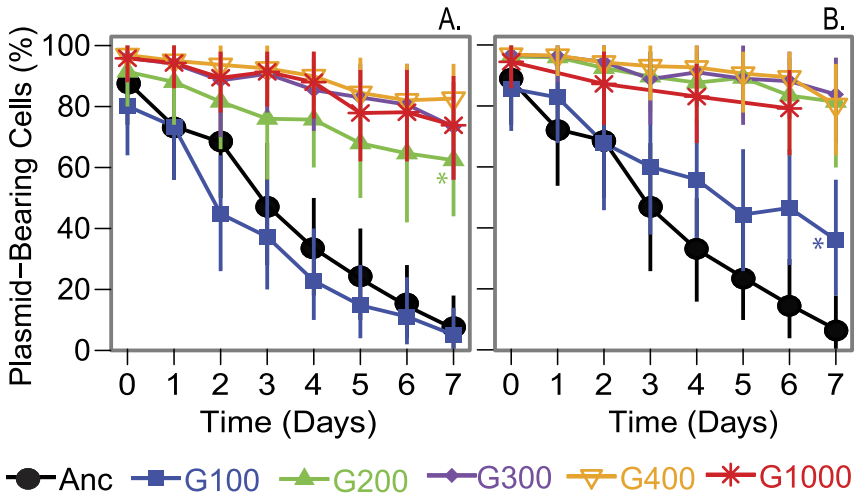

FIG 1 Stability of evolving pMS0506 in coevolving (A) and ancestral (B) host populations from experiment A lineage 5. Anc, ancestral populations; G100 to G400 and G1000, populations that have undergone 100 to 400 and 1,000 generations of growth, respectively. Data represent the posterior means, with the credibility intervals at each of the observed time points. The first time point by which stability was significantly improved is indicated by an asterisk (based on logistic regression). Parallel results for all other lineages of experiments A and B are shown in the supplemental material (see Fig. S1 and S2).

the dynamics of bacterial evolution at the genotype level (16–19). In these studies, only some of the mutations present at high frequencies within the evolving populations were analyzed. Additionally, due to the limitations of current software, only single nucleotide polymorphisms (SNPs) and small insertions and deletions (indels of <3 bp) were detected. Due to their much smaller genomes, phage have been a more amenable model system for experimental evolution studies aimed at determining the molecular mechanisms and evolutionary dynamics of adaptation (20–23). So far, no study has examined the dynamics of plasmid-host evolution at the molecular level. Although it has been typically expected that beneficial mutations will become fixed in a population through periodic selective sweeps, a growing body of literature suggests that a process called "clonal interference" may explain at least some of the dynamics of microbial evolution (20, 24–28). Clonal interference occurs when different genotypes with unique beneficial mutations are present concurrently (16) in a given asexual population and therefore compete with one another. Such competition is expected to result in an increase in the time to fixation and a greater fitness effect of those mutations that eventually become fixed. While recent studies suggest that clonal interference may be an important process in bacterial and phage evolutionary dynamics (20, 24–26, 28), there is still relatively little direct experimental evidence of its occurrence at the genotype level (16, 20, 25, 27). So far, it is not known whether clonal interference plays a role in the dynamics of plasmid evolution.

In a previous study, we found that the nonconjugative BHR IncP-1 minireplicon pMS0506 had drastically improved its stability in novel host *Shewanella oneidensis* MR-1 after 1,000 generations of experimental evolution under selection for the plasmid (kanamycin). This phenotypic change was due to single mutations (point mutations, deletions, or duplications) in the 5′ region of the plasmid-encoded *trfA1* gene, which codes for the replication initiation protein TrfA1 (8). Like most other IncP-1 plasmids, our model plasmid encodes long and short replication initiation proteins, TrfA1 and TrfA2. Since TrfA2 is produced from an internal translational start site within the *trfA1* gene sequence, none of the 5′-end *trfA1* mutations affected TrfA2. In the present study, our goal was to gain insight into the dynamics that underlie the evolution of plasmid stability in a novel host. To do this, we monitored the plasmid stability and genetic changes that occurred in plasmid pMS0506 during evolution of strain MR-1 (pMS0506). We found that plasmid stability significantly improved within as few as 200 cell doublings and that a high diversity of evolved genotypes with mutations in the 5′ region of *trfA1* arose and competed in the evolving populations. These results strongly suggest that clonal interference plays a role in the dynamics of plasmid-host adaptation.

RESULTS

Improvement in plasmid stability occurred rapidly in evolving populations. To determine the dynamics of plasmid adaptation to a new host in an environment that selects for plasmid maintenance, plasmid stability was examined every 100 generations during evolution of 10 replicate lineages of *S. oneidensis* MR-1 carrying plasmid pMS0506. These populations were obtained from two independent experimental evolution studies (evolution experiments A [evolution over 1,000 generations] and B [repeated evolution experiment for 400 generations]). First, plasmid stability was tested in evolving populations at generations 0, 100, 200, 300, 400, and 1,000 (experiment A) and at generations 0 through 400 (experiment B). To determine if the observed improvement in stability was due to plasmid evolution, host evolution, or coevolution, the stability of plasmids from the evolving populations of evolution experiment A was also tested in the ancestral host. A typical example of the rapid improvement in plasmid stability that was observed in all lineages, both in coevolved and ancestral hosts, is shown in Fig. 1. Results for all lineages are presented in the supplemental material (see Fig. S1 and S2). For each stability curve, logistic regression was used to fit the data, and the rate of decline in the plasmid-bearing fraction was calculated at L50, the time when 50% of the population had lost its plasmids (data not shown). This showed that the stability of evolved plasmids in both ancestral and coevolved hosts was significantly improved after 200, and sometimes even 100, generations. The stability patterns of the evolving plasmids were similar between the two independent evolution experiments (see Fig. S1A and S2 in the supplemental material). Thus, within as few as 200 cell doublings, a bacterial population with an unstable drug resistance plasmid can evolve improved plasmid persistence.

Evolutionary dynamics of *trfA* genotype frequencies during plasmid adaptation. With phenotypic evidence for the rapid improvement of plasmid persistence, we then determined which plasmid genotypes rose to dominancy in the evolving populations and what were the evolutionary dynamics of this plasmid adaptation process. In our previous study (8), one randomly chosen evolved plasmid variant from each of the five lineages of experiment A had been completely sequenced after 1,000 generations.

These five plasmids each contained only one mutation, always in the 5′ region of the gene *trfA1*, which codes for the plasmid replication initiation protein. Here, two different approaches were used to monitor the evolutionary dynamics of plasmid adaptation at the molecular level. First, in a colony PCR-based approach, the first 693 bp of the 5′ end of gene *trfA1* were amplified and sequenced from at least 10 clones, randomly chosen from each evolving population sampled at regular time intervals (generations 0 to 400, 600, 800, 1,000). Second, Roche 454 Titanium pyrosequencing was used to sample deeper into the plasmid populations from lineage A5 every 100 generations and to determine if there were other polymorphisms outside this *trfA1* region.

Two questions were addressed using the colony PCR-based approach: (i) how many evolved *trfA1* genotypes arose and (ii) what were the dynamics of genotype frequencies in each evolving population over time. Between evolution experiments A and B, no less than 40 individual bacterial clones with a unique *trfA1* mutation were identified. The detection of identical as well as different genotypes in independently evolved lineages from repeated evolution experiments (see Fig. 2 inset and Table S1 in the supplemental material) indicates some degree of parallel evolution at the nucleotide level. As shown in Fig. 2, the genetic changes consisted of point mutations (6/40), deletions (22/40), and duplications (12/40). All mutations were within the first 340 bp of *trfA1*, leaving the shorter *trfA2* gene intact. The majority of the 22 deletions and all duplications observed were flanked by GC-rich direct repeats of 3 to 10 bp. These results are consistent with previous sequence data obtained at generation 1,000 (8), but a much larger diversity of mutants was obtained here due to deeper sampling at more time points. The results show that plasmid adaptation to host MR-1 occurred by parallel evolution in the *trfA1* gene, likely because of strong antibiotic selection on a large variety of mutations within the same gene region.

We also monitored the changes in relative abundance of *trfA1* genotypes over time (Fig. 3A and B for evolution experiments A and B, respectively). After just 100 generations, only very few evolved genotypes were observed. However, the majority of bacterial clones sampled from generation 200 onward contained a plasmid with an evolved *trfA1* genotype, with an average of 4.5 types per population. Note that dominance of these new genotypes at generation 200 in the 10 populations coincided with the sudden increase in plasmid stability between generations 100 and 200 (compare Fig. 3A and B with Fig. 1 and Fig. S1 and S2 in the supplemental material). Together with our previous findings on the role of the *trfA* mutations (8), this strongly suggests that these mutations were responsible for the stability improvement.

Instead of a series of selective sweeps, multiple *trfA1* genotypes coexisted at various time points at high frequencies. Since only 10 to 15 clones were sampled per population, the detection of a given genotype suggests that it had become dominant. Multiple genotypes thus simultaneously rose to high frequencies within each evolving population. At generation 400, a high diversity of genotypes remained in most lineages even though at the phenotype level, no statistical differences in plasmid stability were seen after generation 200 (data not shown). Strikingly, the frameshift genotype 7 was present in 9 of 10 lineages, often at high frequencies, but never persisted up to generation 1,000. After 1,000 generations, 4 of the 5 populations were dominated by genotype 8, and 1 was dominated by genotype 14 (Fig. 2 and 3A). Interestingly, all genotypes observed after 1,000 generations were in-frame mutations,

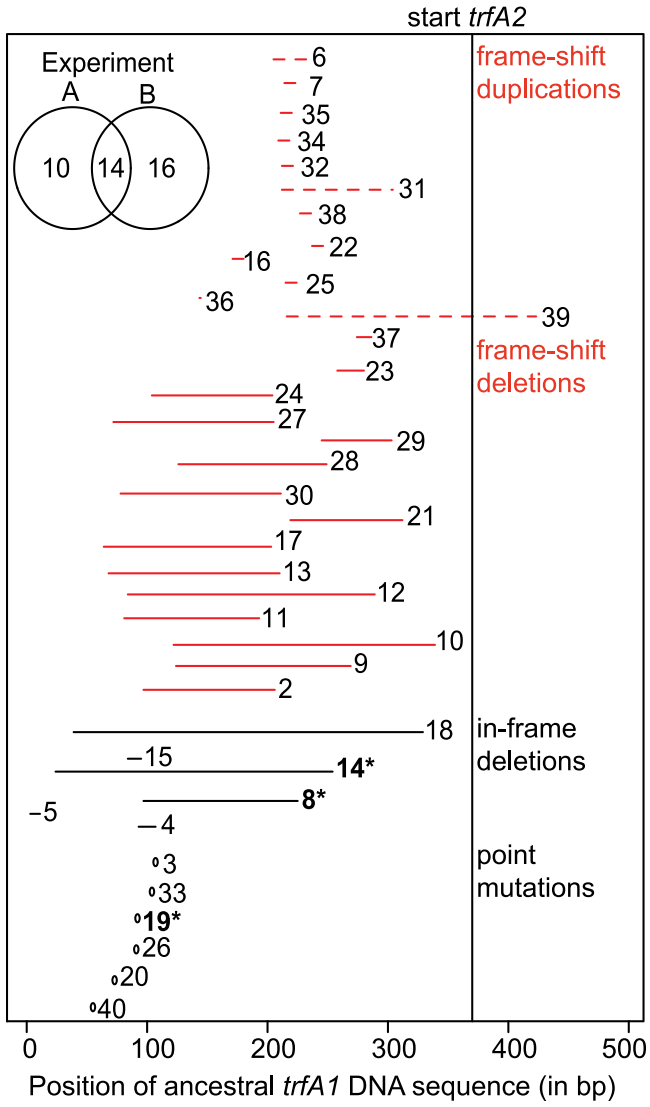

FIG 2 Positions of mutations within the first 500 nucleotides of *trfA1* for 40 evolved genotypes detected by Sanger sequencing. The ancestral sequence is represented by the *x* axis; dashed lines represent duplication mutations, solid lines represent deletions, and ovals represent point mutations. Changes in DNA sequence that resulted in frameshifts are indicated in red, and those that retained the ancestral reading frame are shown in black. The inset shows a Venn diagram with the number of unique *trfA1* genotypes observed in experiments A and B and those observed in at least one lineage of both experiments. Numbers next to mutations correspond to genotypes reported in Table S1 in the supplemental material, with bold numbers indicating that the presence of that genotype was detected via Sanger sequencing at generation 1,000 in lineages A1 to A5.

whereas more than 50% of the mutations observed earlier resulted in frameshifts (Fig. 2). This suggests that maintaining the reading frame of *trfA1* provided a selective advantage. Thus, plasmid-host adaptation through improvement of plasmid stability occurred through complex dynamics, suggesting competition between coexisting hosts with different plasmid genotypes.

We then investigated whether the same evolutionary dynamics described above would be observed when populations were sampled more deeply and sequenced directly (i.e., without cultivation/amplification of individual clones). This was done using the second approach: Roche 454 Titanium pyrosequencing of plasmid

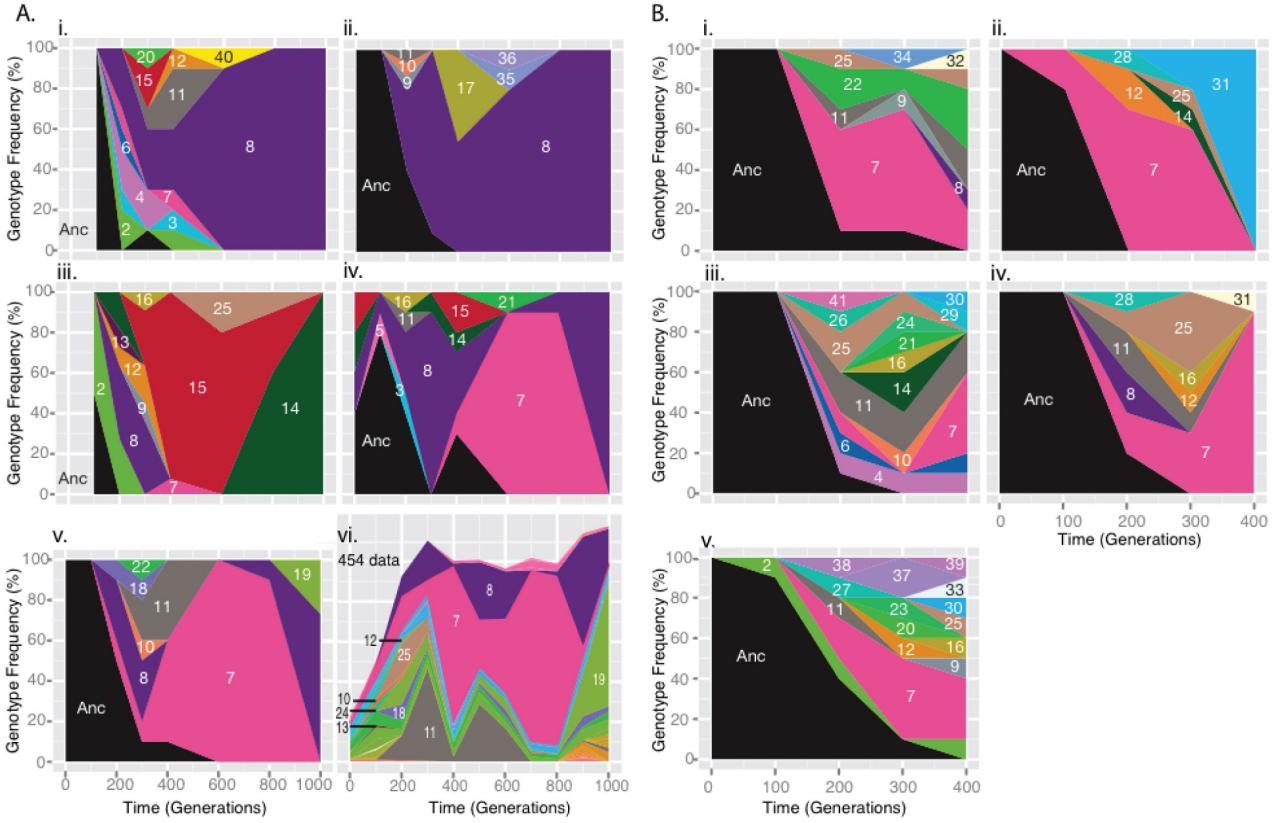

FIG 3 (A, panels i to v) Dynamics of *trfA* genotypes that arose over the course of 1,000 generations in five independently evolved lineages, A1 to A5 (experiment A). Genotype frequencies are based on Sanger sequencing of at least 10 clones from ancestral (Anc) populations (where available) and from populations evolved for 100 to 400 and 600, 800, and 1,000 generations. Numbers correspond to genotypes reported in Fig. 2; see also Table S1 in the supplemental material. (A, panel vi) Dynamics of *trfA* genotypes (not including the ancestral type) that arose in lineage A5, as observed by deep sequencing of plasmid populations. Numbers represent the genotypes previously observed via Sanger sequencing (A, panels i to v; see Fig. 2 and Table S1). Point mutations and indels of less than 4 bp that were detected at <1% were assumed to be sequencing errors and removed. See supplemental Methods S1 for an explanation of why total frequencies of the ancestral type cannot be accurately displayed and why the sum of genotype frequencies can exceed 100%. (B, panels i to v) Dynamics of *trfA* genotypes that arose in five independently evolved lineages, B1 to B5 (experiment B). Genotype frequencies are based on Sanger sequencing of 10 clones from ancestral (Anc) populations and populations evolved for 100 to 400 generations (note that the time scale is different from that in panel A).

DNA purified from 11 archived populations of lineage A5 (generations 0 to 1,000). Coverage of the 5′ *trfA1* region ranged from 180× (generation 100) to 1,412× (generation 600), with an average for all populations of 569×. The *trfA* genotypes and the dynamics in their frequencies were similar to those observed with the first sequencing approach, with some differences (Fig. 3A, panels vi and v). As expected, several *trfA1* genotypes were detected at high frequencies within the evolving populations, and a much higher number of them was identified through deep sequencing in this one lineage: a total of 104 compared to only 6 through the colony PCR approach. Until generation 200, evolved *trfA* genotypes were detected at relatively low frequencies (<7% each). As observed with the colony PCR approach, many rose to higher frequencies at subsequent time points. The frequencies of most also declined by generation 1,000, at which point the population was dominated by the same two genotypes, yet their relative proportions were reversed (~70/30% for genotypes 8/19 through colony PCR, and ~17/60% through pyrosequencing) (Fig. 3A, panels v and vi). Thus, regardless of the approach used, multiple *trfA* genotypes were observed simultaneously in a given evolving population at almost any time point, with a decline in frequency of most types by generation 1,000, resulting in one or two "winners."

Mutations detected in the remainder of the plasmid. To determine if mutations arose outside the 5′ end of *trfA1* and, if so, at what proportions in the populations, the plasmid population sequences obtained by Roche 454 Titanium sequencing were analyzed over the entire plasmid genome for all 11 time points of lineage A5. Our read cleanup procedure resulted in an average read length of 337 nucleotides. The average plasmid coverage ranged from 183× (generation 100) to 1,331× (generation 600), with a mean of 523×. While mutations were observed outside the *trfA1* 5′ region, none of these reached a frequency above 30% percent (see Table S2 in the supplemental material). Of the 11 mutations that obtained the highest frequency (>10%) in at least one time point, 5 were already observed in the ancestral population at frequencies between 14 and 25% and were 1-bp insertions or deletions in homopolymer (and typically intergenic) regions. It is not yet clear at this point whether these differences in sequences represent true mutations or sequencing errors. Four other mutations occurred at different locations in *oriV*, though none persisted in observable frequencies over more than 2 consecutive time points (100 generations). Thus, besides the mutations in the *trfA* gene, not a single mutation in the rest of the 13-kb plasmid seemed

to have a large enough beneficial effect on host fitness to sweep through the population.

DISCUSSION

A major concern in the alarmingly rapid spread of antibiotic resistance genes among bacterial pathogens is the persistence of antibiotic resistance plasmids even in the absence of the drugs. In order to avoid further escalation of this medical crisis, we must understand the evolutionary mechanisms by which plasmids become more persistent in novel hosts, as well as the underlying dynamics. Several studies suggest that plasmid cost and instability are key factors that facilitate the replacement of drug-resistant bacteria by drug-sensitive bacteria in a population (29–32). Additionally plasmid specialization to a nonpathogenic host can result in a shift in host range away from a human pathogen (8). The underlying evolutionary dynamics of such host range shifts have not previously been studied. In the current study, we showed that a plasmid can adapt to a novel host within 200 cell doublings under the given experimental conditions. Improving our knowledge of the rate and dynamics of plasmid adaptation will help inform antibiotic use policies, such as antibiotic cycling and short-term discontinuation regimens (29, 33).

A striking result of this study was the high diversity of clones with unique mutations in the first 340 bp of the *trfA1* gene over the course of two evolution experiments, with multiple genotypes simultaneously present in the same population. This result has at least two important implications. First, it highlights the high rate at which plasmids can diversify due to point mutations as well as small and large deletions and duplications. The 5′ end of the *trfA1* gene, known to encode an unstructured tail of the TrfA1 protein (8), seems to be a particular hot spot for mutations. Many of the deletions and duplications occurred between short direct repeats, probably due to slipped-strand mispairing, and were more frequently detected than point mutations (Fig. 2; see also Table S1 in the supplemental material). Many more *trfA1* mutations were detected here, in addition to those that were described by Sota et al. at generation 1,000 in the same 340-bp region and shown to improve plasmid stability (8) (Fig. 2; see also Table S1). We tested a few of the evolved plasmids with new *trfA1* mutations individually and showed a significant improvement in stability compared to pMS0506 (see Fig. S3 in the supplemental material). This high diversity generated in real time under laboratory conditions is consistent with the large sequence divergence observed in this gene region among natural IncP-1 plasmids. Indeed, while the shorter *trfA2* gene is conserved enough to allow alignments and infer phylogenetic history, the 5′ end of *trfA1* is not (8, 34).

The second ramification of the simultaneous presence of multiple evolved genotypes at high frequencies is the importance of clonal interference in plasmid evolution dynamics (Fig. 3; see Table S1 in the supplemental material). Indeed, the data suggest that under strong selection from an antibiotic on a strain that poorly maintained its resistance plasmid, competition between bacteria with distinct plasmid genotypes occurred. Unlike some studies that showed that stable polymorphisms persist due to metabolic interactions (35), here the polymorphism seemed to disappear after 1,000 generations, and one genotype often won the competition. Since our plasmid was nonconjugative, the bacterial populations could not exchange plasmid DNA by conjugation and therefore most likely evolved asexually. Our study provides one of the few instances of direct experimental evidence for clonal interference in evolving asexual populations at the molecular level (16–20, 27). In asexual populations whose evolutions are influenced by clonal interference, beneficial mutations from different parent cells do not have the opportunity to co-occur in daughter cells. Therefore, competition arises between variants with beneficial mutations, which slows the rate of fixation of any given genotype. This in turn allows more time for mutations of larger fitness effect to arise and become dominant in the populations (24). Even though clonal interference may thus have slowed down the rate at which individual stability-improving mutations swept through the populations, the plasmid population as a whole adapted rapidly to the novel host and already reached maximum stability after 200 generations of growth. In a separate study, a mathematical model was developed that allows simulating bacterial evolution in the presence of clonal interference. The first simulations of evolution of our 13.1-kb plasmid in a large bacterial population show that multiple mutants arise and compete, much like the dynamics observed in this experimental study (Z. Abdo, M. J. Stein, A. Wojtowicz, J. M. Hughes, K. Pepin, and E. M. Top, in preparation).

To determine the differences in fitness between evolved genotypes, we measured maximum growth rates as a proxy for fitness, using optical density measurements as previously described (8). We were unable to detect any significant differences between the few *trfA1* variants tested (genotypes 7, 8, 10, 11, 18, 19, and 22), but all growth rates were significantly higher than that of the ancestral host with ancestral plasmid (data not shown). Interestingly, only in-frame mutations fixed in the populations of experiment A after 1,000 generations, while most genotypes detected earlier were frameshifts (Fig. 2). Yano et al. recently confirmed that at least one of these frameshift mutants no longer produced the TrfA1 protein but replicates by virtue of the shorter TrfA2 (34). The cost of these plasmid variants may thus be slightly different, but our growth assays may not have been sensitive enough to detect small differences in maximum growth rates. Alternatively, there may be another parameter that explains fitness differences in serial batch cultures, such as shortened lag phase or improved survival in stationary phase. In future competition experiments, we will try to determine subtle differences in fitness between clones with different *trfA1* mutations in their plasmids. It should be noted that in our previous study, hosts with *trfA1* frameshift mutations were able to sweep through two of the five populations after 1,000 generations when the plasmid was transferred back to the ancestor every 100 generations (host-switching protocol), yet not in a serial passage protocol (non-host-switching, experiment A) (8). The number of lineages in each protocol (five) is however too small to draw conclusions from this observation. Studies to determine the molecular mechanisms underlying the plasmid stability improvement for evolved plasmid with frameshift and in-frame mutations are under way.

In addition to parallel evolution at the gene level, parallel evolution at the nucleotide level was also frequently observed, as several plasmid genotypes were found repeatedly in independently evolved lineages and in both evolution experiments (Fig. 2; see Table S1 in the supplemental material). Some of these swept through multiple, independently evolved populations. For example, the same 129-bp deletion (genotype 8) swept through 4 of the 5 populations after 1,000 generations (Fig. 3), even though 23 other unique plasmid genotypes were observed at earlier time points through colony PCR. One possible explanation for this is that the prevalence of only in-frame mutations by generation

1,000 was a combination of chance or sampling error. Stochastic processes are important in determining the fate of beneficial mutations, as previously predicted by theoretical models and deduced from experimental phage evolution data (20). Yet given the dominance of one specific *trfA* genotype in four of five independent populations, it seems unlikely that its fitness effect was the same as all other *trfA* mutations and its fate was due solely to chance. Another explanation could be that the 129-bp deletion occurred at a higher frequency than other mutations because it was likely due to slipped-strand mispairing (36). Yet many other observed deletions and duplications occurred between direct repeats (see Table S1). Moreover, frameshift mutations caused by small duplications seemed to occur earlier than this deletion, suggesting that they may arise more frequently and not less, which contradicts this explanation. We hypothesize that some mutations like this deletion conferred a slightly greater fitness benefit to the host than others and that we were unable to measure that fitness difference so far. Future studies, including competition assays, simulation experiments, and statistical modeling of our data, will help determine how differences in mutation frequency and selection coefficients between mutations can best explain the observed evolutionary dynamics.

It is likely that *trfA1* mutations arose prior to generation 200, given that these mutations were observed at frequencies of more than 10% by that point in all lineages. Indeed, *trfA* mutations were observed even in the ancestral population of lineages A4 and A5 using PCR and 454 sequencing, respectively. The high frequency of multiple plasmid genotypes at time 0 may be surprising, yet genetic diversity in stock cultures has been recently reported for other bacteria (37). It can be explained in part by the additional freezing, thawing, and cultivation steps that occurred between transformation of the ancestral plasmid into the ancestral host and sequencing these populations. This is particularly true for the 454 sequencing data of lineage A5, from which plasmid DNA was extracted after growing large volumes of culture. Therefore, our genotype frequency observations for the ancestral and evolved populations of that lineage reflect a population structure that is several generations removed from the point of archival. This slight shift can in fact be seen when comparing the dynamics in Fig. 3A, panels vi and v. To avoid the bias introduced through freezing and regrowing populations, evolution experiment B was conducted; it was initiated with five clones of host MR-1 after transformation with plasmid pMS0506. Although no deep-level sequencing was performed on these populations, not a single *trfA* variant was observed at generation 0 in any of the five lineages (Fig. 3B). In summary, *trfA* mutations arose very early in the evolution experiment, and some may already have been present in the ancestral populations.

In lineage A5, analyzed by 454 sequencing, several putative mutations were also identified outside *trfA* in the evolving plasmid populations. However, even though 11 of these mutations were occasionally detected at frequencies higher than 10%, none of them swept through or even consistently increased in frequency over time (see Table S2 in the supplemental material). When Sota et al. sequenced one evolved plasmid from each of their independently evolved lineages, after 1,000 generations (experiment A), they observed no mutations at all outside *trfA* (8). Therefore, the nucleotide changes we have identified here (see Table S2) represent either sequencing errors or mutations with no or small positive fitness effects, since none of them seemed to be linked on the same plasmid with the most dominant *trfA1* mutation by generation 1,000. Notably, 9 of the 11 single nucleotide changes were 1-bp deletions in intergenic homopolymer regions. Their high frequencies may well be overestimated or completely false due to a known high error rate of 454 sequencing of homopolymers (36). Future studies will have to confirm the putative mutations observed here and determine their potential importance in plasmid stability improvement.

Thanks to the greater depth of coverage afforded by Roche 454 Titanium deep sequencing of plasmid DNA extracted from entire populations, we were able to discern a larger variety of mutants present at lower frequencies than with the colony PCR method (Fig. 3A, panel vi). This approach also diminished the probability of missing certain genotypes due to possible PCR priming bias in cases where mutations (in particular deletions) may have been in primer binding sites. One limitation of both approaches, however, is the fact that some small deletions or point mutations may not have been accounted for if the genome in question acquired a deletion mutation that encompassed the previous mutation. While this is a possibility to keep in mind, double mutations of this sort may not have occurred frequently enough to significantly change estimated genotype frequencies. Also of note is that the results of the temporal dynamics of genotype frequencies were quite similar between the PCR and deep-sequencing approaches for lineage A5, except that the relative abundances of genotypes 8 (129-bp deletion) and 19 (point mutation) were reversed at generation 1,000. It is unlikely that this difference is due to differences in plasmid copy number affecting the proportions of *trfA* genotypes measured by the deep-sequencing approach. We previously showed that the plasmid copy numbers of the two genotypes in the ancestral hosts were 4.7 and 5.7. For the observed reversion in relative frequency to be due to copy number only, the difference in copy number would have to be 8-fold. The high proportion of genotype 19 detected by the deep-sequencing method is more consistent with the fact that Sota et al. (8) determined the plasmid sequence for a random clone from that same population and found it to have the point mutation of genotype 19. Thus, despite some limitations, the deep-sequencing method can be a useful and cost-effective tool to monitor the evolutionary dynamics of evolving bacterial plasmid populations.

The tools available for 454 sequence data analysis typically do not permit analyzing multiple mutations of different types that occur at the same loci, when these include moderate to large indels (up to 413 bp in our study) (16, 38, 39). Most of those studies interested in bacterial evolution have so far reported only genotypes of dominant members of the population (18, 26). Phage evolution studies have been able to sample deeper but were also typically limited to less than 30 complete genomes per population (23, 40, 41). In this study, we designed a procedure that reports the presence and frequency of point mutations as well as both large and small indels in population sequence data, even when they occur at the same position in the genome within a population. Our script is available upon request.

In conclusion, we showed that a broad-host-range plasmid of the incompatibility group IncP-1 improved its persistence in a novel host as quickly as within 200 generations. During this short time, a surprisingly high diversity of genotypes arose with mutations in the replication initiation gene *trfA*. Moreover, these evolved variants seemed to compete for dominance within populations, indicating that clonal interference can play an important

role in the evolution of antibiotic resistance plasmids. These observed evolutionary dynamics were highly repeatable between independent evolution experiments, suggesting strong selection on a rapidly evolving plasmid replication initiation gene.

MATERIALS AND METHODS

Plasmid and bacterial strain. The 13.1-kb minireplicon pMS0506 was constructed previously (8) from the natural cryptic IncP-1β plasmid pBP136, found in *Bordetella pertussis* (42). This was done by removing an entire section containing the two transfer regions and inserting a kanamycin resistance gene. This minireplicon thus contains only the genes necessary for plasmid replication, maintenance, and control and is not self-transferable. The host used in this study was *Shewanella oneidensis* MR-1, ATCC 700550 (from here on referred to as MR-1), a gammaproteobacterium shown previously to poorly maintain the plasmid in the absence of kanamycin (8). The copy number of the plasmid in the ancestral host was approximately 3 plasmids per chromosome (8).

Evolution experiment. The first set of five replicate-evolved populations of MR-1 (pMS0506) was obtained from a previously published evolution experiment carried out using the "nonswitching host protocol," which simply consisted of daily serial batch transfers at 10 doublings per day (8). The maximum population size was approximately 2.5×10^{10} CFU (5 ml at 5×10^9 CFU/ml), and the daily bottleneck was 2.5×10^7 CFU (1,000-fold dilutions). From here on, this first evolution experiment is referred to as "experiment A," and these five evolving populations (lineages) are named A1 to A5. We then repeated the first 400 generations of the evolution experiment (named "experiment B"). The same procedure was followed, except that a concentration of 50 μg/ml kanamycin (Km50) instead of 100 μg/ml was used. To start experiment B, plasmid pMS0506 was reintroduced into host MR-1 through electroporation, and five resulting clones were chosen at random to inoculate 5 ml of Luria broth (LB) with 50 μg/ml of kanamycin (LB Km50). These cultures were incubated at 30°C with shaking and constituted the five ancestral populations of lineages B1 to B5 of experiment B.

Plasmid stability assays. Plasmid stability in the absence of selection for the plasmid was used as a measure of plasmid adaptation, and stability assays were conducted as described previously (8). Briefly, to assess the change in fraction of plasmid-bearing cells over time in ancestral and evolving populations, three aliquots from each population under study were inoculated in three test tubes containing 5 ml LB Km50. In the case of experiment A, approximately 2- to 5-μl aliquots were taken from archived populations; in experiment B, the ancestral and evolving populations were sampled and analyzed immediately, before archiving. Cultures were transferred daily to LB without antibiotics (4.88 μl in 5 ml), and every day or every other day they were diluted and plated on LB agar (LBA). Fifty randomly chosen clones were then replicated onto LBA and LBA Km50, and the fraction of Km-resistant clones was considered the fraction of plasmid-bearing cells.

To determine the stability of evolved plasmids in the ancestral host, these plasmids were first transferred to the ancestral host as follows. Plasmid DNA was extracted from individual populations archived after 100, 200, 300, 400, and 1,000 generations of evolution by a standard plasmid extraction protocol based on alkaline lysis and introduced into plasmid-free ancestral cells of MR-1 through electroporation (43). The pools of transformants were resuspended in 5 ml LB Km50 and used as starting populations for plasmid stability assays, as described above. Plasmid stability was also tested for randomly chosen evolved clones representing each *trfA* genotype from lineage A5.

Statistical analysis of plasmid stability. Logistic regression (44) was used to evaluate changes in plasmid stability within evolving lineages over the course of an evolution experiment. We assumed that the observed lineages represented a random sample from a collection of all possible lineages. Accordingly, we constructed a random-effects, hierarchical logistic regression model that allowed for different slopes and intercepts of the stability curves for every 100-generation interval within each lineage, while allowing for differences between lineages and accounting for possible overdispersion between replicate stability assays. A Bayesian framework was used to fit and evaluate this model. The analyses were done on all available data for the entire course of the stability assays, including time points beyond those shown in Fig. 1 and Fig. S1 and S2 in the supplemental material for some of the lineages and generation times. These additional time points are not shown in the graphs for reasons of consistency and clarity of presentation. Details of this model, its implementation, and a simplified variation (used to determine differences in stability between different plasmid genotypes within a certain lineage) can be found in the supplemental material.

Sanger sequencing and analysis of *trfA1*. To determine the evolutionary dynamics of plasmid adaptation, we first determined the nucleotide sequence of the 5′ end of *trfA1* for 10 to 15 randomly chosen clones per population. This was done for all populations archived after 0, 100, 200, 300, 400, 600, 800, and 1,000 generations from all five lineages of evolution experiment A. However, we were unable to revive the archived ancestral cultures used to initiate lineages 1 and 3. The DNA sequence of this same gene region was also determined for all five lineages in the repeated evolution experiment B, for 10 clones per population evolved for 0, 100, 200, 300, and 400 generations. Partial *trfA1* amplification was performed as described previously through amplification by colony PCR (8). Nucleotide sequences were determined with an ABI PRISM model 3730 DNA analyzer (Applied Biosystems, Foster City, CA). Sequence analysis was performed with software programs CLC Sequence Viewer (CLC bio A/S, Cambridge, MA) and FinchTV (Geospiza Inc., Seattle, WA).

Deep Roche 454 pyrosequencing of plasmid pMS0506 in evolving populations. To perform shotgun pyrosequencing, plasmid pMS0506 DNA was first isolated and purified from archived populations of lineage A5. Given the poor viability of the archived cultures (in 20% glycerol), a specific volume was sampled from the frozen cultures to obtain at least 10,000 cells. This was inoculated in 5 ml LB Km50, and cultures were grown overnight at 30°C with shaking. Subsequently, 30 ml LB Km50 was inoculated with 4.8 ml of the overnight culture and grown for 2.5 h at 30°C with shaking. Then, 26 ml of this culture was transferred to 500 ml fresh LB Km50 and grown for another 2 h. Plasmid DNA was extracted from these exponentially growing cultures using a standard alkaline lysis protocol without phenol-chloroform extractions (43). Since strain MR-1 harbors a 161.6-kb native plasmid, pMR-1 (45), pMS0506 was separated from both the native plasmid and any chromosomal DNA by agarose gel electrophoresis and electroelution (43). Residual ethidium bromide was extracted by using a 1:1 vol of 100% *N*-butanol (43). For each of the eleven samples, a Roche 454 pyrosequencing library was prepared as described in the Roche rapid library preparation protocol and sequenced on the Roche 454 Genome Sequencer FLX using a separate multiplex identifier (MID) tag for each sample.

Roche 454 sequencing reads were then cleaned and processed to identify MID and adapter sequences and to remove poor-quality regions. We then developed a custom procedure to identify unique variants (SNPs and indels) and determine their relative frequencies by counting the number of variant occurrences divided by the average coverage for that region. These procedures are further described in detail in supplemental Methods S1.

Due to the increased error rates associated with 454 pyrosequencing in homopolymer regions, all homopolymer regions were identified in the reference plasmid genome using a custom python script. In order to reduce the number of error rates categorized as mutations, we chose a strict cutoff value for small mutations. Any small mutation (up to 3 bp) that occurred within a homopolymer region was not included in further analysis if its frequency was below 1%.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at http://mbio.asm.org/lookup/suppl/doi:10.1128/mBio.00077-12/-/DCSupplemental.

Methods S1, DOCX file, 0.1 MB.

Figure S1, EPS file, 0.5 MB.

Figure S2, EPS file, 0.4 MB.

Figure S3, EPS file, 0.3 MB.
Table S1, PDF file, 0.1 MB.
Table S2, PDF file, 0.1 MB.

ACKNOWLEDGMENTS

We thank L. Forney and H. Yano for useful suggestions. We also thank L. Rogers, D. Sen, J. Krol, and K. Turner for technical assistance.

This work was supported by NIH grants R01 AI084918 from the National Institute of Allergy and Infectious Diseases (NIAID) and P20 RR16448 from the COBRE program of the National Institute for General Medical Sciences (NIGMS) and in part by the National Science Foundation BEACON Center for the Study of Evolution in Action under cooperative agreement no. DBI-0939454. Support was also provided in the form of two undergraduate research fellowships to B. K. Lohman (B. and G. Hill and Biological Sciences, University of Idaho). We also acknowledge support from the IBEST Computational and Genomics Resources Core facilities and staff supported by NIH COBRE grant P20RR16448.

REFERENCES

1. **Gogarten JP, Doolittle WF, Lawrence JG.** 2002. Prokaryotic evolution in light of gene transfer. Mol. Biol. Evol. **19**:2226–2238.
2. **Jain R, Rivera MC, Moore JE, Lake JA.** 2002. Horizontal gene transfer in microbial genome evolution. Theor. Popul. Biol. **61**:489–495.
3. **Gogarten JP, Townsend JP.** 2005. Horizontal gene transfer, genome innovation and evolution. Nat. Rev. Microbiol. **3**:679–687.
4. **Thomas CM, Smith CA.** 1987. Incompatibility group P plasmids: genetics, evolution, and use in genetic manipulation. Annu. Rev. Microbiol. **41**:77–101.
5. **Krishnapillai V.** 1988. Molecular genetic analysis of bacterial plasmid promiscuity. FEMS Microbiol. Lett. **54**:223–237.
6. **Norman A, Hansen LH, Sørensen SJ.** 2009. Conjugative plasmids: vessels of the communal gene pool. Philos. Trans. R. Soc. Lond. B Biol. Sci. **364**: 2275–2289.
7. **De Gelder L, Ponciano JM, Joyce P, Top EM.** 2007. Stability of a promiscuous plasmid in different hosts: no guarantee for a long-term relationship. Microbiology **153**:452–463.
8. **Sota M, et al.** 2010. Shifts in the host range of a promiscuous plasmid through parallel evolution of its replication initiation protein. ISME J. **4**:1568–1580.
9. **Heuer H, Fox RE, Top EM.** 2007. Frequent conjugative transfer accelerates adaptation of a broad-host-range plasmid to an unfavorable *Pseudomonas putida* host. FEMS Microbiol. Ecol. **59**:738–748.
10. **De Gelder L, Williams JJ, Ponciano JM, Sota M, Top EM.** 2008. Adaptive plasmid evolution results in host-range expansion of a broad-host-range plasmid. Genetics **178**:2179–2190.
11. **Lenski RE, Bouma JE.** 1987. Effects of segregation and selection on instability of plasmid pACYC184 in *Escherichia coli* B. J. Bacteriol. **169**: 5314–5316.
12. **Bouma JE, Lenski RE.** 1988. Evolution of a bacteria/plasmid association. Nature **335**:351–352.
13. **Dionisio F, Conceição IC, Marques ACR, Fernandes L, Gordo I.** 2005. The evolution of a conjugative plasmid and its ability to increase bacterial fitness. Biol. Lett. **1**:250–252.
14. **Lenski RE, Simpson SC, Nguyen TT.** 1994. Genetic analysis of a plasmid-encoded, host genotype-specific enhancement of bacterial fitness. J. Bacteriol. **176**:3140–3147.
15. **Modi RI, Adams J.** 1991. Coevolution in bacterial-plasmid populations. Evolution **45**:656–667.
16. **Barrick JE, Lenski RE.** 2009. Genome-wide mutational diversity in an evolving population of *Escherichia coli*. Cold Spring Harb. Symp. Quant. Biol. **74**:119–129.
17. **Barrick JE, et al.** 2009. Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. Nature **461**:1243–1247.
18. **Imhof M, Schlötterer C.** 2001. Fitness effects of advantageous mutations in evolving *Escherichia coli* populations. Proc. Natl. Acad. Sci. U. S. A. **98**:1113–1117.
19. **Maharjan R, Seeto S, Notley-McRobb L, Ferenci T.** 2006. Clonal adaptive radiation in a constant environment. Science **313**:514–517.
20. **Pepin KM, Wichman HA.** 2008. Experimental evolution and genome sequencing reveal variation in levels of clonal interference in large populations of bacteriophage ΦX 174. BMC Evol. Biol. **8**:85.
21. **Wichman HA, Millstein J, Bull JJ.** 2005. Adaptive molecular evolution for 13,000 phage generations: a possible arms race. Genetics **170**:19–31.
22. **Wichman HA, Badgett MR, Scott LA, Boulianne CM, Bull JJ.** 1999. Different trajectories of parallel evolution during viral adaptation. Science **285**:422–424.
23. **Miller CR, Joyce P, Wichman HA.** 2011. Mutational effects and population dynamics during viral adaptation challenge current models. Genetics **187**:185–202.
24. **Gerrish PJ, Lenski RE.** 1998. The fate of competing beneficial mutations in an asexual population. Genetica **102**:127–144.
25. **De Visser JA, Rozen DE.** 2006. Clonal interference and the periodic selection of new beneficial mutations in *Escherichia coli*. Genetics **172**: 2093–2100.
26. **Fogle CA, Nagle JL, Desai MM.** 2008. Clonal interference, multiple mutations and adaptation in large asexual populations. Genetics **180**: 2163–2173.
27. **Lang GI, Botstein D, Desai MM.** 2011. Genetic variation and the fate of beneficial mutations in asexual populations. Genetics **188**:647–661.
28. **Miralles R, Gerrish PJ, Moya A, Elena SF.** 1999. Clonal interference and the evolution of RNA viruses. Science **285**:1745–1747.
29. **Lenski RE.** 1997. The cost of antibiotic resistance—from the perspective of a bacterium, p 131–151. *In* Chadwick DJ, Goode J (ed), Antibiotic resistance: origins, evolution, selection and spread. Wiley, Chichester, United Kingdom.
30. **Spratt BG.** 1996. Antibiotic resistance: counting the cost. Curr. Biol. **6**:1219–1221.
31. **Levin BR.** 2001. Minimizing potential resistance: a population dynamics view. Clin. Infect. Dis. **33**:S161–S169.
32. **De Gelder L, et al.** 2004. Combining mathematical models and statistical methods to understand and predict the dynamics of antibiotic-sensitive mutants in a population of resistant bacteria during experimental evolution. Genetics **168**:1131–1144.
33. **Levy SB.** 1994. Balancing the drug-resistance equation. Trends Microbiol. **2**:341–342.
34. **Yano H, Deckert GE, Rogers LM, Top EM.** 2012. The role of long and short replication initiation proteins in the fate of IncP-1 plasmids. J. Bacteriol. **194**:1533–1543.
35. **Rosenzweig RF, Sharp RR, Treves DS, Adams J.** 1994. Microbial evolution in a simple unstructured environment: genetic differentiation in *Escherichia coli*. Genetics **137**:903–917.
36. **Albertini AM, Hofer M, Calos MP, Miller JH.** 1982. On the formation of spontaneous deletions: the importance of short sequence homologies in the generation of large deletions. Cell **29**:319–328.
37. **Nahku R, et al.** 2011. Stock culture heterogeneity rather than new mutational variation complicates short-term cell physiology studies of *Escherichia coli* K-12 MG1655 in continuous culture. Microbiology **157**: 2604–2610.
38. **Nusbaum C, et al.** 2009. Sensitive, specific polymorphism discovery in bacteria using massively parallel sequencing. Nat. Methods **6**:67–69.
39. **Dalca AV, Rumble SM, Levy S, Brudno M.** 2010. VARiD: a variation detection framework for color-space and letter-space platforms. Bioinformatics **26**:i343–i349.
40. **Poullain V, Gandon S, Brockhurst MA, Buckling A, Hochberg ME.** 2008. The evolution of specificity in evolving and coevolving antagonistic interactions between a bacteria and its phage. Evolution **62**:1–11.
41. **Betancourt AJ.** 2009. Genomewide patterns of substitution in adaptively evolving populations of the RNA bacteriophage MS2. Genetics **181**: 1535–1544.
42. **Kamachi K, et al.** 2006. Plasmid pBP136 from *Bordetella pertussis* represents an ancestral form of IncP-1β plasmids without accessory mobile elements. Microbiology **152**:3477–3484.
43. **Sambrook J, Russell DW.** 2001. Molecular cloning: a laboratory manual, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
44. **Gelman A, Hill J.** 2007. Data analysis using regression and multilevel/hierarchical models, p 625. Cambridge University Press, New York, NY.
45. **Heidelberg JF, et al.** 2002. Genome sequence of the dissimilatory metal ion–reducing bacterium *Shewanella oneidensis*. Nat. Biotechnol. **20**: 1118–1123.