

Directed evolution of *Escherichia coli* with lower-than-natural plasmid mutation rates

Daniel E. Deatherage, Dacia Leon, Álvaro E. Rodriguez, Salma K. Omar and Jeffrey E. Barrick*

Department of Molecular Biosciences, Center for Systems and Synthetic Biology, The University of Texas at Austin, Austin, TX 78712, USA

Received May 01, 2018; Revised August 03, 2018; Editorial Decision August 06, 2018; Accepted August 08, 2018

ABSTRACT

Unwanted evolution of designed DNA sequences limits metabolic and genome engineering efforts. Engineered functions that are burdensome to host cells and slow their replication are rapidly inactivated by mutations, and unplanned mutations with unpredictable effects often accumulate alongside designed changes in large-scale genome editing projects. We developed a directed evolution strategy, Periodic Reselection for Evolutionarily Reliable Variants (PResERV), to discover mutations that prolong the function of a burdensome DNA sequence in an engineered organism. Here, we used PResERV to isolate *Escherichia coli* cells that replicate ColE1-type plasmids with higher fidelity. We found mutations in DNA polymerase I and in RNase E that reduce plasmid mutation rates by 6- to 30-fold. The PResERV method implicitly selects to maintain the growth rate of host cells, and high plasmid copy numbers and gene expression levels are maintained in some of the evolved *E. coli* strains, indicating that it is possible to improve the genetic stability of cellular chassis without encountering trade-offs in other desirable performance characteristics. Utilizing these new antimutator *E. coli* and applying PResERV to other organisms in the future promises to prevent evolutionary failures and unpredictability to provide a more stable genetic foundation for synthetic biology.

INTRODUCTION

Populations of cells engineered to function as factories or biosensors experience a failure mode that is peculiar to living systems: they evolve. Unwanted evolution is a foundational problem for bioengineering that limits the efficiency and predictability of metabolic and genome engineering efforts (1–5). Often an engineered function diverts critical resources from cellular replication or otherwise interferes

with growth or homeostasis (6,7). In these cases, ‘broken’ cells with mutations that inactivate the engineered function can rapidly outcompete the original design (8–10). The rate at which an engineered function decays within a cell population in this manner can be summarized as an evolutionary lifetime or half-life (8) or defined in terms of an evolutionary landscape by the rates at which various mutational failure modes occur and their respective fitness benefits (9,10).

It is sometimes possible to edit a genome to eliminate or reduce the rate at which certain types of mutations occur (11–16) or to devise a way of reducing the burden of an engineered function (2,17). However, given the complexity of DNA replication and repair processes and the multifarious ways that an engineered function can burden a host cell, a point is generally reached at which it is difficult to further improve upon the reliability of a cell. Directed evolution is an effective strategy for optimizing the performance of complex systems with many interacting components, even when they include unknown factors. For example, it has been used to engineer novel enzymes that outperform their natural counterparts (18) and to tune artificial gene circuits to effectively perform logic operations (19).

Given the similarly complex constraints underlying cellular mutagenesis and the fitness burdens of diverse engineered functions, we reasoned that a directed evolution procedure, Periodic Reselection for Evolutionarily Reliable Variants (PResERV) (Figure 1), could be an effective strategy for improving the evolutionary reliability of an engineered cell. In PResERV, one artificially selects for mutant cells that exhibit improved maintenance of a burdensome engineered function over tens to hundreds of cell divisions. We expected that PResERV might isolate cells with mutations that either reduced the rate at which failure mutations occurred or the fitness burden of the engineered function, or both, possibly in ways that would generalize to stabilizing other engineered functions in the evolved cells. Here, we describe *Escherichia coli* strains evolved by PResERV that exhibit lower-than-natural mutation rates for genes encoded on high-copy plasmids, thereby stabilizing them against unwanted evolution.

*To whom correspondence should be addressed. Tel: +1 512 471 3247; Fax: +1 512 471 2149; Email: jbarrick@cm.utexas.edu

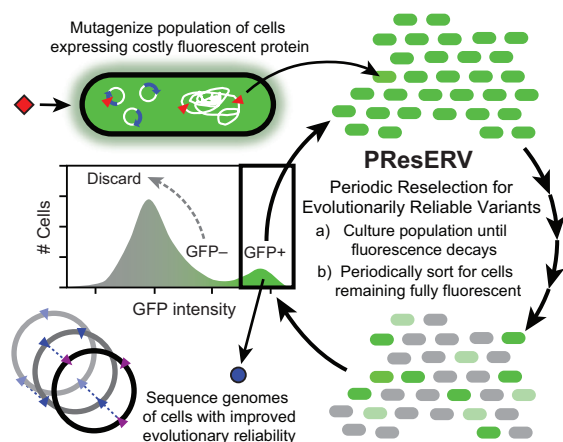


Figure 1. Periodic Reselection for Evolutionarily Reliable Variants (PResERV) method. PResERV begins with a population of cells expressing GFP to such a high level that it imposes a significant fitness burden. After mutagenesis, the population is cultured through enough cell doublings that mutants with reduced GFP expression arise and outcompete other cells. Periodically, the population is sorted to retain only those cells that remain as fluorescent as the original strain, enriching for mutant host cells with reduced mutation rates or a lower fitness cost for GFP expression. Once the evolutionary stability of GFP expression increases, fluorescent cells are isolated and their genomes are sequenced to identify and characterize the genetic changes that contribute to this improvement. Regrowth of cells during PResERV implicitly selects for only those mutants that achieve improved genetic stability without introducing any trade-offs that significantly reduce cellular growth rates.

MATERIALS AND METHODS

Culture conditions

Escherichia coli was grown as 10 ml cultures in 50 ml Erlenmeyer flasks with incubation at 37°C and 120 rpm orbital shaking over a diameter of 1 in. unless otherwise noted. The Miller formulation of Lysogeny Broth (LB) was used (10 g/l tryptone, 5 g/l yeast extract, and 10 g/l NaCl). Media were supplemented with 100 µg/ml carbenicillin (Crb), 20 µg/ml chloramphenicol (Cam), 50 µg/ml kanamycin (Kan), 100 µg/ml rifampicin (Rif) and 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), as indicated. Bacterial cultures were frozen at -80°C after adding glycerol as a cryoprotectant to a final concentration of 13.3% (v/v).

Strains and plasmids

The progenitor strain (BW25113) of the Keio knockout collection (20) was transformed with pSKO4. This plasmid contains the redesigned I7101 (R0010+E0240) circuit, which was edited to remove unstable repeat sequences in a prior study by Sleight *et al.*, on the BioBrick cloning vector pSB1A2 backbone (21). It is a ColE1 group plasmid with a pBR322 origin of replication. Plasmid pTEM-1.D254tag encodes TEM-1 β-lactamase with the codon for an amino acid at a surface-exposed position in the enzyme's structure at which multiple amino acid substitutions are compatible with enzyme function replaced with a TAG stop codon (22). pTEM-1.D254tag has a pBR322 origin of replication and additionally encodes the *rop* protein.

UV mutagenesis

BW25113 cells containing pSKO4 were cultured overnight to stationary phase in LB-Crb. Then, these cultures were pelleted by centrifugation and resuspended in an equal volume of sterile saline. Eleven 120 µl droplets of these cell suspensions were spotted on petri dishes and subjected to 27 500 µJ/cm² of 254 nm UV radiation in a UVP CL-1000 crosslinker. After UV exposure, 100 µl from each droplet was combined and pelleted by centrifugation to collect ~2.5 × 10⁶ surviving cells. These cells were inoculated into 10 ml of LB-Crb and grown to a final density of ~2 × 10⁹ cells/ml. This mutagenized library was archived as a frozen stock.

PResERV directed evolution procedure

All growth steps were conducted in 10 ml of LB-IPTG-Crb. We used 0.1 ml of the mutagenized library to found the experimental population. After overnight growth to saturation, we propagated the population through daily 1:1000 dilutions of saturated cultures into fresh media followed by regrowth. GFP expression was monitored using a BD Fortessa flow cytometer. Periodically, overnight cultures were diluted to ~2.5 × 10⁶ cells/ml in HPLC grade water and stained with 150 nM of the nucleic acid dye SYTO 17 (Life Technologies). The GFP+ portion of the population was calculated as the percentage of SYTO 17 positive cells with at least the ancestral level of GFP fluorescence by flow cytometry. When fewer than 25% of cells in the population were GFP+, instead of a normal transfer, the population was diluted to ~2.5 × 10⁶ cells/ml in HPLC grade water and between ~4 × 10⁴ and ~5 × 10⁵ GFP+ cells were sorted into 10 ml of fresh LB-IPTG-Crb using a BD FACSaria IIIu. The SYTO 17 dye was found to decrease cell viability, so we did not add this counterstain in the sorting steps, at the cost of less efficient enrichment of GFP+ over GFP- cells.

Isolation of evolved cells and plasmids

The evolved population was plated on LB-IPTG-Crb and six visibly GFP+ colonies were selected at random for further study. Each of these clonal isolates was grown overnight in LB-IPTG-Crb before isolating its plasmid and creating a frozen stock. To select for plasmid loss, we also diluted these cultures 1:1000 into media lacking Crb (LB-IPTG). After overnight growth, dilutions of the resulting cultures were then plated on LB-IPTG agar. GFP- colonies were patched onto LB-Crb agar to ensure the lack of fluorescence was caused by loss of plasmid rather than a GFP mutation. One colony which had been cured of its plasmid was selected for each of the original evolved clones and re-transformed with the wild-type plasmid. We were unable to cure one evolved strain (AER7) of the plasmid in this way. The wild-type BW25113 strain was separately transformed with the plasmids isolated from each of the six evolved clones. GFP decay experiments were carried out on the resulting eleven strains to examine them for evidence of increased evolutionary stability.

GFP decay curves

For each of the strains tested, individual colonies were used to inoculate nine replicate *E. coli* populations. In order to more accurately estimate the number of cell doublings elapsed since the single-cell bottleneck, care was taken to include all cells in each colony in the first liquid culture by excising and transferring the piece of agar underneath and around each colony. Each population was then subjected to daily transfers under the same conditions as the PResERV experiment, while monitoring GFP fluorescence using SYTO17 staining and a BD Fortessa flow cytometer as describe above. For creating graphs of the percentage of cells remaining GFP⁺ over time, flow cytometry data were analyzed in R using the flowCore Bioconductor package (v1.42.3) (23). Among the events exhibiting a SYTO17 signal, cells were classified as GFP⁺ if they were above a signal intensity threshold that was set based on the distribution of fluorescence values observed for the wild-type strain-plasmid combination in that experiment. For graphing and comparing the initial GFP fluorescence in each strain, median intensity values for the GFP⁺ subpopulation were log₂ transformed before performing statistical analyses.

Genome sequencing

DNA was extracted from stationary phase *E. coli* cultures using the PureLink Genomic DNA Mini kit from Life Technologies. Purified DNA was fragmented using the Covaris AFA system, and samples were prepared using the NEB-Next DNA Library Prep Reagent Set for Illumina kit from New England Biolabs. An Illumina HiSeq 2500 was used to generate 2 × 125 paired-end reads from each sample at The University of Texas at Austin Genome Sequencing Analysis Facility (GSAF). FASTQ files have been deposited in the NCBI sequence read archive (SRP090775). Mutations in each of the evolved strains were predicted by using *breseq* (version 0.32.0a) (24) to compare the Illumina reads to the *E. coli* BW25113 reference genome (GenBank:CP009273.1). Several genetic differences between this reference sequence and all four sequenced samples were assumed to have existed in the ancestral *E. coli* strain used to initiate the PResERV experiment and are not reported.

Strain reconstruction

Donor strains from the i-Deconvoluter library (25) were used to revert the three candidate evolved alleles we tested in two evolved clones (AER8 and AER12) back to wild type sequences using P1 transduction as previously described (26), except that only 2 μl of lysate was used in each transduction. Lysate from SMR20954, SMR20794, and SMR20838 was combined with the appropriate evolved strain with a mutation in *polA*, *polB*, or *rne*, respectively, and plated on LB agar plates containing Kan. Resultant colonies were screened for correct replacement via Sanger sequencing before FLP recombinase was used to remove the linked Kan^R cassette used for selection of transductants as previously described (27). The strain with *rne* reverted to the wild-type allele was subjected to a second round of P1 transduction to also revert the *polB* mutation, and the Kan^R cassette was again removed via FLP recombination.

Mutation rate measurements

Luria-Delbrück fluctuation tests were carried out to measure mutation rates (28). For plasmid mutation rates, strains cured of plasmid pSKO4 were transformed with plasmid pTEM-1.D254tag after making any genetic modifications to revert evolved alleles. Cultures were grown in LB-Cam to select for retention of this plasmid and plated on LB-Cam agar additionally supplemented with 500 μg/ml Crb to select for mutants. Thus, this assay measures the aggregate rate of all mutations that revert this stop codon to a permitted sense codon. Cells with the original, unmutated pTEM-1.D254tag plasmid were somewhat resistant to Crb, presumably due to some translational readthrough of the stop codon inserted into the β-lactamase gene on this high-copy plasmid. This background resistance is why an unusually high concentration of Crb (five times the level normally used to select for plasmid retention in this strain) was necessary to select for Crb^R mutants. For chromosomal mutation rates, LB agar containing 100 μg/ml rifampicin (Rif) or 60 μg/ml D-cycloserine (DCS) was used for the selective condition. Rif resistance requires specific point mutations in the *rpoB* gene (29). In minimal media, DCS resistance requires a loss-of function mutation specifically in the *cycA* gene (30), but mutations in additional targets may also be possible in the rich LB media used here.

Mutation rates were determined by taking an overnight culture of a strain and transferring ~1000 cells from a dilution in sterile saline to each separate fluctuation test culture. After growth of the replicate cultures to saturation, the entire volume of each one was either plated on a selective LB agar plate or a dilution was plated on a non-selective LB agar plate. For Crb^R plasmid and Rif^R chromosomal mutation rates comparing wild-type, evolved, and reconstructed strains, we used 0.2 ml cultures in 18 × 150 mm test tubes containing non-selective media and incubated these cultures with orbital shaking. A total of six non-selective and 48 or 12 selective plates were used for plasmid and chromosomal mutations, respectively. For DCS chromosomal mutation rates, 1.0 ml LB cultures and 12 selective media plates were used. Additionally, only a portion (25 μl) of these cultures were plated on the selection LB agar. For comparing Crb^R plasmid mutation rates of wild-type and evolved clones (AER8 and AER12), we conducted three separate sets of fluctuation tests using different growth formats. In the first, we grew 1 ml cultures in test tubes with orbital shaking and used 4 nonselective and 12 selective plates for each strain. In the second, we grew 200 μl cultures in test tubes with orbital shaking and used 12 nonselective and 48 selective plates for each strain. In the third, we grew 200 μl cultures in 96-well deep well microplates with no shaking and used 12 nonselective and 51 selective plates for each strain. In all cases, the liquid cultures in nonselective media were grown for 24 h, and mutant colonies on the selective plates were counted after 48 h of incubation. Colony counts on selective and non-selective plates were used to estimate mutation rates using the rSalvador R package (v1.7) (31). We used its likelihood ratio methods for calculating confidence intervals and the statistical significance of differences between mutation rate estimates for two strains.

Plasmid copy number determination

Cells containing the pTEM-1.D254tag plasmid were revived in 5 ml LB-Cam with overnight growth from frozen stocks. These cultures were diluted 1000-fold into 10 ml of fresh LB-Cam and 1 ml of each culture was harvested when its growth reached exponential phase ($OD_{600} \sim 0.5$). Three cell pellets from different biological replicates were collected for each strain. Mixed plasmid and chromosomal DNA was isolated using the PureLink Genomic DNA Mini Kit (Invitrogen). Total DNA concentrations in these samples were determined using a Qubit Fluorometer (ThermoFisher).

Primer pairs for qPCR were designed to amplify products from either the antibiotic resistance gene (*cat*) in the pTEM-1.D254tag plasmid or the *ftsZ* gene in the *E. coli* chromosome. The *cat* primers were 5'-GTGAGCTGGTGATATGGGATAG and 5'-CCGGAAATCGTCGTGGTATT. The *ftsZ* primers were 5'-GCAAGGTATCGCTGA ACTGA and 5'-CGTAGCCCATCTCAGACATTAC. For each DNA sample, separate amplification reactions with each of the two primer pairs were conducted using Power SYBR Green PCR Master Mix reagents and a ViiA7 Real-Time PCR System (ThermoFisher). These reactions were performed in 96-well PCR plates with a final volume of 15 μ l per well and 500 nM of each primer.

Standard curves for plasmid and chromosomal DNA were used to calculate the absolute concentrations of each type of DNA from Ct values (32). The plasmid standard curve was constructed by amplifying pTEM-1.D254tag plasmid DNA isolated using the PureLink Quick Plasmid Miniprep Kit (Invitrogen). The chromosomal DNA standard curve was made by amplifying *E. coli* DNA isolated using the PureLink Genomic DNA Mini Kit (Invitrogen) from cells without a plasmid. Each curve consisted of a series of 10-fold dilutions of template DNA. Plasmid copy number was calculated by averaging values from technical replicates for each biological replicate and then using the standard curves to estimate its plasmid:chromosome ratio. Graphing and statistical analyses were performed using log2 transformed values of copy number estimates for each biological replicate. The mean copy number estimated for the pTEM-1.D254tag plasmid in the wild-type BW25113 strain was 410.

Scaling of apparent mutation rates with plasmid copy number

We used numerical simulations to examine how the apparent per-cell mutation rates estimated from our fluctuation tests with the pTEM-1.D254tag plasmid would be expected to scale if there were a change in the copy number of this plasmid in an evolved *E. coli* strain. Python scripts for performing the simulations are available online (<https://github.com/barricklab/plasmrs>). In these simulations, a population begins with a single wild-type cell that contains a set number of copies (N_p) of the mutational reporter plasmid. The population growth process is modeled by iteratively picking a random cell from the population to divide and replacing this cell with its two daughter cells. When a cell divides, its complement of plasmids is replicated by iteratively picking a random plasmid from the current population of plasmids in the cell to copy until there are a total of $2N_p$ plasmids. Each time a wild-type plasmid replicates there is a chance

(μ_p) that the new copy is a mutant plasmid that has restored the β -lactamase reading frame. The resulting collection of $2N_p$ plasmids, including any mutant plasmids that may have replicated or have newly arisen during division of this cell, are randomly allocated such that each daughter cell inherits exactly half of the plasmids. After cells divide enough times to reach a final population size (N), the number of mutant cells that would yield colonies on Crb agar (N_m) is counted as the number of cells that contain at least a minimum number of mutant plasmids needed to yield a resistant colony (N_r). After one hundred replicate cultures were simulated for each condition, we estimated the apparent mutation rate per cell (μ) from the observed distribution of mutant colony counts per culture (N_m) and the total number of cells per culture (N) using the rSalvador R package, in the same way that we analyzed experimental data.

We specifically used these simulations to examine how the reduction in plasmid copy number observed in the AER12 PResERV strain with the *polA* mutation would be expected to impact its apparent mutation rate if the evolved strain maintained the same per-plasmid mutation rate as the wild-type BW25113 strain. Because it was not computationally feasible to simulate *E. coli* populations as large as those used in our actual fluctuation tests (with $\sim 5.0 \times 10^7$ to $\sim 2.5 \times 10^9$ cells) we performed simulations that scaled the apparent per-cell mutation rate (μ) upward such that equivalent values of the expected number of antibiotic resistant mutant cells per culture ($m = N\mu$) were reached with smaller values of N . We performed four different sets of simulations matching the μ and m parameter combinations for the wild-type BW25113 strain that we observed in fluctuation tests in the four different experimental blocks shown in the results section that compared the plasmid mutation rates of the wild-type and AER12 strains. These combinations were: $\mu = 4.69 \times 10^{-8}$ and $m = 104$; $\mu = 2.89 \times 10^{-7}$ and $m = 27.7$; $\mu = 2.36 \times 10^{-8}$ and $m = 1.53$; and $\mu = 2.14 \times 10^{-7}$ and $m = 20.5$.

For each set of simulations, we first determined an underlying per-plasmid mutation rate (μ_p) that matched the experimentally measured apparent per-cell mutation rate (μ) to within 5% by performing a series of simulations using $N = 3.2 \times 10^4$ cells and $N_p = 410$ plasmids per cell. We repeated this procedure for each of three different values of N_r (1, 3, and 10). Then we performed five new simulations with each μ_p value corresponding to a different N_r across five different values of N (10^4 , 1.8×10^4 , 3.2×10^4 , 5.6×10^4 , and 10^5). Finally, we performed new simulations at these same fifteen combinations of μ_p , N , and N_r with $N_p = 185$ and all other parameters left unchanged, in order to determine what the apparent mutation rate would have been in the fluctuation tests if plasmid copy number had decreased without any change in the underlying plasmid mutation rate. For each set of five pairs of simulations differing only in N_p , we calculated R , the ratio of the apparent mutation rate for $N_p = 185$ to that for $N_p = 410$. We found no significant dependence of the logarithm of R on the logarithm of N across the range of values tested in any of these sets ($P > 0.05$, Bonferroni-corrected test for a nonzero slope in a linear regression model), justifying our inversely proportional rescaling of N and μ for the purpose of making

the simulations feasible. We further found that the values of the logarithm of R within each of the three sets of simulations at a fixed N_T that varied μ and N were indistinguishable from one another ($P > 0.05$, Kruskal–Wallis tests). Therefore, we report only the overall mean of the log-transformed R values for each set of 20 simulations at the same N_T and a bootstrap confidence interval on this statistic constructed from 100 000 resampled sets.

RESULTS

PResERV experiment with a ColE1 plasmid in *E. coli*

We applied PResERV to *E. coli* K-12 strain BW25113 (20) transformed with pSKO4, a high-copy-number pBR322 plasmid (ColE1-type origin) that encodes GFP under control of an inducible promoter (21). GFP expression is a generic proxy for a costly engineered function in this scenario. A UV-mutagenized library of cells containing pSKO4 was propagated through daily 1:1000 serial transfers in the presence of antibiotic selection for plasmid retention. Under these conditions, cells with mutations in pSKO4 that inactivate or reduce costly GFP expression evolve, outcompete fully fluorescent cells, and constitute a majority of the population within a few days (21). GFP fluorescence of cells in the PResERV population was periodically monitored by flow cytometry using the red fluorescent nucleic acid dye SYTO17 as a counterstain to improve detection of cells with a low GFP signal. When 75% or more of the cells exhibited reduced GFP signal, cell sorting was used to isolate $\sim 10^5$ cells that remained at least as fluorescent as the ancestor to continue the population. We subjected this population to a total of 8 sorts spread throughout 30 regrowth cycles (Figure 2A).

Six *E. coli* clones designated AER7–AER12 were isolated from the final population for further characterization. Five of these maintained more fully-fluorescent cells for more cell doublings than the unevolved wild-type strain with the wild-type pSKO4 plasmid (Figure 2B). Mutations in the plasmid, the *E. coli* chromosome, or both could have been responsible for these improvements. To determine which was the case, we cured these cells of their plasmids and retransformed them with the wild-type pSKO4 plasmid, and we also isolated plasmids from each of the evolved strains and transformed them into unevolved wild-type *E. coli* cells. For four of these strains (AER7, AER8, AER9 and AER12), the improvement in the evolutionary lifetime of GFP expression appeared to be mainly due to mutations in the *E. coli* chromosome rather than mutations in the pSKO4 plasmid (Figure 2C).

Mutations in PResERV strains

We sequenced the genomes of these four evolved clones to understand the genetic basis of their improved reliability (Figure 3). In agreement with the re-transformation tests, no mutations were found in the pSKO4 plasmid in any of these strains. Each contained from four to ten mutations in the *E. coli* chromosome. These mutations could theoretically lead to the improved maintenance of GFP expression that we observed by reducing the burden of GFP expression from the plasmid or by reducing the rate at which mu-

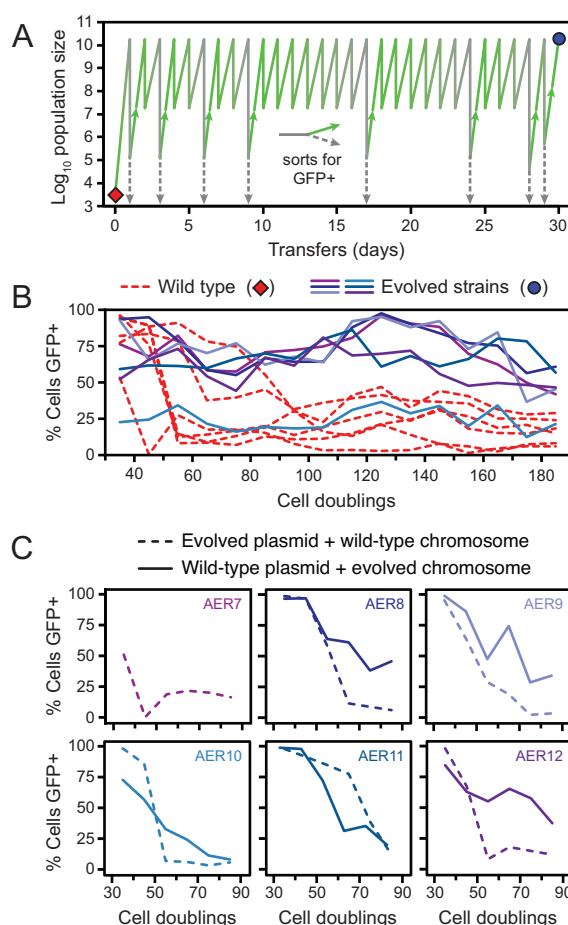


Figure 2. PResERV applied to an *E. coli* plasmid. (A) Propagation and sorting regimen used to perform PResERV on an *E. coli* population in which GFP was expressed from plasmid pSKO4, a high-copy plasmid with a pBR322 origin of replication. The red diamond denotes the wild-type strain that was UV-mutagenized prior to beginning PResERV. Dashed gray and solid green bifurcating lines show when the population was sorted to retain fully fluorescent cells (GFP+). The blue circle indicates when fluorescent clones were isolated and sequenced. (B) Populations initiated from six different clones isolated at the end of the PResERV evolution experiment (solid lines in shades of blue and purple) were allowed to evolve alongside six replicates of the non-mutagenized, wild-type *E. coli* strain-plasmid combination (red dashed lines). Cells were considered GFP+ if they maintained a fluorescent intensity as measured by flow cytometry that was above a threshold level that was kept constant across all tested strains. (C) For each evolved PResERV strain, its plasmid was isolated and transformed into the wild-type strain containing no plasmid (dashed lines), and the evolved strain was cured of its plasmid and re-transformed with the wild-type pSKO4 plasmid (solid lines). Populations initiated from these strains were propagated and monitored as in B. The stability of AER7 re-transformed with the wild-type plasmid was not determined because of difficulty curing the evolved plasmid from this strain. In panels B and C, the same colors are used for each PResERV strain. In both experiments, the percentage of GFP+ cells was measured by flow cytometry after the first 35 cell doublings, corresponding to growth of the initial culture from a single cell on an agar plate, and then every 10 cell doublings afterward, corresponding to regrowth after daily subculturing steps that used a 1:1000 dilution into fresh media.

tations that inactivate GFP arise. Therefore, we examined the lists of mutations in these strains to see if they hit any genes known to be involved in these processes.

Position	Mutation	Evolved Strains			Annotation	Gene	Description
63,992	G→A	AER7	AER8	AER9	H597Y (CAT→IAT)	<i>polB</i> ←	DNA polymerase II
64,108	G→T	AER7	AER8	AER9	S558* (TCG→TAG)	<i>polB</i> ←	DNA polymerase II
516,812	A→G			AER12	E314G (GAA→GGA)	<i>ybbP</i> →	Putative ABC transporter permease
645,462	A→T	AER7	AER8	AER9	L162H (CTC→CAC)	<i>citE</i> ←	Citrate lyase (beta) subunit
1,139,159	A→G	AER7	AER8	AER9	L222S (TTA→TCA)	<i>rne</i> ←	Ribonuclease E
1,240,803	2 bp→AA	AER7	AER8	AER9	coding (188-189/441 nt)	<i>ycgY</i> →	Uncharacterized protein
1,444,811	G→A	AER7	AER8	AER9	L266L (CTG→ITG)	<i>tynA</i> ←	Tyramine oxidase, Cu-requiring
1,771,323	G→A	AER7	AER8	AER9	D494N (GAT→AAT)	<i>ydlF</i> →	Putative acetyl-CoA:acetoacetyl-CoA transferase
2,196,845	G→A			AER12	G1030R (GGG→AGG)	<i>yehl</i> →	DUF4132-containing protein
3,368,284	G→A	AER7	AER8	AER9	L19L (CTG→CTA)	<i>dcuD</i> →	Putative transporter
3,452,455	C→T			AER12	S253F (TCC→TTC)	<i>gspE</i> →	General secretory pathway component, cryptic
4,042,525	C→T			AER12	H734Y (CAC→IAC)	<i>polA</i> →	DNA Polymerase I
4,166,148	A→G			AER9	I93V (ATC→GTC)	<i>tufB</i> →	Translation elongation factor EF-Tu 2
4,327,401	IS1 (-) +8 bp			AER9	coding (339-346/762 nt)	<i>adiY</i> ←	Transcriptional activator, <i>adi</i> system

Figure 3. Mutations in PResERV strains. The genomes of four evolved strains were re-sequenced to identify mutations that accumulated during the directed evolution procedure. The position column shows the coordinate of the first affected base pair defined relative to the *E. coli* K-12 BW25113 genome (GenBank: CP009273.1). The mutation column shows base changes on the top strand of the genome, except for the *IS1* element in AER9 that inserted in the reverse direction and duplicated bases 4,327,401–4,327,408 at the target site on each side of the new IS copy. The annotation column shows the amino acid changes and codon changes caused by single-base substitutions or the locations of bases affected within a gene for other mutations. The gene column includes arrows showing the genomic strand on which each mutated gene is located.

Two of these strains (AER7 and AER8) had eight identical mutations while a third strain (AER9) had these same eight mutations plus two additional ones. All three shared mutations in *polB* and *rne* that were candidates for affecting evolutionary stability. PolB (Pol II) is a stress-induced DNA polymerase that participates in translesion synthesis and nucleotide excision repair. The *polB* gene sustained two mutations in these PResERV strains: a missense mutation (H597Y) and a nonsense mutation earlier in the reading frame (S558*). The full-length PolB protein is 783 amino acids in length, and the stop codon mutation truncates the protein within its catalytic core (33). Presumably, this results in a complete loss of Pol II activity in the mutant. Deletion of *polB* in the clean-genome *E. coli* strain MDS42 has been shown to lead to ~30% lower chromosomal mutation rates (12). The *rne* gene (encoding RNase E) contains a missense mutation (L222S) in all three strains. RNase E regulates the copy number of ColE1 origin plasmids in *E. coli* by processing the RNA I antisense regulator of the RNA II replication primer (34,35). Cells defective in *rne* accumulate higher levels of RNA I and have reduced plasmid copy number (36). The site of the PResERV mutation is within the RNaseH-like domain of RNase E, which is involved in determining its RNA substrate selectivity, but its effect on the activity of this enzyme is not clear from the structural context (37).

The fourth sequenced strain (AER12) had a completely different set of four mutations, which included a missense mutation in *polA* (H734Y). PolA (Pol I) is the DNA polymerase that is utilized primarily for filling gaps during lagging strand synthesis and in DNA repair in *E. coli*. It is also responsible for extending the primer derived from RNA II during replication of ColE1-type plasmids (35). Antimutator variants of PolA that lower the frequencies of mutations observed on a reporter plasmid have been identified previously by screening a sequence library created by mutagenizing an exo^- PolA variant lacking 3'→5' exonuclease proof-reading activity (38). The exact same substitution that we observed (H734Y) was found among the 592 active polymerase variants characterized in that study, but the effects

of this specific mutation on polymerase function were not reported. H734 is located near the phosphate groups of the dNTP substrate when it is bound to the Klenow fragment of Pol I (39), indicating that the PResERV mutation may have an effect on nucleotide binding.

Evolutionary stability and mutation rates in PResERV and reconstructed strains

To test whether these three mutations contributed to the increased evolutionary reliability of the PResERV strains, we tested *E. coli* strains in which we reverted the evolved alleles back to their wild-type sequences. We then propagated replicate populations of wild-type *E. coli*, two focal evolved clones (AER12, the strain with the *polA* mutation; and AER8, one of the three strains containing mutations in *polB* and *rne*), and four revertant strains (one for each mutation and also a strain in which *polB* and *rne* were both reverted) under the same conditions as the initial evolution experiment and monitored the loss of fluorescence over the course of ~100 cell doublings (Figure 4). Mutations in *polA* and *rne* appeared to be responsible for most or all of the improved stability, as reverting these mutations reduced the evolutionary lifetime of GFP expression back to a level similar to that observed in the wild-type strain. In contrast, reverting the *polB* mutation alone or reverting it in a strain that also had the *rne* mutation reverted did not appreciably affect how rapidly GFP expression decayed.

We next used Luria-Delbrück fluctuation tests (28) to determine if the increase in evolutionary reliability in these strains was associated with a decrease in mutation rates. We first measured the rates of point mutations that reverted a stop codon in a β -lactamase gene cloned into another pBR322-based plasmid designated pTEM-1.D254tag (22). Mutation rates to carbenicillin resistance, which requires mutating this stop codon to a sense codon, were significantly lower in each of the two focal evolved clones compared to wild type in multiple experiments (Figure 5). In agreement with the changes in the evolutionary stability of

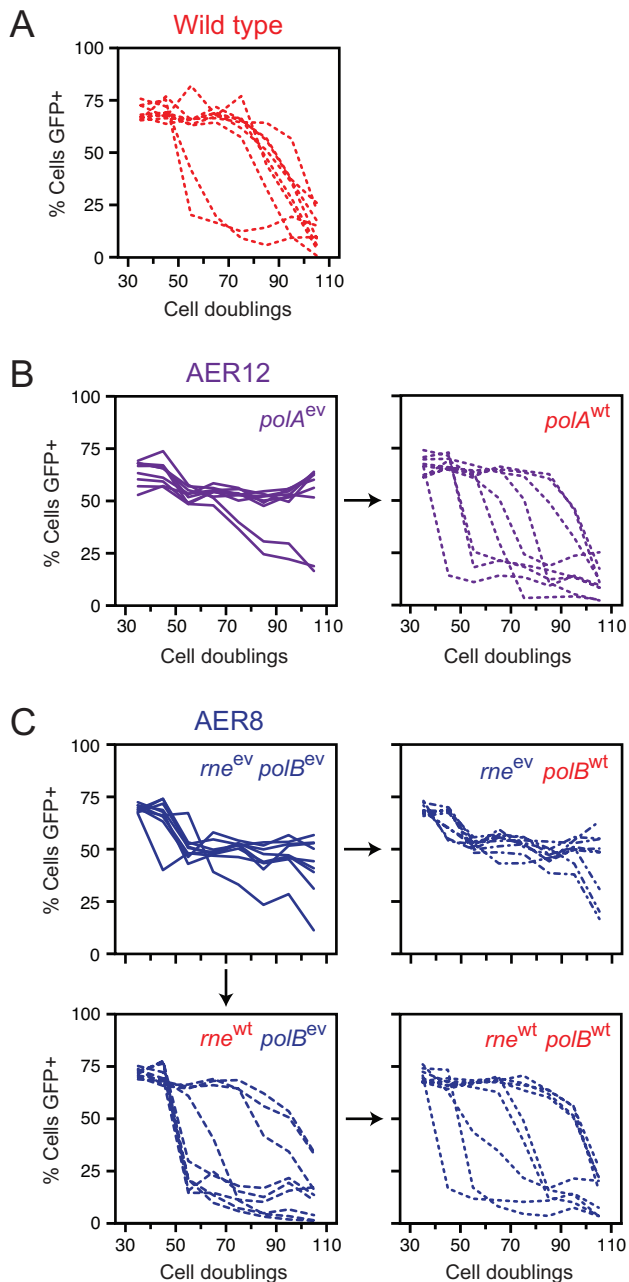


Figure 4. Evolutionary stability and mutation rates in PResERV and reconstructed strains. (A) Wild-type strain. (B) Evolved strain AER12 and a derived strain with its evolved *polA* allele reverted to the wild-type sequence. (C) Evolved strain AER8 and derived strains with its evolved *polB* and *rne* alleles reverted, singly and in combination. In each panel, the strains being tested were first transformed with plasmid pSKO4. Then, nine independent populations were initiated from single colonies of each strain. The prevalence of GFP-expressing cells within each population was monitored by flow cytometry over multiple daily serial transfers. Evolved strains are shown with solid lines. Dashed lines indicate that a strain contains one or more wild-type alleles, as indicated in red type. Measurements were made after the first 35 cell doublings and then every 10 cell doublings thereafter.

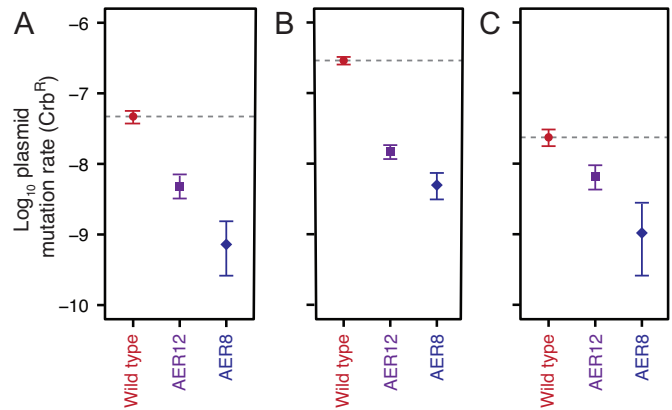


Figure 5. Plasmid mutation rates in PResERV strains. Mutation rates to carbenicillin resistance (Crb^R) due to reversion of a stop codon in the TEM-1.D254tag plasmid were measured using Luria-Delbrück fluctuation tests. Wild type and the two focal evolved strains were compared in three experiments under different growth conditions: (A) in 1 ml cultures in test tubes incubated with orbital shaking (B) in 200 μ l test cultures in test tubes incubated with orbital shaking, and (C) in 200 μ l cultures incubated in 96-well microplates with no shaking. Each experiment included the wild-type *E. coli* strain for comparison (dashed horizontal lines). Error bars are 95% confidence intervals.

GFP expression, reversion of either the *polA* or *rne* mutation raised the mutation rate to that of the wild-type *E. coli* strain, and reversion of the *polB* mutation had no detectable effect on the mutation rate (Figure 6A). We also measured mutation rates in two further sets of fluctuation tests, selecting either for resistance to rifampicin or to D-cycloserine, which require mutations in genes located on the *E. coli* chromosome in both cases. We did not find any significant improvements versus the wild-type strain in these assays (Figure 6B). Thus, it appears that PResERV discovered *E. coli* mutants that primarily display lower plasmid mutation rates, with much less of an effect, if any, on mutation rates in the chromosome.

Plasmid copy number and GFP fluorescence in evolved strains

Given previous reports of lower plasmid copy number when *rne* function is reduced in a temperature-sensitive mutant (36) and that *polA* antimutator mutations can lead to slower rates of DNA replication (38), we were concerned that a decrease in plasmid copy number in the PResERV evolved cells could give a false signal of improvement in our two assays. First, having fewer plasmids per cell would lower the GFP expression burden and thereby increase the number of cell doublings it would take for new cells that arise with mutated plasmids to outcompete cells with wild-type plasmids (i.e., it would increase the apparent evolutionary stability). Second, with fewer plasmids per cell there would be a smaller chance that any given cell would experience a mutation in one of its plasmids that would lead to resistance in the β -lactamase stop codon reversion assay (i.e., it would reduce the apparent mutation rate per cell). Therefore, we measured the copy number of the pTEM-1.D254tag plasmid in the two focal evolved clones and four reconstructed strains using qPCR (Figure 7A), and we also examined the

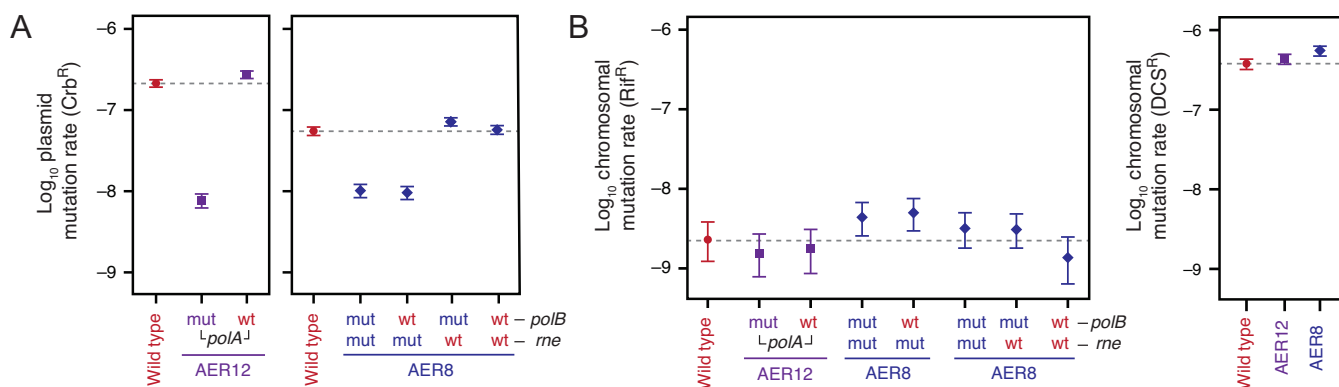


Figure 6. Plasmid and chromosomal mutation rates in PResERV and reconstructed strains. (A) Mutation rates to carbenicillin resistance (Crb^R) due to reversion of a stop codon in the TEM-1.D254tag plasmid were measured using Luria-Delbrück fluctuation tests. Wild-type and evolved strains with mutant *polA*, *polB* and *rne* alleles (mut) reverted to wild-type sequences (wt), individually or in combination, were examined. Strains related to the evolved clone with a *polA* mutation (AER12) were tested in one experiment (left panel), and strains related to the evolved clone with *polB* and *rne* mutations (AER8) were tested in another experiment (right panel). (B) Mutation rates to rifampicin resistance (Rif^R) (left panel) and D-cycloserine resistance (DCS^R) (right panel) were measured using Luria-Delbrück fluctuation tests. Both of these resistance phenotypes require mutations in genes located on the *E. coli* chromosome. Each experiment included wild-type *E. coli* for comparison (dashed horizontal lines). Error bars are 95% confidence intervals.

per-cell GFP fluorescence from the pSKO4 plasmid in each strain (Figure 7B).

We found that the *polA* mutation did reduce plasmid copy number somewhat. The evolved *polA* strain (AER12) had marginally fewer plasmids per chromosomal DNA copy when compared to wild-type ($P = 0.0666$, one-tailed *t*-test on log-transformed values) and also exhibited reduced GFP fluorescent intensity ($P = 0.0142$, one-tailed *t*-test on log-transformed GFP⁺ subpopulation medians). Interestingly, other mutations in the evolved strain appeared to counteract the effects of the *polA* mutation, as reverting just this mutation to the wild-type allele increased both copy number ($P = 0.0011$) and GFP intensity ($P = 0.0003$). GFP signal ($P = 0.0108$) and perhaps copy number ($P = 0.0893$) were even greater in this *polA* revertant that still contained all other evolved mutations than they were in the original wild-type *E. coli* strain. Overall, these results suggest that there is a trade-off in the evolved *polA* mutant between plasmid copy number and mutation rate.

In contrast, plasmid copy number did not vary when comparing wild-type, the evolved strain with *rne* and *polB* mutations (AER8), and all three reconstructed strains reverting those mutations singly and in combination (one-way ANOVA, $F_{4,10} = 0.439$, $P = 0.778$). Here, too, there was evidence that other mutations in this evolved strain may have increased GFP intensity, as all four of the AER8-derived strains considered together had indistinguishable fluorescence intensities (one-way ANOVA, $F_{3,32} = 1.153$, $P = 0.343$) that were, as a group, significantly greater than that of the wild-type ($P = 0.0112$, one-tailed *t*-test). Therefore, the evolved *rne* allele reduced plasmid mutation rates with no detectable trade-off in terms of plasmid copy number or gene expression.

To determine whether reduced plasmid copy number in the AER12 strain containing the *polA* mutation could explain the reduction of 20- to 60-fold in the plasmid mutation rate measured for this strain (Figures 5 and 6A), we performed numerical simulations of the growth of cell pop-

ulations that included multicopy plasmid replication, mutation, and segregation (see Materials and Methods). Our qPCR results indicate that the copy number of the pTEM-1.D254tag plasmid was reduced from ~410 plasmids per *E. coli* chromosome in the wild-type strain to ~185 in AER12. We simulated the results of fluctuation tests with 410 mutational reporter plasmids per cell and with other parameters chosen to match the observed numbers of mutant cells per culture for each of our four different sets of mutation rate measurements comparing wild type and AER12. Then, we performed a new set of simulations with the same parameters but reducing the plasmid copy number to 185 to determine by how much this would reduce the apparent plasmid mutation rate inferred from the Luria-Delbrück analysis. We found that a reduction in copy number of 2.2-fold is expected to yield a roughly proportional change in the apparent mutation rate. The result varies slightly if one changes how many mutant plasmids in a cell are necessary for it to give rise to a mutant colony, a parameter that is unknown in our system but is likely one or a just a few plasmids with restored β -lactamase copies per cell. For simulations requiring one mutant plasmid per cell we predicted a reduction of 2.00-fold (1.93–2.07, 95% bootstrap confidence interval, see Methods) in the apparent mutation rate in AER12. For three copies the reduction was 2.12-fold (2.05–2.20), and for ten copies it was 2.37-fold (2.24–2.51). We conclude that the reduction in plasmid copy number in AER12 is not sufficient to explain a majority of the reduction in plasmid mutation rates in that strain.

DISCUSSION

Mutation rates in microbial populations reflect a dynamic balance between different evolutionary forces and inherent constraints on organisms that have DNA as their genetic material. On one side, there is a universal selection pressure to minimize mutation rates because most new mutations are far more likely to be deleterious to fitness than beneficial (40–42). This risk associated with deleterious mutations

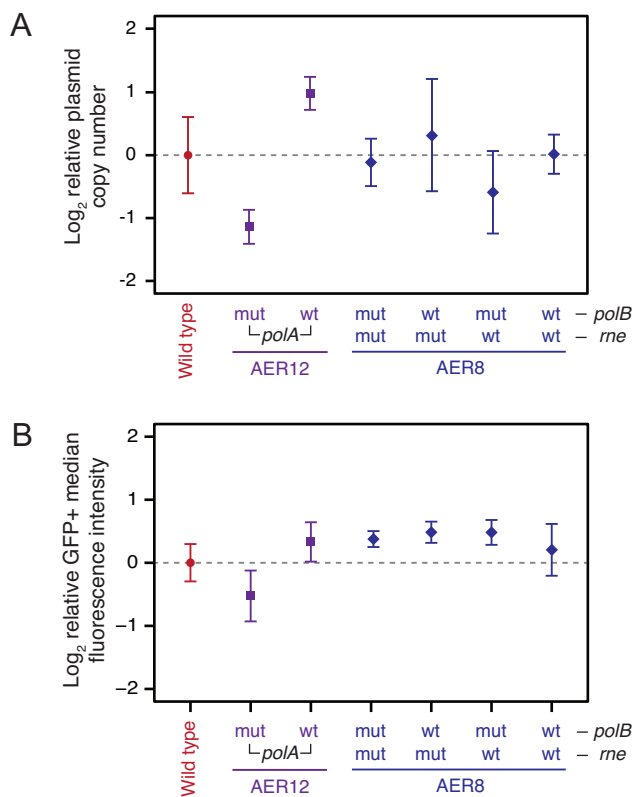


Figure 7. Plasmid copy number and GFP fluorescence in PResERV and reconstructed strains. (A) Plasmid copy number for wild-type, evolved, and reconstructed strains determined by qPCR. Wild-type and evolved strains with mutant *polA*, *polB* and *rne* alleles (mut) reverted to wild-type alleles (wt), individually or in combination, were tested. The horizontal dashed line indicates the estimated copy number in the wild-type strain. Error bars show the standard error of the mean on log-transformed values from three biological replicates. (B) Initial GFP fluorescence of wild-type, evolved, and reconstructed strains as measured by flow cytometry. The median per-cell fluorescence intensity of the GFP+ subpopulation of cells was determined for each of nine replicate cultures immediately after outgrowth in liquid culture (after ~35 cell doublings). The graphed values are the log-averaged values of these medians. Error bars are 95% confidence intervals calculated assuming the logarithms of the medians are normally distributed. The horizontal dashed line shows the value for the wild-type strain.

contributes to genetic load. That is, there is a fitness cost associated with a given mutation rate in terms of the fraction of an organism's offspring that will experience lethal or deleterious mutations that lead to their immediate or eventual extinction. If selection to reduce genetic load were the only evolutionary force in play, then a mutation rate of zero would be optimal. On the other side of the balance, there are at least three different forces or barriers that will prevent the evolution of lower mutation rates past a certain point in microbial populations: second-order selection for evolvability and limits imposed by the strength of genetic drift and by physiological constraints.

New mutations may be a bad bet on average, but they do—more rarely—generate beneficial genetic diversity that is necessary to fuel adaptive evolution. Thus, under certain circumstances, mutation rates can evolve to rebalance the potential for beneficial mutations against the risk of dele-

terious mutations (43–45). For example, laboratory populations of bacteria and yeast often evolve hypermutation (elevated mutation rates) (46,47) because they experience strong and constant selection pressures that can indirectly favor more evolvable lineages that have a greater chance of sampling rare adaptive mutations (48,49). The simplicity of laboratory environments compared to nature also means that there is less of a deleterious genetic load associated with evolving a high mutation rate in these experiments. Many mutations that would be lethal under other circumstances (e.g., that disrupt pathways for utilizing alternative nutrients or stress responses for contingencies that are never experienced) will be effectively neutral in the comparatively monotonous environments of these experiments (50,51). For similar reasons, hypermutators also commonly evolve in populations of bacteria during the long-term progression of chronic infections treated with antibiotics, such as for *Pseudomonas aeruginosa* in the lungs of cystic fibrosis patients (52–54).

The molecular basis for the evolution of bacterial hypermutators in the laboratory and in the clinic is usually straightforward. Mutations disrupt major housecleaning enzymes (e.g. *mutT*) or DNA repair pathways (e.g. *mutS*), often leading to an increase in point mutations with a characteristic base substitution spectrum (46,47,53,55). Interestingly, some experiments have shown that experimental populations that are started with or that spontaneously evolved hypermutation can subsequently evolve reduced mutation rates (56–59). This can occur when hypermutator populations are propagated through severe population bottlenecks, which exacerbates the genetic load associated with a given mutation rate while reducing the chances of sampling adaptive mutations (56,57). Reduced mutation rates have also been observed to evolve after a population becomes well-adapted to its environment and opportunities for beneficial mutations diminish relative to the risk of deleterious mutations (58,59). When the molecular mechanisms have been examined in detail, the evolution of lower mutation rates in these experiments has been found to occur through new mutations that partially compensate in some way for the defect in the hypermutator (58), or by exact reversion of the mutation responsible for hypermutation (59). The evolution of cells with a mutation rate that is lower than that of the ancestral, wild-type microbe has not been observed in these experiments.

Despite the potential for evolving hypermutation, wild-type microbes isolated from nature almost always have very low mutation rates (60). The uniformity of these baseline rates across many species isolated from diverse environments suggests that a different balance of evolutionary forces than the one between genetic load and the potential for beneficial mutations is normally responsible for setting mutation rates in nature. If one assumes that wild-type microbes are already well-adapted to the combinations of complex and varying environments that they regularly experience, then there may be little or no benefit possible from further mutations. Under these circumstances there will only be selection to minimize genetic load. What then would set the lower bound on mutation rates? Because baseline mutation rates have been found to scale inversely with effective population sizes across many organisms, it

has been argued that genetic drift is the dominant force opposing selection for even lower mutation rates (61,62). This 'drift barrier' arises because once the mutation rate is sufficiently low, the very small and indirect marginal benefit for a new mutation that leads to an even lower mutation rate becomes so insignificant that it looks effectively neutral to natural selection. That is, natural selection does not have the power to favor this hypothetical new antimutator allele such that it will reliably increase in frequency on its merits and eventually win out in the population.

Another potential barrier to the evolution of lower mutation rates considers the molecular biology of DNA replication and repair. Biochemical and genetic studies of bacteria over the past several decades have mapped a complex suite of pathways dedicated to maintaining genome integrity via overlapping and redundant mechanisms (63). There is a direct fitness cost to a cell for expressing any protein (6,7), and there may be other fitness costs associated with increased surveillance for DNA damage and enforcing replication fidelity, such as off-target promiscuous activities of housecleaning enzymes (64) or slower rates of DNA polymerization in exchange for increased proofreading (38). Thus, there must be a point at which the direct fitness cost of evolving additional molecular machinery outweighs the diminishing indirect benefit of further reducing the genetic load from deleterious mutations. However, there is no evidence that this physiological lower limit on mutation rates has been reached in natural microbes. Antimutators with reduced point mutation rates have been identified by genome-wide genetic screens (65), targeted mutagenesis (38,66,67), gene disruption (12), and gene overexpression (68). Some of these antimutators do exhibit growth trade-offs (38,65), but some do not appear to have any deleterious side-effects (12,38,69), at least in the laboratory environments in which they have been tested. Yet, due to the intrinsic instability of DNA, which can be chemically damaged or miscopied in an dizzying variety of ways (63), there must exist some finite, non-zero mutation rate at which this physiological genetic stability limit is reached.

In this study, we show that imposing artificially strong selection for bacterial cells that are less likely to give rise to mutations in a reporter gene on a plasmid can overcome the selection pressures and other barriers that normally oppose the evolution of reduced mutation rates. Specifically, we developed and used a Periodic Reselection for Evolutionarily Reliable Variants (PResERV) directed evolution approach to isolate *E. coli* host strains with mutations in their chromosomes that lead to lower-than-natural mutation rates in genes encoded on high-copy vectors such as pUC and pBR322 from the ColE1 plasmid incompatibility group. We sequenced the genomes of four improved PResERV strains to better understand the molecular basis for the improvements and found mutations in three key genes (*polA*, *polB*, and *rne*). Then, we characterized the effects these mutations had on the evolutionary stability of burdensome GFP expression, on plasmid and chromosomal mutation rates, and on plasmid copy number and gene expression in the evolved *E. coli* strains.

One PResERV strain had a mutation in DNA polymerase I (*polA*) that reduced plasmid mutation rates by ~30-fold. Pol I is required for the normal replication of ColE1 plas-

mids in *E. coli* (35). Both hypermutator and antimutator variants of this polymerase have been shown to affect the fidelity of DNA replication (38,70,71), so it was not surprising that we identified a mutation in *polA* that increased genetic stability. In fact, the exact amino acid substitution in Pol I recovered by PResERV (H734Y) was found previously in a library of mutagenized *exo*⁻ Pol I sequences (38). Though the specific effects of the H734Y mutation were not reported individually in that study, many of the mutagenized Pol I variants were antimutators. Their improved fidelity was attributed to increased selectivity for the incoming nucleotide. The *polA* mutation that we recovered may act similarly, as it is located close to the binding site for the incoming dNTP (39). Thus, our identification of a *polA* mutation was in essence a positive control that PResERV could successfully isolate generalizable antimutator alleles, as opposed to mutations that increased the evolutionary stability of just the pSKO4 plasmid in an idiosyncratic way (e.g., by reducing the cost of expressing GFP).

Pol I has an outsized role in replicating ColE1 plasmids compared to its relatively minor roles in lagging-strand synthesis and DNA repair during normal replication of the *E. coli* chromosome. Pol I initiates plasmid DNA replication by extending a primer that is processed from the RNA II transcript derived from the plasmid origin. In the canonical model of plasmid replication, the DNA polymerase holoenzyme involved in chromosomal replication, which utilizes Pol III, takes over plasmid replication after Pol I extends the primer by ~400–500 nucleotides (35). However, there is evidence that Pol I also replicates other portions of ColE1 plasmids, at least some of the time. When a hypermutator Pol I variant was expressed in *E. coli* cells, plasmid mutation rates were most elevated close to the origin of replication, within the expected 400–500 bp window, but mutation rates were still much higher than normal within a region extending at least 3700 bp downstream of the origin (70).

The reading frame for GFP on the pSKO4 plasmid used in PResERV is located from ~350 to ~750 base pairs downstream of the origin (as measured from the typical end of the RNA II transcript after nucleolytic processing). This places at least part of the GFP gene within the region known to be heavily replicated by Pol I, meaning that we would expect to observe a particularly strong effect of a *polA* antimutator allele during PResERV and in subsequent decay experiments in which we monitored the evolutionary stability of GFP expression over multiple growth cycles. In contrast, the stop codon in β -lactamase on the pTEM-1.D254tag reporter plasmid that we used in fluctuation assays to measure plasmid mutation rates is located ~3000 bp downstream of the end of RNA II. We still see greatly reduced mutation rates in this reporter, corroborating the prior observations that changes in Pol I fidelity impact mutation rates across most or all of the sequence of a ColE1-type plasmid. Also in broad agreement with previous studies, which report that Pol I variants have less of an effect on chromosomal mutation rates compared to plasmid mutation rates (70,71), we found no significant difference in chromosomal mutation rates from the PResERV *polA* antimutator allele.

Our overall goal is to construct an *E. coli* cell that is more robust against unplanned evolution to serve as an improved chassis for synthetic biology and biotechnology ap-

plications. For this purpose, the most useful antimutator alleles are those that increase the genetic stability of an engineered DNA sequence in a host cell without any trade-offs in other desirable traits. In one important respect, the PResERV *polA* mutation and many other *polA* antimutator alleles pass this test: they do not negatively affect *E. coli* growth rates (71). However, the PResERV allele does exhibit a significant trade-off that diminishes its potential utility. Copy number of the pTEM-1.D254tag plasmid used to measure mutation rates was reduced by ~55% in strains with the evolved *polA* allele. This decrease is consistent with a reduction in GFP expression from the pSKO4 plasmid when the evolved *polA* mutation was present in a strain. We used numerical simulations to show that this slight reduction in plasmid copy number can explain only ~2-fold of the ~30-fold reduction in plasmid mutation rates we observed in the evolved strain. Still, as ColE1 plasmids are widely used for cloning and protein overexpression, where maximal yield of plasmid DNA or a protein encoded on the plasmid is the primary goal, this trade-off of much lower plasmid mutation rates at the expense of reduced plasmid copy number would not be favorable, on balance, for many biotechnology applications.

Other mutations in Pol I have previously been found to reduce the copy number of high-copy ColE1 plasmid variants (72). It has been hypothesized that they might have this effect by decreasing the frequency of initiation of DNA synthesis from the RNA II primer, by reducing the speed of DNA polymerization, or by some combination of the two. In biochemical assays, some of the exo^- Pol I antimutator variants have been reported to have significantly reduced rates of polymerization (38). They would presumably also reduce ColE1 plasmid copy number, though this has not been tested. However, other antimutator Pol I variants apparently retain wild-type enzyme activity (38). It is possible that these *polA* mutations do not exhibit any trade-off in plasmid copy number and would ultimately be more useful than the PResERV allele for constructing an improved *E. coli* host strain.

The three other sequenced PResERV strains all shared mutations in two genes that also have known roles in DNA replication fidelity and regulation of plasmid replication: DNA polymerase II (*polB*) and RNase E (*rne*). Pol II is a repair polymerase induced by the SOS and RpoS responses (73,74). We observed a mutation that creates a premature stop codon in *polB* at amino acid 558. This truncation likely results in a completely inactivated enzyme. There is also a second point mutation later in the *polB* reading frame in these strains that would result in an amino acid substitution (H597Y) if the nonsense mutation were not present. The occurrence of two nearby mutations in the same gene is probably due to our use of UV mutagenesis to create initial genetic diversity in the *E. coli* population at the beginning of PResERV, as clustered mutations can result from long-patch repair of UV damage (75). Despite the fact that UV damage induces the SOS response, we do not expect that these mutations in *polB* were favored due to any direct connection to the mutagenesis procedure. Loss of Pol II function does not appreciably affect cell survival or the overall level of mutagenesis after UV exposure unless other DNA repair pathways are also inactivated (76).

The connection between Pol II activity and *E. coli* mutation rates has multiple facets. On one hand, Pol II can act as a high-fidelity alternative to the other stress-induced polymerases (Pol IV and Pol V). As a consequence, inactivation of *polB* increases the incidence of point mutations arising from the repair of DNA double-strand breaks that occur during long-term carbon starvation (77,78). However, Pol II also participates in mutagenic translesion synthesis pathways that repair other types of DNA damage in an error-prone manner. Accordingly, deletion of Pol II has been reported to have the opposite effect and reduce mutagenesis associated with certain DNA base adducts (79,80) and in cells exposed to antibiotics that can cause DNA damage (81). The mutagenic effect of Pol II activity appears to dominate in *E. coli* cells growing under standard laboratory culture conditions, as incorporating a deletion of *polB* into an engineered reduced-mutation variant of the MDS42 clean-genome *E. coli* strain lowered chromosomal mutation rates by ~30% (12).

Despite these connections between Pol II and mutagenesis, the mutant *polB* gene sequence from PResERV was not associated with a significant change in mutation rates in our assays when we reverted it to the wild-type sequence in the evolved strain or in a strain in which the *rne* mutation found in the same evolved strains was also reverted. It is possible that this is due to epistatic interactions with the other five mutations common to this set of evolved strains, although none of these mutations affect genes with an obvious connection to DNA replication or repair processes. One or both of the Pol II mutations could have been present immediately after UV mutagenesis at the beginning of PResERV and reduced mutation rates in this context. Then, subsequent mutations that arose in this winning lineage during the regrowth cycles of PResERV might have overshadowed the effect of the Pol II mutations by making them redundant. However, we believe this is unlikely to be the only explanation, as the evolved *rne* allele on its own seems to explain all of the reduction in mutation rates, whether or not the evolved *polB* sequence is present. It is also possible that some aspect of the environment experienced by cells during PResERV but not during the mutation rate assays introduced a stress that favored *polB* inactivation. For example, the PResERV cultures were often interrupted by diluting them into water and processing them through a cell sorter before the next growth cycle. In any case, it is clear that knockout of Pol II is not as effective at reducing mutation rates under normal growth conditions as the mutation in RNase E that is present in the same strains.

RNase E is an endoribonuclease with global roles in RNA maturation, processing, and decay (82). It is involved in tRNA, rRNA, and small RNA processing and has been reported to initiate the decay of ~60% of *E. coli* mRNAs (83,84). RNase E is also specifically involved in controlling the copy number of ColE1 plasmids (34,35). It does so by cleaving the regulatory antisense RNA I transcript at a specific site, which converts it into an inactive form that cannot bind to and inhibit processing of RNA II into a productive primer. RNase E is an essential gene, but eliminating its expression using a temperature-sensitive mutant has been shown to reduce plasmid copy number, as is expected from the resulting increase in levels of the active form of the

RNA I inhibitor (36). Despite this connection to the regulation of initiation of ColE1 plasmid replication and unlike the *polA* allele in the other PResERV strain, the evolved *rne* allele did not significantly change plasmid copy number, as measured using qPCR for the pTEM-1.D254tag mutational reporter plasmid or in terms of the fluorescence output from the pSKO4 plasmid. This *rne* allele demonstrates an advantage of using the PResERV directed evolution approach. Although RNase E has a known role in ColE1 plasmid replication, it would not have been an obvious target for rationally engineering a more genetically stable host strain.

The PResERV *rne* allele was responsible for a 6-fold reduction in plasmid mutation rates with no significant effect on chromosomal mutation rates. The altered amino acid (L222S) is located within its split RNaseH-like domain near the embedded 5' sensor domain that is responsible for its preference for RNA substrates with a 5' monophosphate (37). It is unclear how this mutation might affect RNase E activity and lead to a reduction in plasmid mutation rates. It could potentially have a direct effect on processing of RNA I and/or RNA II that alters the balance of different DNA polymerases used to replicate and/or repair ColE1-type plasmids, though it is hard to imagine how this could happen without also affecting plasmid copy number. Alternatively, the *rne* mutation may have an indirect effect by altering the decay or maturation of other RNAs in a cell. RNase E has been shown to affect the biogenesis and activity of small RNAs (84), many of which are involved in stress responses (85), to point out one such possibility among many. It will take future work examining the biochemical effects of this mutation on enzyme activity and its global effects on the *E. coli* transcriptome to decipher why it has an antimutator effect on plasmid replication. Of the three mutations in the PResERV strains that we studied, this mutation in RNase E appears to hold the greatest promise for applications in biotechnology, as the antimutator effect is not associated with any unwanted trade-offs in terms of growth rate or plasmid copy number in the standard culture conditions we used.

Overall, we expect that the PResERV approach will be widely applicable and useful for isolating mutations that make engineered cells more robust against evolutionary failure by lowering mutation rates. One advantage of PResERV is that it is agnostic to the source of mutations and the type of host cell. It will select for mutants that eliminate the dominant cause of mutations inactivating the reporter gene used for cell sorting, and it can be employed iteratively to further eliminate the next-most dominant source of mutations by subsequently introducing new genetic variation into the population and continuing the cycles of cell growth and sorting. When there is genome-wide genetic variation in the cell population, PResERV can discover mutations in genes of unknown function or pathways that do not have obvious connections to mutagenesis, like the *rne* mutation in this study. In the future, PResERV could also be used on libraries of cells that target variation to one key enzyme (e.g. *polA*) or to a suite of genes known to be involved in DNA replication and repair, by using multiplex genome editing methods (86,87). The current study demonstrates proof-of-principle for the PResERV approach, but by examining just six isolates from one mutagenized population,

it has clearly not identified all of the ways that mutations in the *E. coli* genome can lower mutation rates.

An advantage of using directed evolution compared to screening approaches that have been used to isolate antimutators in the past (65,66) is that the cycles of regrowth between cell-sorting steps in PResERV implicitly favors isolating just those antimutator alleles with no trade-off in terms of a reduced growth rate. However, there are potential risks and pitfalls in any directed evolution approach. Selection will yield a reduction in the dominant type of mutation for a particular reporter gene and plasmid in a particular environment and host cell, but these improvements may not translate to other DNA constructs, growth conditions, or genetic backgrounds. In this study, we showed that there are consistent antimutator effects between two distinct ColE1-type plasmids with different reporter genes in the evolved *E. coli* strains. How the antimutator alleles isolated here behave in other environments needs to be further tested to ensure that they do not degrade performance in specific applications. In general, this risk can be mitigated by matching PResERV conditions as closely as possible to those relevant for applications of a strain (e.g. in an industrial bioreactor) or by exposing cells to a variety of different environments during PResERV. It also remains to be seen whether the PResERV mutations would maintain their antimutator effects if they were engineered into other *E. coli* strain backgrounds that are of interest in biotechnology (e.g. BL21 for protein expression).

One critical consideration for applying PResERV is knowing what types of mutations will inactivate the reporter gene used for cell sorting. Certain DNA sequences contain mutational hotspots such that a specific deletion or frameshift dominates among the mutations found to inactivate a reporter gene because it occurs at a rate that is many orders of magnitude higher than the point mutation rate (13). Transposons are the most prominent source of mutations that disrupt other engineered DNA constructs (16,21,88). The presence of any type of dominant mutation in the fluorescent reporter gene will concentrate PResERV on isolating mutants that 'solve' that particular mechanism of failure. In this study, we purposefully used a GFP reporter plasmid that had been edited to remove sequence-based mutational hotspots (21), so that we could recover mutants that reduced point mutation rates. Because transposon mutations can be completely eliminated by using 'clean-genome' strains that have these and other selfish DNA elements deleted from their genomes (11,14–16), it would probably not be a very useful application of PResERV to employ it to find mutations that suppress their activity, at least in bacteria. Rather, we anticipate that PResERV is most useful for neutralizing point mutations, for which it is less obvious how to modify either the sequence of the DNA construct or a cell's genome to improve genetic stability.

One of the main challenges for implementing the PResERV approach in other contexts, as opposed to with a high-copy plasmid in a bacterial cell, is that expression of the reporter gene used to monitor for mutations must impose a large, dominant fitness burden on the host cell. This ensures, first, that cells with mutations in the reporter gene will arise and reach a high frequency within a reasonable

number of growth cycles so that one can complete multiple sorting steps to enrich for antimutator variants. Second, mutations that inactivate the reporter gene that is being monitored will be competing within these populations with other categories of beneficial mutations that improve growth for unrelated reasons (e.g. adaptation to the growth media). If the burden of the reporter gene is too small, then those other mutations will be favored over mutations that change GFP fluorescence, unfocusing evolution from the objective of PResERV. A related challenge, illustrated by the *polA* mutation in this study, is that it may be difficult to guard against a gradual and subtle loss of fluorescence over time during the sorting procedure, which can lead to the enrichment of mutations in the plasmid that modify the expression or burden of the reporter gene with undesirable side-effects.

In summary, we showed that the PResERV directed evolution approach can isolate antimutator *E. coli* variants that exhibit reduced mutagenesis of ColE1-type plasmids. Since these high-copy plasmids are widely used in *E. coli* for cloning and recombinant protein expression, these or similar antimutator alleles may be broadly useful in biotechnology applications. Future applications of PResERV with the burdensome reporter gene encoded in the chromosome or on a plasmid with a different origin of replication, might enrich for host cell variants that have a higher fidelity for replicating other components of a bacterial genome. The PResERV approach could also potentially be applied to other cell types used for industrial bioproduction, such as yeast or Chinese hamster ovary cells, if suitable reporter genes for monitoring genetic stability can be devised for these systems. Despite decades of studying the mechanisms of DNA repair and replication, we do not know the fundamental physiological constraints that determine a lower limit on the mutation rates that could potentially be achieved by tuning or augmenting these processes. Ultimately, this overall strategy of lowering mutation rates to arrest evolution promises to improve the foundations of synthetic biology so that cells engineered for any purpose will function more predictably and reliably.

DATA AVAILABILITY

Genome sequencing files have been deposited with the NCBI Sequence Read Archive under accession number SRP090775.

ACKNOWLEDGEMENTS

We thank Sean Sleight and Herbert Sauro for the gift of pSKO4 and helpful discussions; Drew Tack and Andy Ellington for providing pTEM-1.D254tag; Stratton Georgoulis for contributing to the development of the numerical simulations; and four anonymous reviewers for helpful comments. We acknowledge the Texas Advanced Computing Center (TACC) at The University of Texas at Austin for providing high-performance computing resources.

FUNDING

Defense Advanced Research Projects Agency [HR0011-15-C0095]; National Science Foundation [CBET-1554179];

National Science Foundation BEACON Center for the Study of Evolution in Action [DBI-0939454]; American Society for Microbiology Watkins graduate research fellowship (to D.L.); CONACyT doctoral scholarship [476261 to Á.E.R.]. Funding for open access charge: National Science Foundation [CBET-1554179].

Conflict of interest statement. None declared.

REFERENCES

- Teng, X., Dayhoff-Brannigan, M., Cheng, W.-C., Gilbert, C.E., Sing, C.N., Diny, N.L., Wheelan, S.J., Dunham, M.J., Boeke, J.D., Pineda, F.J. *et al.* (2013) Genome-wide consequences of deleting any single gene. *Mol. Cell*, **52**, 485–494.
- Sandoval, C.M., Ayson, M., Moss, N., Lieu, B., Jackson, P., Gaucher, S.P., Horning, T., Dahl, R.H., Denery, J.R., Abbott, D.A. *et al.* (2014) Use of pantothenate as a metabolic switch increases the genetic stability of farnesene producing *Saccharomyces cerevisiae*. *Metab. Eng.*, **25**, 1–12.
- Gibson, D.G., Glass, J.I., Lartigue, C., Noskov, V.N., Chuang, R.-Y., Algire, M.A., Benders, G.A., Montague, M.G., Ma, L., Moodie, M.M. *et al.* (2010) Creation of a bacterial cell controlled by a chemically synthesized genome. *Science*, **329**, 52–56.
- Wannier, T.M., Kunjapur, A.M., Rice, D.P., McDonald, M.J., Desai, M.M. and Church, G.M. (2018) Adaptive evolution of genomically recoded *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.*, **115**, 3090–3095.
- Bull, J.J. and Barrick, J.E. (2017) Arresting evolution. *Trends Genet.*, **33**, 910–920.
- Scott, M., Gunderson, C.W., Mateescu, E.M., Zhang, Z. and Hwa, T. (2010) Interdependence of cell growth and gene expression: origins and consequences. *Science*, **330**, 1099–1102.
- Borkowski, O., Ceroni, F., Stan, G.-B. and Ellis, T. (2016) Overloaded and stressed: whole-cell considerations for bacterial synthetic biology. *Curr. Opin. Microbiol.*, **33**, 123–130.
- Canton, B., Labno, A. and Endy, D. (2008) Refinement and standardization of synthetic biological parts and devices. *Nat. Biotechnol.*, **26**, 787–793.
- Renda, B.A., Hammerling, M.J. and Barrick, J.E. (2014) Engineering reduced evolutionary potential for synthetic biology. *Mol. Biosyst.*, **10**, 1668–1678.
- Arkin, A.P. and Fletcher, D.A. (2006) Fast, cheap and somewhat in control. *Genome Biol.*, **7**, 114.
- Pósfai, G., Plunkett, G., Fehér, T., Frisch, D., Keil, G.M., Umenhoffer, K., Kolisnychenko, V., Stahl, B., Sharma, S.S., de Arruda, M. *et al.* (2006) Emergent properties of reduced-genome *Escherichia coli*. *Science*, **312**, 1044–1046.
- Csörgö, B., Fehér, T., Timár, E., Blattner, F.R. and Pósfai, G. (2012) Low-mutation-rate, reduced-genome *Escherichia coli*: an improved host for faithful maintenance of engineered genetic constructs. *Microb. Cell Fact.*, **11**, 11.
- Jack, B.R., Leonard, S.P., Mishler, D.M., Renda, B.A., Leon, D., Suárez, G.A. and Barrick, J.E. (2014) Predicting the genetic stability of engineered DNA sequences with the EFM calculator. *ACS Synth. Biol.*, **4**, 939–943.
- Suárez, G.A., Renda, B.A., Dasgupta, A. and Barrick, J.E. (2017) Reduced mutation rate and increased transformability of transposon-free *Acinetobacter baylyi* ADP1-ISx. *Appl. Environ. Microbiol.*, **83**, e01025-17.
- Choi, J.W., Yim, S.S., Kim, M.J. and Jeong, K.J. (2015) Enhanced production of recombinant proteins with *Corynebacterium glutamicum* by deletion of insertion sequences (IS elements). *Microb. Cell Fact.*, **14**, 207.
- Rugbjerg, P., Myling-Petersen, N., Porse, A., Sarup-Lytzen, K. and Sommer, M.O.A. (2018) Diverse genetic error modes constrain large-scale bio-based production. *Nat. Commun.*, **9**, 787.
- Ceroni, F., Algar, R., Stan, G.-B. and Ellis, T. (2015) Quantifying cellular capacity identifies gene expression designs with reduced burden. *Nat. Methods*, **12**, 415–418.
- Romero, P.A. and Arnold, F.H. (2009) Exploring protein fitness landscapes by directed evolution. *Nat. Rev. Mol. Cell Biol.*, **10**, 866–876.

19. Haseltine, E.L. and Arnold, F.H. (2007) Synthetic gene circuits: design with directed evolution. *Annu. Rev. Biophys. Biomol. Struct.*, **36**, 1–19.
20. Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L. and Mori, H. (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.*, **2**, 2006.0008.
21. Sleight, S.C., Bartley, B.A., Lieviant, J.A. and Sauro, H.M. (2010) Designing and engineering evolutionary robust genetic circuits. *J. Biol. Eng.*, **4**, 12.
22. Tack, D.S., Ellefson, J.W., Thyer, R., Wang, B., Gollihar, J., Forster, M.T. and Ellington, A.D. (2016) Addicting diverse bacteria to a noncanonical amino acid. *Nat. Chem. Biol.*, **12**, 138–140.
23. Ellis, B., Haaland, P., Hahne, F., Le Meur, N., Gopalakrishnan, N., Spidlen, J. and Jiang, M. (2018) flowCore: Basic structures for flow cytometry data. R package version 1.42.3.
24. Deatherage, D.E. and Barrick, J.E. (2014) Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using *breseq*. *Methods Mol. Biol.*, **1151**, 165–188.
25. Nehring, R.B., Gu, F., Lin, H.Y., Gibson, J.L., Blythe, M.J., Wilson, R., Bravo Núñez, M.A., Hastings, P.J., Louis, E.J., Frisch, R.L. *et al.* (2015) An ultra-dense library resource for rapid deconvolution of mutations that cause phenotypes in *Escherichia coli*. *Nucleic Acids Res.*, **44**, 1–11.
26. Thomason, L.C., Costantino, N. and Court, D.L. (2007) *E. coli* genome manipulation by P1 transduction. *Curr. Protoc. Mol. Biol.*, doi:10.1002/0471142727.mb0117s79.
27. Datsenko, K.A. and Wanner, B.L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U.S.A.*, **97**, 6640–6645.
28. Rosche, W.A. and Foster, P.L. (2000) Determining mutation rates in bacterial populations. *Methods*, **20**, 4–17.
29. Jin, D.J. and Zhou, Y.N. (1996) Mutational analysis of structure-function relationship of RNA polymerase in *Escherichia coli*. *Methods Enzymol.*, **273**, 300–319.
30. Fehér, T., Cseh, B., Umenhoffer, K., Karcagi, I. and Pósfai, G. (2006) Characterization of *cycA* mutants of *Escherichia coli*. An assay for measuring in vivo mutation rates. *Mutat. Res.*, **595**, 184–190.
31. Zheng, Q. (2017) rSalvador: An R package for the fluctuation experiment. *G3 (Bethesda)*, **7**, 3849–3856.
32. Lee, C., Kim, J., Shin, S.G. and Hwang, S. (2006) Absolute and relative QPCR quantification of plasmid copy number in *Escherichia coli*. *J. Biotechnol.*, **123**, 273–280.
33. Wang, F. and Yang, W. (2009) Structural insight into translesion synthesis by DNA Pol II. *Cell*, **139**, 1279–1289.
34. Lin-Chao, S. and Cohen, S.N. (1991) The rate of processing and degradation of antisense RNA I regulates the replication of ColE1-type plasmids in vivo. *Cell*, **65**, 1233–1242.
35. Kües, U. and Stahl, U. (1989) Replication of plasmids in gram-negative bacteria. *Microbiol. Rev.*, **53**, 491–516.
36. Tomcsányi, T. and Apirion, D. (1985) Processing enzyme ribonuclease E specifically cleaves RNA I. An inhibitor of primer formation in plasmid DNA synthesis. *J. Mol. Biol.*, **185**, 713–720.
37. Callaghan, A.J., Marcaida, M.J., Stead, J.A., McDowall, K.J., Scott, W.G. and Luisi, B.F. (2005) Structure of *Escherichia coli* RNase E catalytic domain and implications for RNA turnover. *Nature*, **437**, 1187–1191.
38. Loh, E., Choe, J. and Loeb, L.A. (2007) Highly tolerated amino acid substitutions increase the fidelity of *Escherichia coli* DNA polymerase I. *J. Biol. Chem.*, **282**, 12201–12209.
39. Beese, L.S., Friedman, J.M. and Steitz, T.A. (1993) Crystal structures of the Klenow fragment of DNA polymerase I complexed with deoxynucleoside triphosphate and pyrophosphate. *Biochemistry*, **32**, 14095–14101.
40. Eyre-Walker, A. and Keightley, P.D. (2007) The distribution of fitness effects of new mutations. *Nat. Rev. Genet.*, **8**, 610–618.
41. Bataillon, T. and Bailey, S.F. (2014) Effects of new mutations on fitness: Insights from models and data. *Ann. N.Y. Acad. Sci.*, **1320**, 76–92.
42. Robert, L., Ollion, J., Robert, J., Song, X., Matic, I. and Elez, M. (2018) Mutation dynamics and fitness effects followed in single cells. *Science*, **359**, 1283–1286.
43. Sniegowski, P.D., Gerrish, P.J., Johnson, T. and Shaver, A. (2000) The evolution of mutation rates: separating causes from consequences. *BioEssays*, **22**, 1057–1066.
44. de Visser, J.A.G.M. (2002) The fate of microbial mutators. *Microbiology*, **148**, 1247–1252.
45. Denamur, E. and Matic, I. (2006) Evolution of mutation rates in bacteria. *Mol. Microbiol.*, **60**, 820–827.
46. Voordeckers, K., Kominek, J., Das, A., Espinosa-Cantú, A., De Maeyer, D., Arslan, A., Van Pee, M., van der Zande, E., Meert, W., Yang, Y. *et al.* (2015) Adaptation to high ethanol reveals complex evolutionary pathways. *PLoS Genet.*, **11**, e1005635.
47. Sniegowski, P.D., Gerrish, P.J. and Lenski, R.E. (1997) Evolution of high mutation rates in experimental populations of *E. coli*. *Nature*, **387**, 703–705.
48. Tenaillon, O., Taddei, F., Radman, M. and Matic, I. (2001) Second-order selection in bacterial evolution: selection acting on mutation and recombination rates in the course of adaptation. *Res. Microbiol.*, **152**, 11–16.
49. Barrick, J.E. and Lenski, R.E. (2013) Genome dynamics during experimental evolution. *Nat. Rev. Genet.*, **14**, 827–839.
50. Funchain, P., Yeung, A., Stewart, J.L., Lin, R., Slupska, M.M. and Miller, J.H. (2000) The consequences of growth of a mutator strain of *Escherichia coli* as measured by loss of function among multiple gene targets and loss of fitness. *Genetics*, **154**, 959–970.
51. Leiby, N. and Marx, C.J. (2014) Metabolic erosion primarily through mutation accumulation, and not tradeoffs, drives limited evolution of substrate specificity in *Escherichia coli*. *PLoS Biol.*, **12**, e1001789.
52. Marvig, R.L., Sommer, L.M., Molin, S. and Johansen, H.K. (2015) Convergent evolution and adaptation of *Pseudomonas aeruginosa* within patients with cystic fibrosis. *Nat. Genet.*, **47**, 57–64.
53. Oliver, A. and Mena, A. (2010) Bacterial hypermutation in cystic fibrosis, not only for antibiotic resistance. *Clin. Microbiol. Infect.*, **16**, 798–808.
54. Oliver, A., Cantón, R., Campo, P., Baquero, F. and Blázquez, J. (2000) High frequency of hypermutable *Pseudomonas aeruginosa* in cystic fibrosis lung infection. *Science*, **288**, 1251–1254.
55. Couce, A., Caudwell, L.V., Feinauer, C., Hindrér, T., Feugeas, J.-P., Weigt, M., Lenski, R.E., Schneider, D. and Tenaillon, O. (2017) Mutator genomes decay, despite sustained fitness gains, in a long-term experiment with bacteria. *Proc. Natl. Acad. Sci. U.S.A.*, **114**, E9026–E9035.
56. McDonald, M.J., Hsieh, Y.Y., Yu, Y.H., Chang, S.L. and Leu, J.Y. (2012) The evolution of low mutation rates in experimental mutator populations of *Saccharomyces cerevisiae*. *Curr. Biol.*, **22**, 1235–1240.
57. Sprouffske, K., Aguilar-Rodríguez, J., Sniegowski, P. and Wagner, A. (2018) High mutation rates limit evolutionary adaptation in *Escherichia coli*. *PLoS Genet.*, **14**, e1007324.
58. Wielgoss, S., Barrick, J.E., Tenaillon, O., Wiser, M.J., Dittmar, W.J., Cruveiller, S., Chane-Woon-Ming, B., Médigue, C., Lenski, R.E. and Schneider, D. (2013) Mutation rate dynamics in a bacterial population reflect tension between adaptation and genetic load. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 222–227.
59. Tenaillon, O., Barrick, J.E., Ribick, N., Deatherage, D.E., Blanchard, J.L., Dasgupta, A., Wu, G.C., Wielgoss, S., Cruveiller, S., Médigue, C. *et al.* (2016) Tempo and mode of genome evolution in a 50,000-generation experiment. *Nature*, **536**, 165–170.
60. Drake, J.W. (1991) A constant rate of spontaneous mutation in DNA-based microbes. *Proc. Natl. Acad. Sci. U.S.A.*, **88**, 7160–7164.
61. Lynch, M. (2010) Evolution of the mutation rate. *Trends Genet.*, **26**, 345–352.
62. Lynch, M., Ackerman, M.S., Gout, J.-F., Long, H., Sung, W., Thomas, W.K. and Foster, P.L. (2016) Genetic drift, selection and the evolution of the mutation rate. *Nat. Rev. Genet.*, **17**, 704–714.
63. Friedberg, E.C., Walker, G.C., Siede, W., Wood, R.D., Schultz, R.A. and Ellenberger, T. (2006) *DNA Repair and Mutagenesis*. 2nd edn. ASM Press, Washington, DC.
64. Galperin, M.Y., Moroz, O.V., Wilson, K.S. and Murzin, A.G. (2006) House cleaning, a part of good housekeeping. *Mol. Microbiol.*, **59**, 5–19.
65. Quiñones, A. and Piechocki, R. (1985) Isolation and characterization of *Escherichia coli* antimutators. A new strategy to study the nature and origin of spontaneous mutations. *Mol. Gen. Genet.*, **201**, 315–322.
66. Fijalkowska, I.J., Dunn, R.L. and Schaaper, R.M. (1993) Mutants of *Escherichia coli* with increased fidelity of DNA replication. *Genetics*, **134**, 1023–1030.

67. Oller, A.R. and Schaaper, R.M. (1994) Spontaneous mutation in *Escherichia coli* containing the *dnaE911* DNA polymerase antimutator allele. *Genetics*, **138**, 263–270.
68. Kamiya, H., Iida, E., Murata-Kamiya, N., Yamamoto, Y., Miki, T. and Harashima, H. (2003) Suppression of spontaneous and hydrogen peroxide-induced mutations by a MutT-type nucleotide pool sanitization enzyme, the *Escherichia coli* Orf135 protein. *Genes Cells*, **8**, 941–950.
69. Fujikawa, K. and Kasai, H. (2002) The oxidized pyrimidine ribonucleotide, 5-hydroxy-CTP, is hydrolyzed efficiently by the *Escherichia coli* recombinant Orf135 protein. *DNA Repair (Amst.)*, **1**, 571–576.
70. Camps, M., Naukkarinen, J., Johnson, B.P. and Loeb, L.A. (2003) Targeted gene evolution in *Escherichia coli* using a highly error-prone DNA polymerase I. *Proc. Natl. Acad. Sci. U.S.A.*, **100**, 9727–9732.
71. Loh, E., Salk, J.J. and Loeb, L.A. (2010) Optimization of DNA polymerase mutation rates during bacterial evolution. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 1154–1159.
72. Yang, Y.L. and Polisky, B. (1993) Suppression of ColE1 high-copy-number mutants by mutations in the *polA* gene of *Escherichia coli*. *J. Bacteriol.*, **175**, 428–437.
73. Frisch, R.L., Su, Y., Thornton, P.C., Gibson, J.L., Rosenberg, S.M. and Hastings, P.J. (2010) Separate DNA Pol II- and Pol IV-dependent pathways of stress-induced mutation during double-strand-break repair in *Escherichia coli* are controlled by RpoS. *J. Bacteriol.*, **192**, 4694–4700.
74. Dapa, T., Fleurier, S., Bredeche, M.-F. and Matic, I. (2017) The SOS and RpoS regulons contribute to bacterial cell robustness to genotoxic stress by synergistically regulating DNA polymerase Pol II. *Genetics*, **206**, 1349–1360.
75. Witkin, E.M. (1976) Ultraviolet mutagenesis and inducible DNA repair in *Escherichia coli*. *Bacteriol. Rev.*, **40**, 869–907.
76. Janel-Bintz, R., Napolitano, R.L., Isogawa, A., Fujii, S. and Fuchs, R.P. (2017) Processing closely spaced lesions during Nucleotide Excision Repair triggers mutagenesis in *E. coli*. *PLoS Genet.*, **13**, e1006881.
77. Escarceller, M., Hicks, J., Gudmundsson, G., Trump, G., Touati, D., Lovett, S., Foster, P.L., McEntee, K. and Goodman, M.F. (1994) Involvement of *Escherichia coli* DNA polymerase II in response to oxidative damage and adaptive mutation. *J. Bacteriol.*, **176**, 6221–6228.
78. Hastings, P.J., Hersh, M.N., Thornton, P.C., Fonville, N.C., Slack, A., Frisch, R.L., Ray, M.P., Harris, R.S., Leal, S.M. and Rosenberg, S.M. (2010) Competition of *Escherichia coli* DNA polymerases I, II and III with DNA Pol IV in stressed cells. *PLoS One*, **5**, e10862.
79. Napolitano, R., Janel-Bintz, R., Wagner, J. and Fuchs, R.P. (2000) All three SOS-inducible DNA polymerases (Pol II, Pol IV and Pol V) are involved in induced mutagenesis. *EMBO J.*, **19**, 6259–6265.
80. Wagner, J., Etienne, H., Janel-Bintz, R. and Fuchs, R.P. (2002) Genetics of mutagenesis in *E. coli*: Various combinations of translesion polymerases (Pol II, IV and V) deal with lesion/sequence context diversity. *DNA Repair (Amst.)*, **1**, 159–167.
81. Cirz, R.T., Chin, J.K., Andes, D.R., De Crécy-Lagard, V., Craig, W.A. and Romesberg, F.E. (2005) Inhibition of mutation and combating the evolution of antibiotic resistance. *PLoS Biol.*, **3**, 1024–1033.
82. Mohanty, B.K. and Kushner, S.R. (2016) Regulation of mRNA decay in bacteria. *Annu. Rev. Microbiol.*, **70**, 25–44.
83. Stead, M.B., Marshburn, S., Mohanty, B.K., Mitra, J., Pena Castillo, L., Ray, D., van Bakel, H., Hughes, T.R. and Kushner, S.R. (2011) Analysis of *Escherichia coli* RNase E and RNase III activity *in vivo* using tiling microarrays. *Nucleic Acids Res.*, **39**, 3188–3203.
84. Chao, Y., Li, L., Girodat, D., Förstner, K.U., Said, N., Corcoran, C., Smiga, M., Papenfort, K., Reinhardt, R., Wieden, H.-J. *et al.* (2017) *In vivo* cleavage map illuminates the central role of RNase E in coding and non-coding RNA pathways. *Mol. Cell*, **65**, 39–51.
85. Gottesman, S., McCullen, C.A., Guillier, M., Vanderpool, C.K., Majdalani, N., Benhammou, J., Thompson, K.M., FitzGerald, P.C., Sowa, N.A. and FitzGerald, D.J. (2006) Small RNA regulators and the bacterial response to stress. *Cold Spring Harb. Symp. Quant. Biol.*, **71**, 1–11.
86. Wang, H.H., Isaacs, F.J., Carr, P.A., Sun, Z.Z., Xu, G., Forest, C.R. and Church, G.M. (2009) Programming cells by multiplex genome engineering and accelerated evolution. *Nature*, **460**, 894–898.
87. Garst, A.D., Bassalo, M.C., Pines, G., Lynch, S.A., Halweg-Edwards, A.L., Liu, R., Liang, L., Wang, Z., Zeitoun, R., Alexander, W.G. *et al.* (2017) Genome-wide mapping of mutations at single-nucleotide resolution for protein, metabolic and genome engineering. *Nat. Biotechnol.*, **35**, 48–55.
88. Umenhoffer, K., Fehér, T., Balikó, G., Ayaydin, F., Pósfa, J., Blattner, F.R. and Pósfa, G. (2010) Reduced evolvability of *Escherichia coli* MDS42, an IS-less cellular chassis for molecular and synthetic biology applications. *Microb. Cell Fact.*, **9**, 38.