

Highly mutagenic continuous evolution in *E. coli* using a Φ 29-based orthogonal replication system

Received: 26 May 2025

Fabian B. H. Rehm , Kim C. Liu , Rongzhen Tian & Jason W. Chin 

Accepted: 31 October 2025

Published online: 24 February 2026

 Check for updates

Organisms that permit hypermutation of target genes without off-target mutagenesis of the host genome enable the accelerated, continuous evolution of genes for new or enhanced functions. We develop and optimize an orthogonal DNA replication system in *Escherichia coli* that uses components from bacteriophage Φ 29. The minimal system requires just two Φ 29 genes to maintain the replicon and replicons can be efficiently engineered in vivo. We generate a highly mutagenic Φ 29 DNA polymerase that introduces mutations at a frequency approaching 10^{-4} per base per generation (one mutation in a 1-kb gene every ten generations). Our system is stable for hundreds of generations and enables the continuous, accelerated evolution of new gene functions. We demonstrate the rapid evolution of a tetracycline resistance gene to confer resistance to tigecycline at higher levels than achieved with previously reported systems. We further evolve a 1,000-fold increase in β -lactamase activity for a third-generation cephalosporin in just 3 days.

The high-fidelity replication of an organism's genome is essential for maintaining genetic integrity and limits mutations that could be detrimental to survival or reproduction; therefore, the natural evolution of new gene function, resulting from the accumulation of mutations and selection within a population, is a slow process.

Directed evolution enables the rapid diversification and selection of genes for a desired phenotype on laboratory timescales¹. Classical directed evolution experiments rely on cycles of in vitro genetic diversification followed by transformation into an appropriate host for selection. These types of experiments are laborious and limited in terms of evolutionary depth and scale. Continuous evolution strategies, enabled by in vivo mutagenesis and selection of target genes, may overcome these limitations and allow for the accelerated exploration of fitness landscapes at scale². However, the sequence that can be explored may be a function of the mutational spectrum that is accessed, the copy number of the genes of interest and whether selective pressure is continuous or intermittent.

Early strategies for continuous evolution included approaches that increased the genome-wide mutation rate^{3,4}. However, these approaches were limited by the fitness burden associated with mutagenesis of essential genes⁵. This restriction limits both the number

of generations for which the selection can be performed and the phenotypes that can be evolved.

Inserting target genes into viral genomes and iteratively infecting fresh mutagenic cells^{6,7} sidesteps some of the limitations of the initial approaches. However, these approaches are limited to evolving small genes that can be packaged into viruses and are often run noncontinuously for practical reasons⁸. Other strategies rely on the localization of a mutagenic moiety, such as an error-prone polymerase or a deaminase, to a target gene^{9–14}. These approaches are generally restricted by their mutagenic window and limited orthogonality; these limitations may lead to off-target mutations that can culminate in selection escape. By contrast, orthogonal replication systems, where a dedicated error-prone DNA polymerase (DNAP) specifically maintains an episome harboring the target genes without interfering with genome replication, can enable straightforward and robust continuous evolution experiments^{15–18}.

Existing orthogonal replication systems primarily rely on DNAPs that replicate linear replicons through a protein-primed mechanism of replication^{15–17,19}. The first such system, OrthoRep, exploits natural linear plasmids from yeast and has enabled a range of continuous evolution experiments at high mutation rates^{20–22}. Most genetic tools,

however, are available in the workhorse of synthetic biology, *E. coli* and are often not compatible with yeast²³. Moreover, *E. coli* grows more rapidly and to higher densities than *S. cerevisiae*, enabling faster generation times—a key parameter governing the speed and scale of continuous evolution experiments. The advantages of *E. coli* as a host were realized by a recent orthogonal replication system, EcOREp; this system uses components from the PRD1 bacteriophage, which naturally infects a range of Gram-negative bacteria, including *E. coli*¹⁶. However, the initial mutation rates achieved by this system were only modestly higher than the genomic mutation rates (2×10^{-7} substitutions per base, per generation) and replicon transformation efficiencies were low.

Here, we report the development of a highly mutagenic orthogonal replication system in *E. coli*. Our minimal system relies only on two genes from bacteriophage Φ 29, which exclusively infects Gram-positive hosts and has served as a model to study protein-primed DNA replication^{24–32}. We achieve high replicon transformation efficiencies, develop an efficient strategy to engineer replicons in vivo and identify an error-prone DNAP that can mutagenize the replicon at rates approaching 10^{-4} substitutions per base, per generation.

Results

Design and establishment of a Φ 29-based replication system in vivo

Bacteriophage Φ 29 is a lytic, double-stranded DNA bacteriophage belonging to the Salasmaviridae family³³. First discovered in 1965 (ref. 34), it remains the smallest known phage that infects *Bacillus*. The Φ 29 genome (19,282 bp) encodes 27 protein-coding genes and a prohead RNA that is essential for genome packaging (Fig. 1a)³³. The encoded proteins are involved in replicating and packaging the genome and lysing the host cell^{33,35,36}. The Φ 29 genome is linear and contains origin of replication sequences (oriL and oriR) at each end. The genome is capped at its 5' ends by covalently bound terminal proteins (TPs); these proteins prime replication by the Φ 29 DNAP^{37,38}. The Φ 29 genome is an important model system for studying protein-primed DNA replication.

We aimed to generate a synthetic system for the controlled replication of a linear Φ 29 replicon in *E. coli*. Toward this goal, we designed a synthetic replication operon consisting of the genes encoding TP (gene 3), DNAP (gene 2), double-stranded DNA-binding protein (DSB; gene 6) and single-stranded DNA-binding protein (SSB; gene 5) (Fig. 1a); a combination of these proteins has been used to amplify portions of the Φ 29 genome or heterologous genes in vitro^{39,40}. We arranged these genes into an operon under the control of an IPTG-inducible promoter, PtacIPTG, on a single-copy plasmid (Fig. 1a). All genes were computationally codon-optimized for use in *E. coli* and their expression was driven by synthetic ribosome-binding site sequences^{41–43}.

We designed a Φ 29 synthetic replicon to be replicated by the synthetic replication operon; the replicon consisted of genes encoding tetracycline resistance (TcR) and GFP, flanked by the left and right Φ 29 origins of replication (oriL and oriR) (Fig. 1a). We attempted to establish the Φ 29 synthetic replicon in *E. coli* by electroporating 5 μ g of the Φ 29 synthetic replicon, generated by PCR amplification, into *E. coli* cells harboring the synthetic replication operon plasmid (Fig. 1b). However, this yielded no tetracycline-resistant colonies (Fig. 1c). This indicated that our initial attempt to establish the replicon system was unsuccessful.

We hypothesized that the establishment of the replicon may be limited by degradation of the electroporated PCR product in *E. coli*. To maximize the chance of establishing the synthetic replicon, we generated helper plasmids that express genes encoding the Gam protein (derived from the lambda phage) to inhibit host nucleases (RecBCD and SbcCD), as well as the Φ 29 DSB and SSB to further protect the electroporated replicon DNA.

Upon electroporation of the Φ 29 synthetic replicon into cells bearing the helper plasmids and the synthetic replication operon, we observed a small number of tetracycline-resistant colonies that exhibited green fluorescence. These experiments suggested that the

Φ 29 synthetic replicon was established in these cells (Fig. 1c); this conclusion was supported by additional genotyping and sequencing experiments (Supplementary Fig. 1) and by experiments where we could reduce the replicon copy number by downregulating expression from the replication operon by targeting the promoter driving the operon with dCas9 (Supplementary Fig. 2). Extracting the established replicons from cells bearing the helper plasmids and the synthetic replication operon and electroporating them into fresh cells enabled substantially improved transformation efficiencies, even when the helper plasmids providing genes encoding Gam and Φ 29 DSB and SSB were not present in the recipient cells (Fig. 1c).

We measured the growth of cells with or without the synthetic Φ 29 replication operon and the Φ 29 synthetic replicon (Supplementary Fig. 3). This revealed the replication operon itself to lead to a mild reduction in cell fitness (doubling times of 49–52 min), whereas cells harboring both the replication operon and the Φ 29 synthetic replicon exhibited substantially slower growth (doubling time of 87 min) and a lower maximum cell density (Supplementary Fig. 3).

The Φ 29 synthetic replicon is stably inherited

To assess the stability of the Φ 29 synthetic replicon through many rounds of cell division, we measured the proportion of cells that maintained the replicon, as indicated by GFP fluorescence, over 550 generations (Fig. 1d and Supplementary Fig. 4). Without tetracycline, the replicon was lost in less than 50 generations. In the presence of tetracycline, the replicon was stably maintained for the entire 550 generations tested. Therefore, the Φ 29 synthetic replicon can be stably maintained for many generations, as required for continuous directed evolution experiments.

Defining the genes required for maintenance of the Φ 29 synthetic replicon

Next, we defined the genes necessary for replication of the Φ 29 synthetic replicon in *E. coli*. We generated a set of single-copy plasmids harboring variants of the Φ 29 synthetic replication operon; in each member of the set, an individual gene (or pair of genes) was disrupted. We then electroporated the Φ 29 synthetic replicon into cells containing each of these plasmids and assessed the formation of colonies that exhibited tetracycline resistance and green fluorescence. This experiment revealed that only two genes, encoding TP and DNAP, are essential for replicative maintenance of the Φ 29 synthetic replicon (Fig. 1e). We passaged cells harboring the Φ 29 synthetic replication operon, with or without the genes encoding the SSB and DSB proteins, and a TcR–GFP Φ 29 synthetic replicon for 50 generations and assessed whether the Φ 29 synthetic replicons remained intact by colony PCR (Supplementary Fig. 5). We only observed full-length replicons, suggesting that the absence of the SSB and DSB did not notably impair replication fidelity. Our results are in agreement with in vitro efforts to replicate the Φ 29 genome using only TP and DNAP and the self-replication of TP-encoding and DNAP-encoding replicons in synthetic protocells^{44–46}.

Efficient engineering of the Φ 29 synthetic replicon

As the establishment of new Φ 29 synthetic replicons from PCR amplicons was a low-efficiency process, we sought to develop a strategy for engineering replicons that were established in cells to introduce new DNA sequences of interest. The lambda Red recombination system, from the lambda phage, has been extensively used for engineering bacterial genomes and plasmids through the introduction of linear DNA flanked by short (>30 bp) stretches of homology to the target site^{47,48}. To assess whether this recombination system could be used to engineer Φ 29 synthetic replicons, we tested replacing the tetracycline resistance gene on the replicon with a kanamycin resistance gene. We electroporated 0.5 μ g of a PCR product consisting of a kanamycin resistance gene flanked by 60-bp homologies into cells harboring a Φ 29 synthetic

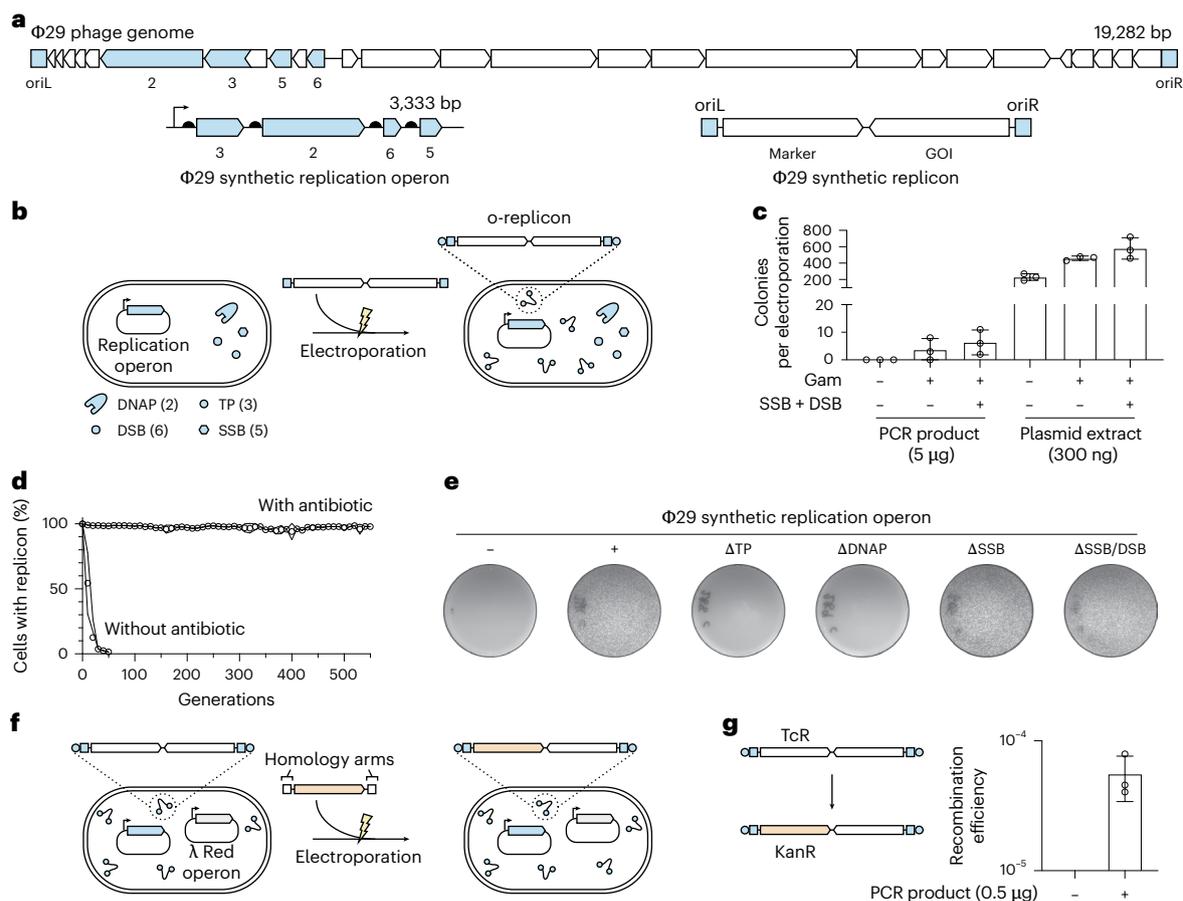


Fig. 1 | Establishing and characterizing a Φ 29-based replication system in *E. coli*. **a**, A synthetic replication operon was designed on the basis of genes 3, 2, 6 and 5, encoding the TP, DNAP, DSB and SSB, respectively. The synthetic replicon was designed to encode an antibiotic resistance gene and a gene of interest (GOI) and was flanked by the left and right origins of replication (oriL and oriR) derived from the Φ 29 genome. **b**, Φ 29 synthetic replicons could be established by electroporating a PCR-derived product into cells harboring the Φ 29 synthetic replication operon on a single-copy plasmid. **c**, Efficiency of establishing Φ 29 synthetic replicons by electroporation. A helper plasmid was used to additionally express *gam* and the genes encoding the Φ 29 SSB and DSB. The Φ 29 synthetic replicons extracted from cells could be transformed with a higher efficiency ($n = 3$ technical replicates—three electroporations into the same batch of electrocompetent cells); error bars indicate the s.d. We note that

‘Plasmid extract’ includes the replicon and copurified plasmid encoding the replication operon; thus, the quantity of replicon transformed is lower than 300 ng. **d**, Stability of the TcR-GFP Φ 29 synthetic replicon over 550 generations with or without tetracycline, as assessed by maintenance of GFP fluorescence using flow cytometry ($n = 4$ biological replicates); lines indicate the s.d. **e**, Essentiality of genes in the synthetic replication operon for transformation of extracted Φ 29 synthetic replicons. **f**, Established Φ 29 synthetic replicons can be engineered using lambda Red recombination by the electroporation of a PCR amplicon flanked by homologies to the replicon. **g**, Efficiency of replacing the tetracycline resistance gene on the Φ 29 synthetic replicon with a kanamycin resistance gene ($n = 3$ technical replicates—three electroporations into the same batch of electrocompetent cells); error bars indicate the s.d.

replication operon plasmid, a tetracycline-resistance-conferring Φ 29 synthetic replicon and a plasmid encoding the lambda Red recombination system components under arabinose-inducible control (Fig. 1f). These experiments gave rise to kanamycin-resistant colonies, where the tetracycline resistance gene was replaced, at a frequency approaching 10^{-4} (Fig. 1g). All modifications of the Φ 29 synthetic replicon performed in this work were conducted using this approach.

Optimization of the Φ 29 synthetic replication operon

Next, we sought to improve the performance of the designed Φ 29 synthetic replication operon by experimentally optimizing DNA sequences that may control the relative levels of the proteins produced from the operon. To achieve this, we generated a Φ 29 synthetic replication operon library in which we varied three regions of the designed operon: we randomized the spacer sequence between the Shine–Delgarno sequence and the start codon of each gene (this sequence can modulate the efficiency of translation initiation, which is rate determining for translation in *E. coli*), we randomized the first

codon following the start codon (to allow for stability tuning according to the N-end rule, which differs between the *Bacillus* species that Φ 29 phage infect and *E. coli*) and we randomized codons 2–7 to synonyms (as the sequence in this region has been shown to influence translational efficiency) (Fig. 2a)^{49–53}.

We passaged cells transformed with the Φ 29 synthetic replication operon library and a Φ 29 synthetic replicon to enrich operons with favorable properties. To select for operons that supported increased replicon copy numbers, we picked colonies exhibiting stronger GFP fluorescence. We identified an operon, Φ 29-opt (Supplementary Fig. 6), with improved performance; this operon supported a higher replicon copy number (Fig. 2b), improved transformation of extracted replicons ($>10^6$ transformants per μ g of transformed replicon; Fig. 2c), improved establishment of new replicons (Supplementary Fig. 7) and faster growth while maintaining a Φ 29 synthetic replicon (doubling time reduced from 87 to 66 min) to higher cell densities (Fig. 2d and Supplementary Fig. 4). We hypothesize that the improved growth is because of the increased capacity to maintain the Φ 29 synthetic

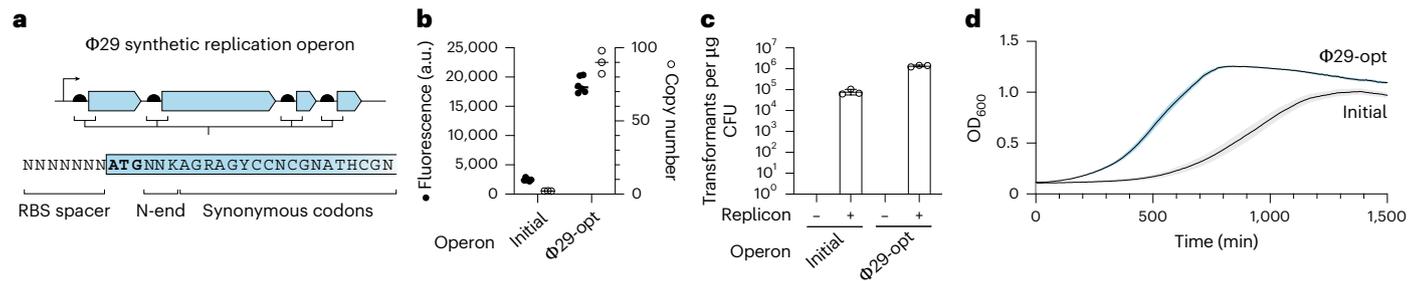


Fig. 2 | Improving $\Phi 29$ synthetic replication operon performance. **a**, A $\Phi 29$ synthetic replication operon library was generated where, for each gene in the operon, the spacer sequence between the Shine–Delgarno sequence and the start codon was randomized, the N-terminal codon following the start codon was saturated and the following six codons were synonymized. **b**, The $\Phi 29$ -opt operon, integrated on a single-copy plasmid, supports a higher replicon copy number, as measured by GFP fluorescence and qPCR ($n = 6$ biological replicates for fluorescence; $n = 3$ biological replicates for qPCR). Fluorescence values were

normalized to culture density. **c**, Extracted $\Phi 29$ synthetic replicons could be efficiently transformed into cells where the $\Phi 29$ -opt operon was genomically integrated ($n = 3$ technical replicates—three electrocompetent cells). Error bars indicate the s.d. CFU, colony-forming unit. **d**, The $\Phi 29$ -opt operon, encoded from a single-copy plasmid, enables improved growth when maintaining a TcR–GFP $\Phi 29$ synthetic replicon ($n = 16$ biological replicates). The boundaries indicate the s.e.m.

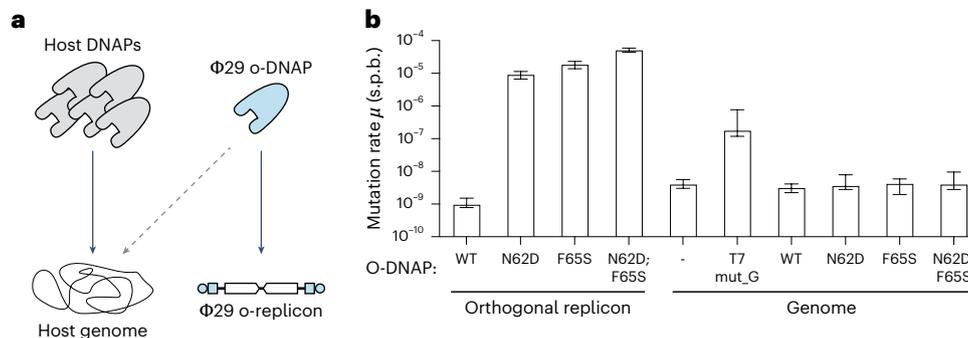


Fig. 3 | The $\Phi 29$ -based replication system is an orthogonal replication system in *E. coli*. **a**, An orthogonal replication system enables hypermutation of a target DNA sequence without interfering with the high-fidelity replication of genomic DNA. **b**, Error-prone mutants of the $\Phi 29$ DNAP do not increase the genomic mutation rate but do increase the replicon mutation rate.

The mutagenic T7 DNAP mutant (mut_G) has previously been reported¹⁸. We used fluctuation analysis to calculate the mutation rate from the frequency of a TAG stop codon reversion in a genomically or replicon-integrated chloramphenicol resistance gene (Q38TAG) ($n = 12$ biological replicates). Error bars indicate the median \pm upper or lower 95% bounds.

replicon, ensuring stable transfer of the replicon to daughter cells at every division.

We next tested whether $\Phi 29$ -opt could support longer $\Phi 29$ synthetic replicons. We used lambda Red recombination to insert variable lengths of the *CFTR* gene and a *CmrR* gene into a KanR–GFP replicons to yield replicons with lengths of 6.6 and 10.5 kb. The longer replicons were established and maintained, as assessed by genotyping and sequencing (Supplementary Fig. 8).

Highly mutagenic $\Phi 29$ -based orthogonal replication system

We generated $\Phi 29$ synthetic replication operons, based on $\Phi 29$ -opt, with variants (N62D and F65S, alone or in combination) of the $\Phi 29$ DNAP; in vitro experiments, these individual DNAP mutants are reported to mutate the DNA that they replicate while retaining strand-displacement activity^{54,55}. We used these operon variants to demonstrate the creation of a highly mutagenic orthogonal replication system in *E. coli* (Fig. 3).

To measure the extent to which these $\Phi 29$ DNAP mutants mutagenize the $\Phi 29$ synthetic replicon, we generated a replicon harboring a chloramphenicol resistance gene with an in-frame stop codon (Q38TAG). After ten generations of growth, with each DNAP variant and a wild-type (WT) control, we measured the fraction of chloramphenicol-resistant cells arising from a point mutation in the TAG stop codon to generate sense codons. These experiments allowed us to determine the apparent mutation rates of the replicon

in the presence of each DNAP using fluctuation analysis^{56,57}. This analysis revealed that the DNAP mutants were highly mutagenic on the replicon (Fig. 3 and Supplementary Fig. 9). The N62D;F65S double mutant outperformed the individual mutants, with a mutation rate approaching 10^{-4} substitutions per base, per generation on the replicon (Fig. 3). To determine the mutational spectra of these error-prone DNAP mutants, we grew cells harboring a $\Phi 29$ synthetic replicon for up to 50 generations and measured the accumulation of mutations using next-generation sequencing (Supplementary Fig. 10). We primarily observed C/G>A/T mutations, although all types of mutations were detected (Supplementary Figs. 10–12). We found mutation rates calculated from this sequencing data to be between 9.6×10^{-5} and 1.8×10^{-4} for the N62D;F65S DNAP double mutant (Supplementary Figs. 11 and 12). Moreover, the introduced mutations did not appear to exhibit positional bias (Supplementary Figs. 11 and 12).

Next, we integrated a chloramphenicol resistance gene containing an in-frame stop codon (Q38TAG) into the genome of *E. coli*. We introduced the $\Phi 29$ DNAP mutants into these cells and measured the fraction of chloramphenicol-resistant cells, arising from point mutations in the genomic TAG stop codon that generate a sense codon, after ten generations. These experiments allowed us to determine the apparent mutation rates of the genome in the presence of each DNAP using fluctuation analysis^{56,57}. This analysis revealed that expression of the WT or mutant $\Phi 29$ DNAPs did not lead to changes in the apparent genomic mutation rate with respect to control cells without a $\Phi 29$

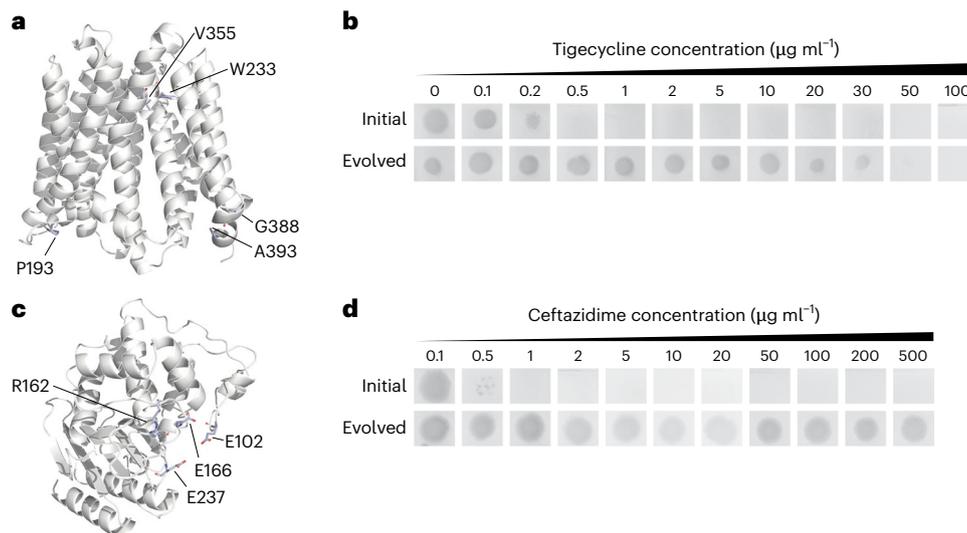


Fig. 4 | Continuous evolution of tigecycline and ceftazidime resistance using the Φ 29-based replication system. **a**, Structure of TetA (AlphaFold prediction) with the residues that were substituted in at least two independent evolution replicates shown in stick representation. All observed substitutions are summarized in Supplementary Data 1. **b**, Pools of cells before and after continuous evolution for enhanced tigecycline resistance. Other pools are shown in Supplementary Fig. 15. Hits from the evolution were also cloned into a standard circular plasmid and validated (Supplementary Fig. 16). **c**, Structure of

the TEM-1 β -lactamase (PDB 1XPB) with the residues that were substituted in at least two independent ceftazidime resistance evolution replicates shown in stick representation. All observed substitutions are summarized in Supplementary Data 2. **d**, Pools of cells before and after continuous evolution for enhanced ceftazidime resistance. Other pools are shown in Supplementary Fig. 19. Hits from the evolution were also cloned into a standard circular plasmid and validated (Supplementary Fig. 20).

DNAP (Fig. 3). By contrast, expression of an error-prone mutant of T7 DNAP led to increased mutagenesis of the genome^{18,58}. We additionally measured genomic mutation rates in the presence of a Φ 29 replicon but did not observe a difference in the apparent mutation rates (Supplementary Fig. 13).

Overall, our data demonstrate that mutagenic Φ 29 DNAPs enable mutation rates approaching 10^{-4} substitutions per base, per generation on the replicon, but do not measurably affect the mutation rate of the genome. We conclude that the Φ 29-based replication system constitutes an orthogonal replication system in *E. coli*.

Continuous evolution of tigecycline, ceftazidime and cefotaxime resistance

Next, we sought to use the Φ 29 orthogonal replication system to continuously evolve new phenotypes. We initially investigated evolving the tetracycline-resistance-conferring *tetA* gene into a gene that confers resistance to tigecycline. This evolution has been carried out using previously reported continuous evolution approaches^{9,16}. We passaged cells containing a *tetA*-GFP replicon and the N62D;F65S error-prone DNAP in increasingly higher concentrations of tigecycline, until a final concentration of $40 \mu\text{g ml}^{-1}$ was reached (12 passages; Supplementary Fig. 14). The resultant pools of cells grew on plates containing tigecycline concentrations as high as $50 \mu\text{g ml}^{-1}$, whereas the pools of cells before continuous evolution grew on $0.2 \mu\text{g ml}^{-1}$ but not on $0.5 \mu\text{g ml}^{-1}$ (Fig. 4 and Supplementary Fig. 15). We sequenced replicons from cells growing on high tigecycline concentrations and found that certain mutations had convergently emerged in our three independent replicates (Supplementary Data 1). Of note are W233, which was substituted to C or S in every clone, A393, which was substituted to S or D in every clone, P193, which was substituted to S, T or N in 79% of clones, G388, which was substituted to V in some clones from all replicates but in all clones from the third replicate, and V355, which was substituted to F in two of the three replicates (Fig. 4a). We cloned select sequences from each replicate into a standard circular plasmid (colE1) to further validate the performance of these mutants (Supplementary Fig. 16). This revealed that all mutants tested could

confer resistance to tigecycline concentrations between 10 and $30 \mu\text{g ml}^{-1}$, whereas the WT *tetA* control was unable to grow on $1 \mu\text{g ml}^{-1}$ tigecycline. We also tested the susceptibility of these evolved variants to counter-selection as mediated by a combination of fusaric acid and ZnCl_2 , revealing a range of sensitivities (Supplementary Fig. 17). The most sensitive variant had 10^3 – 10^4 -fold higher sensitivity compared to the parent TetA and, thus, constitutes an improved dual positive–negative selection marker⁵⁹.

Next, we sought to rapidly evolve TEM-1 β -lactamase for improved activity against the third-generation cephalosporin ceftazidime. Third-generation cephalosporin antibiotics differ from earlier generations because of their improved resistance to β -lactamase activity⁶⁰. We passaged cells harboring a replicon encoding the TEM-1 β -lactamase in increasing concentrations of ceftazidime, ranging from 2 to $500 \mu\text{g ml}^{-1}$ (Supplementary Fig. 18). We completed three passages in 3 days, yielding pools of cells that grew on $500 \mu\text{g ml}^{-1}$ ceftazidime (Fig. 4 and Supplementary Fig. 19). The input pools of cells grew poorly on just $0.5 \mu\text{g ml}^{-1}$ ceftazidime and did not grow on $1 \mu\text{g ml}^{-1}$. We sequenced replicons from cells growing on high ceftazidime concentrations and found that certain mutations had convergently emerged in our four independent replicates (Supplementary Data 2). Of note are E102, which was substituted to K in every sequenced clone, R162, which was substituted to S or N in every sequenced clone, E237, which was substituted to K in some clones in two independent evolution replicates, and E166, which was substituted to D in some clones from three of the four independent evolution replicates (Fig. 4c). We cloned select sequences from each replicate into a standard circular plasmid (colE1) to further validate the performance of these mutants (Supplementary Fig. 20). This revealed that all mutants tested could confer resistance to ceftazidime concentrations between 200 and $500 \mu\text{g ml}^{-1}$, whereas the WT control was unable to grow on $1 \mu\text{g ml}^{-1}$ ceftazidime.

Subsequently, we assessed the possibility of evolving multi-resistant TEM-1 β -lactamase variants. We passaged the pools obtained from the ceftazidime resistance evolution in increasing concentrations of cefotaxime, ranging from 0.5 to $100 \mu\text{g ml}^{-1}$, while simultaneously selecting for both ceftazidime and carbenicillin resistance by

supplementing these antibiotics at $100 \mu\text{g ml}^{-1}$ (Supplementary Fig. 21). We completed four passages in 4 days, yielding pools of cells that grew on 200 to $500 \mu\text{g ml}^{-1}$ cefotaxime while maintaining resistance to both carbenicillin and ceftazidime (Supplementary Fig. 21). By contrast, the ceftazidime-resistant pools, which were the input to this evolution, grew poorly or not at all at greater than $10 \mu\text{g ml}^{-1}$ of cefotaxime. We sequenced replicons from cells growing on high cefotaxime concentrations and found that certain mutations had convergently emerged between replicates 1 and 3 and between replicates 2 and 4 (Supplementary Fig. 21 and Supplementary Data 2). Replicates 2 and 4 had diverged most from the ceftazidime-resistant clones, with P165T, A170D and M180T substitutions in every replicon sequenced from these replicates, in addition to other substitutions. We cloned select sequences from each replicate into a standard circular plasmid (colE1) to further validate the performance of these mutants (Supplementary Fig. 21). This revealed that all mutants tested could confer resistance to cefotaxime concentrations between 50 and $200 \mu\text{g ml}^{-1}$, whereas clone 1 from the ceftazidime resistance evolution grew poorly on $10 \mu\text{g ml}^{-1}$ cefotaxime. All tested clones maintained ceftazidime and carbenicillin resistance (Supplementary Fig. 21). Given that these evolutions are run with multiple replicons per cell, it would in principle be possible for multiresistance to either emerge in a single genotype or for distinct resistances to be conferred by separate genotypes. In this case, all obtained clones fall into the former category. We hypothesize that the improved heritability associated with a single genotype conferring multiresistance may provide a selective advantage.

These experiments demonstrate that the $\Phi 29$ orthogonal replication system enables the rapid and reproducible evolution of new gene function.

Discussion

We established an orthogonal replication system in *E. coli* using a synthetically designed operon encoding components from bacteriophage $\Phi 29$, which naturally infects only Gram-positive bacteria. This phage has served as a long-standing model system to study protein-primed replication. Here, we report sustained in vivo replication using $\Phi 29$ components. We engineered the operon for improved performance and found that only two of the genes are needed for orthogonal replication. Moreover, we used the lambda-phage-derived recombination system to precisely and efficiently engineer the replicons and showed that extracted replicons could be efficiently transformed into fresh cells harboring the optimized synthetic replication operon. We developed an error-prone DNAP with a mutation rate approaching 10^{-4} , exceeding the highest mutation rates reported for EcoRep DNAPs¹⁶. Thus, our $\Phi 29$ -based orthogonal replication system constitutes a highly mutagenic orthogonal replication system for accelerated continuous evolution in *E. coli*. We expect that this system will readily integrate with selection regimes where a desired activity is coupled to bacterial growth.

Future work will focus on using the $\Phi 29$ -based orthogonal replication system to rapidly evolve protein function, predict mutations that will confer resistance to antibiotics and other drugs and rapidly generate synthetic phylogenies for proteins with new and useful functions.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41587-025-02944-x>.

References

- Romero, P. A. & Arnold, F. H. Exploring protein fitness landscapes by directed evolution. *Nat. Rev. Mol. Cell Biol.* **10**, 866–876 (2009).
- Molina, R. S. et al. In vivo hypermutation and continuous evolution. *Nat. Rev. Methods Primers* **2**, 36 (2022).
- Badran, A. H. & Liu, D. R. Development of potent in vivo mutagenesis plasmids with broad mutational spectra. *Nat. Commun.* **6**, 8425 (2015).
- Camps, M., Naukkarinen, J., Johnson, B. P. & Loeb, L. A. Targeted gene evolution in *Escherichia coli* using a highly error-prone DNA polymerase I. *Proc. Natl Acad. Sci. USA* **100**, 9727–9732 (2003).
- Bull, J. J., Sanjuan, R. & Wilke, C. O. Theory of lethal mutagenesis for viruses. *J. Virol.* **81**, 2930–2939 (2007).
- Esvelt, K. M., Carlson, J. C. & Liu, D. R. A system for the continuous directed evolution of biomolecules. *Nature* **472**, 499–503 (2011).
- English, J. G. et al. VEGAS as a platform for facile directed evolution in mammalian cells. *Cell* **178**, 748–761 (2019).
- Miller, S. M., Wang, T. & Liu, D. R. Phage-assisted continuous and non-continuous evolution. *Nat. Protoc.* **15**, 4101–4127 (2020).
- Yi, X., Khey, J., Kazlauskas, R. J. & Travisano, M. Plasmid hypermutation using a targeted artificial DNA replisome. *Sci. Adv.* **7**, eabg8712 (2021).
- Moore, C. L., Papa, L. J. 3rd & Shoulders, M. D. A processive protein chimera introduces mutations across defined DNA regions in vivo. *J. Am. Chem. Soc.* **140**, 11560–11564 (2018).
- Cravens, A., Jamil, O. K., Kong, D., Sockolosky, J. T. & Smolke, C. D. Polymerase-guided base editing enables in vivo mutagenesis and rapid protein engineering. *Nat. Commun.* **12**, 1579 (2021).
- Halperin, S. O. et al. CRISPR-guided DNA polymerases enable diversification of all nucleotides in a tunable window. *Nature* **560**, 248–252 (2018).
- Chen, X. D. et al. Helicase-assisted continuous editing for programmable mutagenesis of endogenous genomes. *Science* **386**, eadn5876 (2024).
- Chu, W. et al. An evolved, orthogonal ssDNA generator for targeted hypermutation of multiple genomic loci. *Nucleic Acids Res.* **53**, gkaf051 (2025).
- Ravikumar, A., Arzumanyan, G. A., Obadi, M. K. A., Javanpour, A. A. & Liu, C. C. Scalable, continuous evolution of genes at mutation rates above genomic error thresholds. *Cell* **175**, 1946–1957 (2018).
- Tian, R. et al. Establishing a synthetic orthogonal replication system enables accelerated evolution in *E. coli*. *Science* **383**, 421–426 (2024).
- Tian, R. et al. Engineered bacterial orthogonal DNA replication system for continuous evolution. *Nat. Chem. Biol.* **19**, 1504–1512 (2023).
- Diercks, C. S. et al. An orthogonal T7 replisome for continuous hypermutation and accelerated evolution in *E. coli*. *Science* **389**, 618–622 (2025).
- Ravikumar, A., Arrieta, A. & Liu, C. C. An orthogonal DNA replication system in yeast. *Nat. Chem. Biol.* **10**, 175–177 (2014).
- Rix, G. et al. Scalable continuous evolution for the generation of diverse enzyme variants encompassing promiscuous activities. *Nat. Commun.* **11**, 5644 (2020).
- Wellner, A. et al. Rapid generation of potent antibodies by autonomous hypermutation in yeast. *Nat. Chem. Biol.* **17**, 1057–1064 (2021).
- Rix, G. et al. Continuous evolution of user-defined genes at 1 million times the genomic mutation rate. *Science* **386**, eadm9073 (2024).
- Blount, Z. D. The unexhausted potential of *E. coli*. *eLife* **4**, e05826 (2015).
- Salas, M., Holguera, I., Redrejo-Rodriguez, M. & de Vega, M. DNA-binding proteins essential for protein-primed bacteriophage $\Phi 29$ DNA replication. *Front. Mol. Biosci.* **3**, 37 (2016).
- Serrano, M. et al. Phage $\Phi 29$ protein p6: a viral histone-like protein. *Biochimie* **76**, 981–991 (1994).

26. Blanco, L. & Salas, M. Characterization and purification of a phage Φ 29-encoded DNA polymerase required for the initiation of replication. *Proc. Natl Acad. Sci. USA* **81**, 5325–5329 (1984).
27. Hermoso, J. M., Mendez, E., Soriano, F. & Salas, M. Location of the serine residue involved in the linkage between the terminal protein and the DNA of phage Φ 29. *Nucleic Acids Res.* **13**, 7715–7728 (1985).
28. Mendez, J., Blanco, L., Esteban, J. A., Bernad, A. & Salas, M. Initiation of Φ 29 DNA replication occurs at the second 3' nucleotide of the linear template: a sliding-back mechanism for protein-primed DNA replication. *Proc. Natl Acad. Sci. USA* **89**, 9579–9583 (1992).
29. Mendez, J., Blanco, L. & Salas, M. Protein-primed DNA replication: a transition between two modes of priming by a unique DNA polymerase. *EMBO J.* **16**, 2519–2527 (1997).
30. Gutierrez, C., Sogo, J. M. & Salas, M. Analysis of replicative intermediates produced during bacteriophage Φ 29 DNA replication in vitro. *J. Mol. Biol.* **222**, 983–994 (1991).
31. Inciarte, M. R., Salas, M. & Sogo, J. M. Structure of replicating DNA molecules of *Bacillus subtilis* bacteriophage Φ 29. *J. Virol.* **34**, 187–199 (1980).
32. Soengas, M. S., Gutierrez, C. & Salas, M. Helix-destabilizing activity of Φ 29 single-stranded DNA binding protein: effect on the elongation rate during strand displacement DNA replication. *J. Mol. Biol.* **253**, 517–529 (1995).
33. Meijer, W. J., Horcajadas, J. A. & Salas, M. Φ 29 family of phages. *Microbiol. Mol. Biol. Rev.* **65**, 261–287 (2001).
34. Reilly, B. E. & Spizizen, J. Bacteriophage deoxyribonucleate infection of competent *Bacillus subtilis*. *J. Bacteriol.* **89**, 782–790 (1965).
35. Camacho, A. et al. Assembly of *Bacillus subtilis* phage Φ 29. 1. Mutants in the cistrons coding for the structural proteins. *Eur. J. Biochem.* **73**, 39–55 (1977).
36. Camacho, A. & Salas, M. Mechanism for the switch of Φ 29 DNA early to late transcription by regulatory protein p4 and histone-like protein p6. *EMBO J.* **20**, 6060–6070 (2001).
37. Salas, M. Protein-priming of DNA replication. *Annu. Rev. Biochem.* **60**, 39–71 (1991).
38. Ito, J. Bacteriophage Φ 29 terminal protein: its association with the 5' termini of the Φ 29 genome. *J. Virol.* **28**, 895–904 (1978).
39. Blanco, L., Lazaro, J. M., de Vega, M., Bonnin, A. & Salas, M. Terminal protein-primed DNA amplification. *Proc. Natl Acad. Sci. USA* **91**, 12198–12202 (1994).
40. Mencia, M., Gella, P., Camacho, A., de Vega, M. & Salas, M. Terminal protein-primed amplification of heterologous DNA with a minimal replication system based on phage Φ 29. *Proc. Natl Acad. Sci. USA* **108**, 18655–18660 (2011).
41. Cetnar, D. P. & Salis, H. M. Systematic quantification of sequence and structural determinants controlling mRNA stability in bacterial operons. *ACS Synth. Biol.* **10**, 318–332 (2021).
42. Farasat, I. et al. Efficient search, mapping, and optimization of multi-protein genetic systems in diverse bacteria. *Mol. Syst. Biol.* **10**, 731 (2014).
43. Salis, H. M., Mirsky, E. A. & Voigt, C. A. Automated design of synthetic ribosome binding sites to control protein expression. *Nat. Biotechnol.* **27**, 946–950 (2009).
44. Blanco, L. & Salas, M. Replication of phage Φ 29 DNA with purified terminal protein and DNA polymerase: synthesis of full-length Φ 29 DNA. *Proc. Natl Acad. Sci. USA* **82**, 6404–6408 (1985).
45. Blanco, L. et al. Highly efficient DNA synthesis by the phage Φ 29 DNA polymerase. Symmetrical mode of DNA replication. *J. Biol. Chem.* **264**, 8935–8940 (1989).
46. Abil, Z. et al. Darwinian evolution of self-replicating DNA in a synthetic protocell. *Nat. Commun.* **15**, 9091 (2024).
47. Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl Acad. Sci. USA* **97**, 6640–6645 (2000).
48. Murphy, K. C. Use of bacteriophage lambda recombination functions to promote gene replacement in *Escherichia coli*. *J. Bacteriol.* **180**, 2063–2071 (1998).
49. Bhattacharyya, S. et al. Accessibility of the Shine–Dalgarno sequence dictates N-terminal codon bias in *E. coli*. *Mol. Cell* **70**, 894–905 (2018).
50. Kudla, G., Murray, A. W., Tollervey, D. & Plotkin, J. B. Coding-sequence determinants of gene expression in *Escherichia coli*. *Science* **324**, 255–258 (2009).
51. Cambray, G., Guimaraes, J. C. & Arkin, A. P. Evaluation of 244,000 synthetic sequences reveals design principles to optimize translation in *Escherichia coli*. *Nat. Biotechnol.* **36**, 1005–1015 (2018).
52. Tian, R. et al. Synthetic N-terminal coding sequences for fine-tuning gene expression and metabolic engineering in *Bacillus subtilis*. *Metab. Eng.* **55**, 131–141 (2019).
53. Robertson, W. E. *Escherichia coli* with a 57-codon genetic code. *Science* **390**, eady4368 (2025).
54. de Vega, M., Lazaro, J. M., Salas, M. & Blanco, L. Mutational analysis of Φ 29 DNA polymerase residues acting as ssDNA ligands for 3'–5' exonucleolysis. *J. Mol. Biol.* **279**, 807–822 (1998).
55. de Vega, M., Lazaro, J. M., Salas, M. & Blanco, L. Primer-terminus stabilization at the 3'–5' exonuclease active site of Φ 29 DNA polymerase. Involvement of two amino acid residues highly conserved in proofreading DNA polymerases. *EMBO J.* **15**, 1182–1192 (1996).
56. Luria, S. E. & Delbruck, M. Mutations of bacteria from virus sensitivity to virus resistance. *Genetics* **28**, 491–511 (1943).
57. Foster, P. L. Methods for determining spontaneous mutation rates. *Methods Enzymol.* **409**, 195–213 (2006).
58. Chen, S. et al. CRISPR–DNA polymerase assisted targeted mutagenesis for regulable laboratory evolution. *Adv. Sci. (Weinh.)* <https://doi.org/10.1002/advs.202511448> (2025).
59. Li, X. T., Thomason, L. C., Sawitzke, J. A., Costantino, N. & Court, D. L. Positive and negative selection using the tetA–sacB cassette: recombineering and P1 transduction in *Escherichia coli*. *Nucleic Acids Res.* **41**, e204 (2013).
60. Rains, C. P., Bryson, H. M. & Peters, D. H. Ceftazidime. An update of its antibacterial activity, pharmacokinetic properties and therapeutic efficacy. *Drugs* **49**, 577–617 (1995).

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

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2026

Methods

Strains and antibiotics

All strains used in this work were derived from *E. coli* DH10B. We genomically integrated the Φ 29 synthetic replication operons, under the control of a PtaIPTG promoter, using lambda Red recombination by the plasmid pEcCas, as described previously^{47,61}. All strains were grown in 2×YT medium. Plasmid and primer sequences are provided in Supplementary Data 3.

Design of the Φ 29 synthetic replication operon

Genes 2, 3, 5 and 6 from bacteriophage Φ 29 were codon-optimized and arranged into a synthetic operon with synthetic ribosome-binding sites using the operon calculator from de novo DNA^{41–43,62–64}. The resultant design was ordered as gBlocks from Integrated DNA Technologies, assembled using overlap extension PCR and cloned into a single-copy plasmid backbone (bacterial F plasmid replication origin) using HiFi Gibson assembly (New England Biolabs) or genomically integrated.

Construction of Φ 29 synthetic replicons

Linear replicons used for electroporation were obtained by PCR using PrimeSTAR Max DNAP (Takara). The initial Φ 29 synthetic replicon was designed to encode tetracycline resistance (*tetA*) and GFP, flanked by the full-length left and right origins of replication from the Φ 29 genome. The primers used for amplifying the replicon annealed at the terminal ends of both origins of replication. The PCR products were purified using a QIAquick PCR Purification Kit (Qiagen).

Construction of circular plasmids

To assist the transformation of the replicon PCR product, we used pFR160 (encoding *gam*) and pFR261 (encoding *gam* and the Φ 29 SSB and DSB). Both plasmids had a CloDF13 origin of replication and a gentamicin resistance marker.

To test the essentiality of the four genes in our Φ 29 synthetic replication operon, we generated plasmids where the genes encoding SSB and/or DSB were disrupted or deleted. Like the intact replication operon, these plasmids had a bacterial F plasmid replication origin and a chloramphenicol resistance gene.

All circular plasmids were constructed using HiFi Gibson assembly (New England Biolabs) from multiple fragments.

Electroporation protocol for establishing Φ 29 synthetic replicons

Colonies of strains harboring the synthetic replication operon, on a single-copy plasmid or the genome, with or without a helper plasmid (pFR160 or pFR261) were inoculated into 10 ml of 2×YT and grown overnight at 37 °C with shaking (220 rpm). Then, 5 ml of this culture was transferred into 250 ml of 2×YT in a 2-L flask and grown at 37 °C with shaking (220 rpm) until an optical density at 600 nm (OD_{600}) of 0.3–0.4 was reached. When helper plasmids were used, we supplemented the medium with 10 mM arabinose at this point and grew the cells for a further 30 min. Subsequently, the cultures were chilled on ice before being pelleted by centrifugation (4,500g, 10 min), washed twice with cold 10% glycerol and resuspended in a final volume of 500 μ l. Then, 100 μ l of the resultant electrocompetent cells were added to an electroporation cuvette (2-mm gap; SLS scientific), the synthetic replicon (PCR product or extract) was added to the cells and electroporation was performed using an Eppendorf Eporator at 2,500 V. Immediately after electroporation, 1 ml of SOB medium was added to the cuvette and the cells were recovered in a 2-ml tube at 37 °C with shaking for 2 h before being plated on an agar plate (2×YT supplemented with the necessary antibiotics). For the initial Φ 29 synthetic replication operon, colonies could be observed after 1–2 days of incubation at 37 °C. For Φ 29-opt, colonies appeared after 16–24 h of incubation. We note that how fast colonies appear might be influenced by what is encoded by the replicon.

Extracting established Φ 29 synthetic replicons from cells

Extracted replicons could be transformed with a much higher efficiency than PCR products and without needing to provide a helper plasmid. To extract these replicons, a 10-ml overnight culture of a strain harboring a Φ 29 synthetic replicon was pelleted by centrifugation (4,500g, 10 min). Next, we used the QIAprep spin miniprep kit to extract the replicon, following the manufacturer's guidelines. We used Millipore-filtered water for the elution and heated the columns, to which the water was added at 55 °C for 10 min before the final centrifugation. All extracted replicons were stored at –20 °C. For the initial Φ 29 synthetic replication operon, colonies could be observed after 1–2 days of incubation at 37 °C. For Φ 29-opt, colonies appeared after 16–24 h of incubation. We note that how fast colonies appear might be influenced by what is encoded by the replicon.

Measuring the stability of the Φ 29 synthetic replicons

To assess the stability of the synthetic replicons, we passaged cells in 2 ml of 2×YT with or without tetracycline (10 μ g ml^{–1}) to select for the replicon. Replicates were passaged in 24-well deep-well plates at 37 °C with shaking. For each passage, we transferred 2 μ l into 2 ml of fresh medium and allowed the cells to grow for at least 16 h to reach the stationary phase. Each 1,000-fold dilution and growth to stationary phase was considered to be ten generations of growth ($\sim 2^{10}$ -fold increase in cells). All samples were analyzed by flow cytometry using a BD LSR-Fortessa cell analyzer (Becton Dickinson) to determine the fluorescence of single cells. Samples were manually gated for single bacterial cells. Gating of green fluorescent cells was set relative to the negative control not expressing GFP. The orthogonal replicon stability was calculated on the basis of the proportion of single GFP-fluorescing cells.

Measuring GFP fluorescence

Cultures to be measured were first pelleted by centrifugation before resuspension in PBS. These samples were then transferred into 96-well flat-bottom clear plates to measure GFP fluorescence ($\lambda_{\text{excitation}}$: 485 nm; $\lambda_{\text{emission}}$: 520 nm) and OD_{600} using a PHERAstar FS plate reader (BMG Labtech). In all cases, we report the GFP fluorescence normalized by OD_{600} .

Measuring Φ 29 synthetic replicon copy number

Cells for which we wanted to measure the replicon copy number were lysed using QuickExtract DNA extraction solution following the manufacturer's guidelines. Briefly, we pelleted 5 μ l of culture, resuspended it in 5 μ l QuickExtract and incubated it at 65 °C for 10 min and 98 °C for 5 min. We diluted this lysate 30-fold with Millipore-filtered water and used it as a template for real-time qPCR to test target DNA copy number. The reference gene was the *dxs* gene on the genome, as it has been shown to be a stable one-copy reference in *E. coli*. The target gene used was GFP or *kanR*. As a control, we used a fused DNA template containing the *dxs* gene with both the *gfp* and the *kanR* genes to make sure that the copy number of the genes was equal (sequence in Supplementary Data 3). We used qPCR to determine the DNA copy numbers for all samples with a Vii7 real-time PCR system with 384-well block (Thermo Fisher).

Primer sequences used:

dxs: 5'-CTTCATCAAGCGTTTCACA-3' and 5'-CGAGAACTGGCGATCCTTA-3'

kanR: 5'-CATGGCAAAGGTAGCGTTGCC-3' and 5'-CCATGCATCATCAGGAGTACG-3'

gfp: 5'-TTATTGCTCAGCGGTGGCAGCAGCCAAC-3' and 5'-ATGGTGCTGCTGGAATTTGTTACCGC-3'

Engineering Φ 29 synthetic replicons using lambda Red recombination

Cells harboring a synthetic replication operon and a replicon to be modified were transformed with plasmid pLF118 encoding the lambda

Red components under the control of an arabinose-inducible promoter⁶⁵. Transformant colonies were then inoculated into 10 ml of 2×YT and grown overnight at 37 °C with shaking (220 rpm). Then, 5 ml of this culture was transferred into 250 ml of 2×YT in a 2-L flask and grown at 37 °C with shaking (220 rpm) until an OD₆₀₀ of 0.3 was reached. We then added arabinose to a final concentration of 10 mM to the culture and grew the cells for a further 30 min. Subsequently, the cultures were chilled on ice before being pelleted by centrifugation (4,500g, 10 min), washed twice with cold 10% glycerol and resuspended in a final volume of 500 µl. Then, 100 µl of the resultant electrocompetent cells were added to an electroporation cuvette (2-mm gap; SLS scientific), the purified PCR product consisting of the gene of interest flanked by at least 30-bp homologies to the existing replicon was added to the cells and electroporation was performed using an Eppendorf Eporator at 2,500 V. Immediately after electroporation, 1 ml of SOB medium was added to the cuvette and the cells were recovered in 10 ml of SOB in a 50-ml tube at 37 °C with shaking for 1 h, transferred to 50 ml of 2×YT supplemented with gentamicin (10 µg ml⁻¹) to select for pLF118 and grown for 1 h at 37 °C with shaking, before being plated on an agar plate (2×YT supplemented with the necessary antibiotics). When determining the efficiency of the recombination, we used 0.5 µg of purified PCR product; however, for other experiments, lower quantities (~50 ng) were also sufficient.

Optimization of the Φ29 synthetic replication operon

To construct a library of the Φ29 synthetic replication operon, we generated a PCR product for each gene in the operon with overlaps to the adjacent gene(s). We used primers that randomized the RBS spacer region, saturated (NNK) the first codon following the ATG start codon and synonymized the following six codons of each gene. We then assembled these genes by overlap extension PCR using primers that amplified the entire operon and appended overlaps to the plasmid backbone and cloned them into a plasmid backbone with a bacterial F plasmid replication origin and a spectinomycin resistance gene using HiFi Gibson assembly (New England Biolabs). The resultant plasmid library was then electroporated into cells harboring the initial Φ29 synthetic replication operon on a chloramphenicol-resistance-conferring plasmid with the same F plasmid origin of replication and a TetA–GFP Φ29 synthetic replicon. We selected for maintenance of the replicon and the plasmid library members by supplementing tetracycline and spectinomycin. To select for operon variants with improved growth, we cultured the cells in 250 ml in 2-L flasks at 37 °C with shaking for three passages (5 ml transferred per passage), before plating the cells for single colonies. We then picked large colonies with bright green fluorescence to ultimately identify the Φ29-opt operon. The Φ29-opt operon was used for all experiments after Fig. 2.

Genomic integration of the Φ29-opt operon

The Φ29-opt operon and the hygromycin resistance marker were PCR-amplified from plasmid pFR329 with primers Gen-integration_opt_F and Gen-integration_opt_R (Supplementary Data 3). The resultant PCR amplicon was then genomically integrated using lambda Red. This was achieved using the same protocol as in the above section detailing Φ29 replicon engineering, except that the strain used was DH10B transformed with pLF118 without a replicon.

Growth measurements

Bacterial colonies were grown overnight at 37 °C in 2×YT with the relevant antibiotics in a 96-well plate. Overnight cultures were diluted 1:100 and monitored for growth in a 200-µl volume in a clear flat-bottom 96-well plate. Measurements of OD₆₀₀ were taken every 5 min on a Tecan Infinite M1000 microplate reader set at 37 °C with shaking. Doubling times were calculated by fitting the data with the logistic growth equation in GraphPad Prism and taking the $\ln 2/k$, where k is the rate constant.

Mutation rate measurements

We used Luria–Delbrück fluctuation analysis to measure the genomic and replicon mutation rates of all DNAPs^{56,57}. We introduced a stop-codon-terminated CmR(Q38TAG) gene into the genome of *E. coli* DH10B. Point mutations that convert the TAG stop codon to sense codons can confer chloramphenicol resistance. The insertion site was adjacent to the *lacI* gene, distal from the origin of replication, where the copy number is expected to be approximately one. We transformed this strain with the plasmids encoding each of the DNAPs to be tested. Untransformed cells were used as a negative control. For the Φ29 DNAPs, the plasmids encoded the entire operon; identical plasmids were used to test the replicon mutation rates. The error-prone T7 DNAP (mut_G) was cloned under the control of a salicylic acid-inducible promoter (pRT307) but leaky (that is uninduced) expression was sufficient to observe substantial mutagenesis of the genome¹⁸. We inoculated transformant colonies into 2 ml of 2×YT supplemented with the necessary antibiotics in 24-well plates and grew the cells overnight at 37 °C with shaking. A total of 12 biological replicates were performed for each strain. All samples were then diluted and plated on plates with or without 20 µg ml⁻¹ chloramphenicol, to measure the proportion of chloramphenicol-resistant cells. FALCOR⁶⁶ was used to calculate the mutation frequency (m) as a function of the cell numbers on the selective and non-selective plates for all 12 biological replicates. The mutation rate per generation per base pair μ was calculated using $\mu = m / (R \times C)$, where the parameter $R = 8/3$ and C (copy number) = 1.

To measure the replicon mutation rates, we generated a KanR–CmR(Q38TAG) replicon by lambda Red recombination. The CmR(Q38TAG) gene used was identical to the one that was genomically integrated. Cells harboring this replicon and a single-copy plasmid (conferring hygromycin resistance) harboring the WT Φ29-opt operon were transformed with a single-copy, spectinomycin resistance-conferring plasmid (encoding the Φ29-opt operon with WT DNAP or the N62D, F65S or N62D;F65S mutants) to replace the initial plasmid. The resultant transformants were inoculated into 2 ml of 2×YT supplemented with the necessary antibiotics (spectinomycin and kanamycin) in 24-well plates and grown overnight at 37 °C with shaking. A total of 12 biological replicates were performed for each strain. All samples were then diluted and plated on plates with or without 100 µg ml⁻¹ chloramphenicol to measure the proportion of chloramphenicol-resistant cells. FALCOR⁶⁶ was used to calculate the mutation frequency (m) as described above.

Measuring mutational spectra using Illumina unique molecular identifier (UMI) consensus sequencing

Four Φ29 DNAPs (WT, F65S, N62D and N62D;F65S) maintaining a KanR–CmR(Q38TAG) replicon were subjected to passaging in at least three biological replicates. At 0 (initial colony from the plate), 20 and 50 generations (1,000-fold dilution of the culture per ten generations), the linear replicon was purified by minimal PCR amplification (Q5 polymerase, New England Biolabs) and prepared for Illumina sequencing with UMIs using the NEB Ultra II FS library prep kit (E7805S, New England Biolabs) and NEB unique dual index UMI adaptors (E7395, New England Biolabs) according to manufacturer's instructions. Libraries were quantified, pooled and sequenced on an Illumina NextSeq2000 with P3, 200-cycle XLEAP reagent kit, reading a 12-mer UMI. A 0.33% PhiX spike-in gave a measured error rate of 0.25%.

The NextSeq2000-generated fastq files were demultiplexed with DRAGEN BCL-Convert (version 4.2.7), separating UMIs with override cycle settings 'Y104;I8U12;I8;Y104'. Reads were trimmed and filtered for quality (cutadapt⁶⁷), aligned to the KanR–CmR(Q38TAG) replicon construct (Bowtie2)⁶⁸, and filtered (SAMtools view)⁶⁹ to keep only unselected regions of the replicon (that is, excluding the *kanR* gene and 100 bp of terminal ends under selection). To generate UMI consensus sequences, UMIs were first summarized using the umi-tools⁷⁰ group function and UMIs with three or more copies were identified. For each

>3-member UMI family, SAMtools consensus was used to generate a consensus sequence. Consensus sequences were then realigned and files containing bases or indels at each position were generated with *igvtools count*⁷¹.

The logic of mutation rate calculation is detailed in Supplementary Fig. 10. In brief, for every DNAP and generation count, files containing bases or indels at each position were processed to generate a mutation frequency for every base mutation (for example, A>C/G/T, insertions and deletions; Ns were ignored). Each specific mutation was processed as follows: the mutation rate per base per generation was the slope of linear regression of mutation frequency across 0, 20 and 50 or 0 and 20 generations. This yielded a set of specific mutation rates per generation for every position, for every mutation, for every DNAP. As this method has no strand information, base-pair substitutions were aggregated (for example, C>A with G>T).

Mutational spectra show the median of positive mutation rates for each substitution mutation and the overall DNAP mutation rate was taken as the mean of all the summed median mutation substitution rates of A, C, G and T.

Continuous evolution of tigecycline, ceftazidime and cefotaxime resistance

Cells harboring a genomic Φ 29-opt synthetic replication operon and a Φ 29 replicon encoding TetA (tigecycline evolution) or TEM-1 β -lactamase (ceftazidime evolution) were transformed with plasmid pFR364 encoding Φ 29-opt with the N62D;F65S error-prone DNAP. Transformant colonies (three or four biological replicates) were inoculated into 2×YT supplemented with the necessary antibiotics (spectinomycin to select for pFR364 and tetracycline or carbenicillin, as appropriate, to select for the replicon). After overnight growth, the resultant cells were tenfold diluted into 5 ml of 2×YT supplemented with spectinomycin and ceftazidime or tigecycline in 24-well plates and incubated at 37 °C with shaking. Tetracycline (10 μ g ml⁻¹) was supplemented throughout the evolution for increasing tigecycline resistance. Carbenicillin (100 μ g ml⁻¹) was supplemented at the lowest concentration of ceftazidime in the ceftazidime evolution but not at subsequent steps. The evolution for cefotaxime resistance began from the final ceftazidime-resistant pools. Ceftazidime and carbenicillin (both 100 μ g ml⁻¹) were supplemented during the evolution for cefotaxime resistance. For each passage, we transferred 500 μ l of culture into 5 ml of fresh medium.

After completing the passages, we plated the cells on agar plates supplemented with tigecycline, cefotaxime or ceftazidime and sequenced individual colonies using Nanopore sequencing of colony PCR products. We also cloned select sequences into a ColE1 plasmid backbone for validation.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Data supporting this study can be found within the article and its Supplementary Information. Sequencing data generated in this study were deposited to the Sequence Read Archive under BioProject PRJNA1335565.

Code availability

Scripts for analyzing mutation rates of polymerases are available from GitHub (<https://github.com/JWChin-Lab/phi29-Mutation-rates>).

References

61. Li, Q. et al. A modified pCas/pTargetF system for CRISPR–Cas9-assisted genome editing in *Escherichia coli*. *Acta Biochim Biophys. Sin. (Shanghai)* **53**, 620–627 (2021).

62. Halper, S. M., Hossain, A. & Salis, H. M. Synthesis success calculator: predicting the rapid synthesis of DNA fragments with machine learning. *ACS Synth. Biol.* **9**, 1563–1571 (2020).
63. Ng, C. Y., Farasat, I., Maranas, C. D. & Salis, H. M. Rational design of a synthetic Entner–Doudoroff pathway for improved and controllable NADPH regeneration. *Metab. Eng.* **29**, 86–96 (2015).
64. Reis, A. C. & Salis, H. M. An automated model test system for systematic development and improvement of gene expression models. *ACS Synth. Biol.* **9**, 3145–3156 (2020).
65. Zurcher, J. F. et al. Continuous synthesis of *E. coli* genome sections and Mb-scale human DNA assembly. *Nature* **619**, 555–562 (2023).
66. Hall, B. M., Ma, C. X., Liang, P. & Singh, K. K. Fluctuation analysis CalculatOR: a web tool for the determination of mutation rate using Luria–Delbruck fluctuation analysis. *Bioinformatics* **25**, 1564–1565 (2009).
67. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **17**, 10–12 (2011).
68. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
69. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
70. Smith, T., Heger, A. & Sudbery, I. UMI-tools: modeling sequencing errors in unique molecular identifiers to improve quantification accuracy. *Genome Res.* **27**, 491–499 (2017).
71. Robinson, J. T. et al. Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).

Acknowledgements

We thank C. Piedrafita and S. Grazioli for helpful discussions in the development of mutation rate analyses. This work was supported by the Medical Research Council UK (MC_U105181009 and MC_UP_A024_1008). F.B.H.R. was supported by a UK Research and Innovation Marie Skłodowska-Curie Actions guarantee fellowship (EP/Y014154/1) and an Investigator Grant (GNT2018461) from the National Health and Medical Research Council Australia.

Author contributions

F.B.H.R., R.T. and K.C.L. performed the experiments. K.C.L. generated and analyzed the next-generation sequencing data. F.B.H.R. conceptualized and established the Φ 29-based orthogonal replication system. J.W.C. supervised the project. F.B.H.R. and J.W.C. wrote the paper with input from all other authors.

Competing interests

The Medical Research Council has filed a provisional patent application related to this work, on which F.B.H.R., R.T. and J.W.C. are listed as inventors. J.W.C. is the founder of and a shareholder in Constructive Bio. K.C.L. declares no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41587-025-02944-x>.

Correspondence and requests for materials should be addressed to Fabian B. H. Rehm or Jason W. Chin.

Peer review information *Nature Biotechnology* thanks Heinz Neumann, Vitor Pinheiro and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

Data analysis

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

Data supporting this study can be found within the article, Supplementary Data and Supplementary Information. Sequencing data generated in this study were deposited at the Sequencing Read Archive under the BioProject ID PRJNA1335565.

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender	na
Reporting on race, ethnicity, or other socially relevant groupings	na
Population characteristics	na
Recruitment	na
Ethics oversight	na

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	The indicated number of biological or technical replicates is shown for each experiment.
Data exclusions	No data was excluded.
Replication	Replicates were generally biological, with at least three replicates. In some cases, as indicated, we report technical replicates. Or technical measurements from pooled biological replicates. All attempts at replication were successful.
Randomization	Samples were not randomized. This is not relevant to the straightforward measurements and characterization of a synthetic genetic system.
Blinding	Blinding was not relevant to our study. Our study consists of straightforward measurements and characterization of a synthetic genetic system.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern
<input checked="" type="checkbox"/>	<input type="checkbox"/> Plants

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Plants

Seed stocks

na

Novel plant genotypes

na

Authentication

na