nature microbiology

Article

https://doi.org/10.1038/s41564-024-01706-w

Engineering plasmid copy number heterogeneity for dynamic microbial adaptation

Received: 27 April 2023

Accepted: 19 April 2024

Published online: 18 June 2024

Check for updates

Shalni Kumar[1,4] ✉, Andrew Lezia[1,4] & Jeff Hasty[1,2,3]

Natural microbial populations exploit phenotypic heterogeneity for survival and adaptation. However, in engineering biology, limiting the sources of variability is a major focus. Here we show that intentionally coupling distinct plasmids via shared replication mechanisms enables bacterial populations to adapt to their environment. We demonstrate that plasmid coupling of carbon-metabolizing operons facilitates copy number tuning of an essential but burdensome construct through the action of a stably maintained, non-essential plasmid. For specific cost–benefit situations, incompatible two-plasmid systems can stably persist longer than compatible ones. We also show using microfluidics that plasmid coupling of synthetic constructs generates population-state memory of previous environmental adaptation without additional regulatory control. This work should help to improve the design of synthetic populations by enabling adaptive engineered strains to function under changing growth conditions without strict fine-tuning of the genetic circuitry.

Heterogeneous gene expression in cell populations is typically viewed as a challenge to synthetic biologists; however, it is pervasive in natural systems[1]. Across all domains of life, phenotypic variation aids population adaptation. Even genetically homogenous bacterial populations strategically rely on heterogeneity to improve population survival via division of labour or bet hedging—a strategy where some cells have better fitness in potential future environments compared with the current one[2,3]. Division of labour is often used by biofilms to balance growth versus maintenance[4], and bet hedging is associated with spore formation, carbon metabolism and antibiotic persistence[5]. The widespread use of population heterogeneity in nature suggests that synthetic gene circuits could exploit variability to adapt to different environmental conditions and carry out complex functions[6,7]. This strategy could help respond to the long-standing need for the development of synthetic communities that are robust to changing environments across fields such as biomanufacturing and biosensing[8].

One common source of noise in synthetic biology is plasmid copy number (PCN) variation. Previously overlooked for building circuits, PCN control has now been demonstrated, including creating strain libraries with varying copy numbers (CNs)[9], designing inducible PCNs[10–12] and building a CN oscillator[13]. In pioneering work, Tang and Liu[14] used two plasmids with coupled CNs to generate sustained ratiometric readouts of environmental events. In their study, small-molecule induction led to CRISPR-mediated plasmid targeting, which directly altered the CN ratio between two plasmids. In our work, we also employ plasmid coupling, but to generate phenotypic heterogeneity within a bacterial population.

PCN is natively controlled by built-in negative feedback that inhibits replication when the CN exceeds a set point[15]. The widely used ColE1 origin of replication uses an antisense RNA to selectively inhibit replication priming[16]. Since plasmids with the same origin of replication share their regulation strategy, they cannot be distinguished during replication, which increases CN heterogeneity. Although it is widely accepted that plasmids with the same origin of replication are unstable, research suggests that some can persist at a high rate (near 100%) over multiple growth cycles without selection[17]. In addition, previously determined

[1]Department of Bioengineering, University of California, San Diego, La Jolla, CA, USA. [2]Molecular Biology Section, Division of Biological Sciences, University of California, San Diego, La Jolla, CA, USA. [3]Synthetic Biology Institute, University of California, San Diego, San Diego, CA, USA. [4]These authors contributed equally: Shalni Kumar, Andrew Lezia. ✉e-mail: s8kumar@ucsd.edu

https://doi.org/10.1038/s41564-024-01706-w

plasmid loss rates may be overestimated due to overshadowing of inherent loss rates by fitness effects[18].

In this work, we show that two-plasmid systems with shared origin types can create population heterogeneity and enable environmental adaptation (Fig. 1). We show that such populations rapidly adapt their CNs to stressors such as antibiotics and engineered cell lysis. We also show that in these shared-origin systems a non-essential plasmid can act as a tuning knob for the CN of an essential plasmid—a concept we name plasmid buffering. Under certain environments, plasmid buffering leads to increased stability for shared-origin plasmids compared with compatible ones, contradicting conventional wisdom that plasmid incompatibility always leads to faster plasmid loss. We also quantify the dynamics of CN adaptation for coupled plasmids of varied origin types and demonstrate retained memory of population CN following transient environmental selection, particularly for high-copy origins. Overall, this work shows how plasmid coupling can create heterogeneity in a bacterial population, enabling CN adaptation to different environments.

Results

Duplicate origins create selectable population heterogeneity

To test whether plasmids with shared origins have coupled CNs and generate phenotypic heterogeneity, we designed two plasmids, one carrying chloramphenicol resistance (CmR) and constitutive red fluorescent protein (RFP) and the other carrying spectinomycin resistance (SpecR) and green fluorescent protein (GFP) (Fig. 2a).

To confirm that each plasmid was distributed equally across the population under non-selective conditions, we performed quantitative PCR (qPCR). The relative read abundance of fluorescent targets divided by the origin showed that in the duplicate-origin case the CN of each plasmid was approximately half of the total CN (Fig. 2b). A single-plasmid control carrying both resistance and fluorescent markers was used as a reference.

Next, we measured the underlying, single-cell CN distribution of duplicate-origin and different-origin systems. Conceptually, duplicate origins should have unconstrained CNs between two plasmid types, allowing a wide distribution of phenotypes, whereas different-origin systems should be constrained by each origin type's replication control. We compared the ColE1 duplicate-origin strain (Dup_Ori) with a control strain with an RFP ColE1 plasmid and a GFP sc101 plasmid (Diff_Ori) (Fig. 2a). To approximately match the ColE1 plasmid's CN, all Diff_Ori strains were made using an M78I sc101-origin variant[19]. To obtain single-cell fluorescence measurements, we imaged 3 µl of 1:10 diluted culture under an agarose pad. In both the Dup_Ori strain and the Diff_Ori strain, we saw considerable variability in GFP and RFP expression. For the Diff_Ori strain, GFP and RFP expression were well correlated ($R^2 = 0.91$), suggesting that variability was mainly due to overall differences in cell state affecting both plasmids similarly[20] (Fig. 2c). Conversely, for the Dup_Ori strain, GFP and RFP expression showed no correlation ($R^2 = -0.80$), suggesting that a shared replication mechanism between the plasmids leads to divergent GFP and RFP expression across the population (Fig. 2d).

We then developed a probabilistic model of plasmid replication and partitioning to evaluate whether a theory of shared replication inhibition was sufficient to generate the above results (Fig. 2e). Internal cell variables carrying PCNs for each plasmid type affected the death and division probability of each cell at every time step. Equations describing plasmid-encoded effects are described in the Methods. Plasmids were partitioned between daughter cells by equal probability binomial trials. First, we simulated the growth of a duplicate-origin and different-origin population, as was done in the experiments above. Under default parameters (Supplementary Table 4), our model recapitulated the divergent heterogeneity found experimentally with duplicate-origin plasmids versus different-origin controls (Fig. 2f).

We then tested whether heterogeneous Dup_Ori populations could shift their relative PCN ratio in response to different environmental

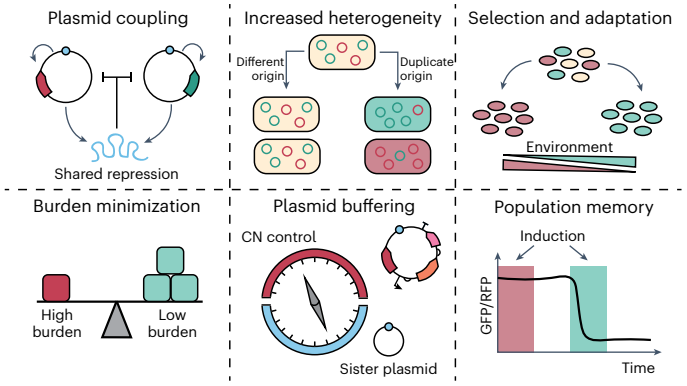


Fig. 1 | PCN heterogeneity can drive fitness-mediated population adaptation. Two plasmids are coupled via shared CN regulation mechanisms. This coupling leads to increased phenotypic heterogeneity in a population. Different cells are selected for under varied environments due to fitness advantages. Systems with two incompatible plasmids allow for burden minimization of high-cost plasmids. They also enable buffering of an essential plasmid of interest by a non-essential, minimal sister plasmid. Lastly, coupled plasmid systems can maintain population memory following selection.

conditions. We cultured the Dup_Ori strain in 50× the working concentration of chloramphenicol or spectinomycin (1 and 3 mg ml^{-1}, respectively). As expected, the population shifted its CN distribution towards the favourable plasmid, as determined by qPCR (Fig. 2g). Measurements of bulk GFP/RFP fluorescence followed the same trend as qPCR. While direct measurement by qPCR or sequencing is the gold standard[21], validation of fluorescence as a suitable metric for relative CN change allows investigation with higher throughput and time-lapse microscopy. In response to different antibiotic concentrations, the Diff_Ori strain had no clear trend in GFP/RFP fluorescence (Extended Data Fig. 1). In summary, we showed that plasmids with the same origin have coupled CNs leading to increased heterogeneity and fitness-based environmental selection.

Plasmid buffering decreases burden and increases stability

Metabolism in *Escherichia coli* is tightly regulated by the composition of the surrounding environment, and relevant enzymes have complicated cost–benefit relationships. For instance, meticulous genetic experiments revealed that lac permease activity drives cost for *lac* operon expression[22]. To investigate whether duplicate-origin strains can use CN flexibility to tune metabolic operon expression in response to altered nutrient availability, we built a Dup_Ori_Met strain. This strain carries the arabinose operon (araOp) and RFP coupled to the lactose operon (lacOp) and GFP within a JS006 host that has the genomic operons knocked out[23] (Fig. 3a and Extended Data Fig. 2a).

We naively expected the population to shift its CN in favour of the plasmid metabolizing the present carbon source; however, we saw the opposite trend. Specifically, when grown on arabinose minimal media without antibiotic, we observed increased GFP/RFP relative to that with glucose or lactose (Fig. 3b). A control strain with araOP on sc101 (Diff_Ori_Met) followed the same trend at a smaller scale, whereas Dup_Ori_Met plasmids in an MG1655 host with functional genomic operons showed pronounced adaptation. These results suggest that the burden induced by operon expression at high CNs drives adaptation rather than additional benefit.

Next, we cultured the Dup_Ori_Met strain in a microfluidic device[24] to dynamically switch the carbon source (Fig. 3c). Cells were cultured without antibiotics in M9 glucose, followed by lactose and then arabinose. Traps contained a majority of RFP-dominant cells when grown on lactose, indicating higher araOP relative to lacOP CNs. Once switched to arabinose, growth temporarily halted in all traps, but resumed in ~15% of

https://doi.org/10.1038/s41564-024-01706-w

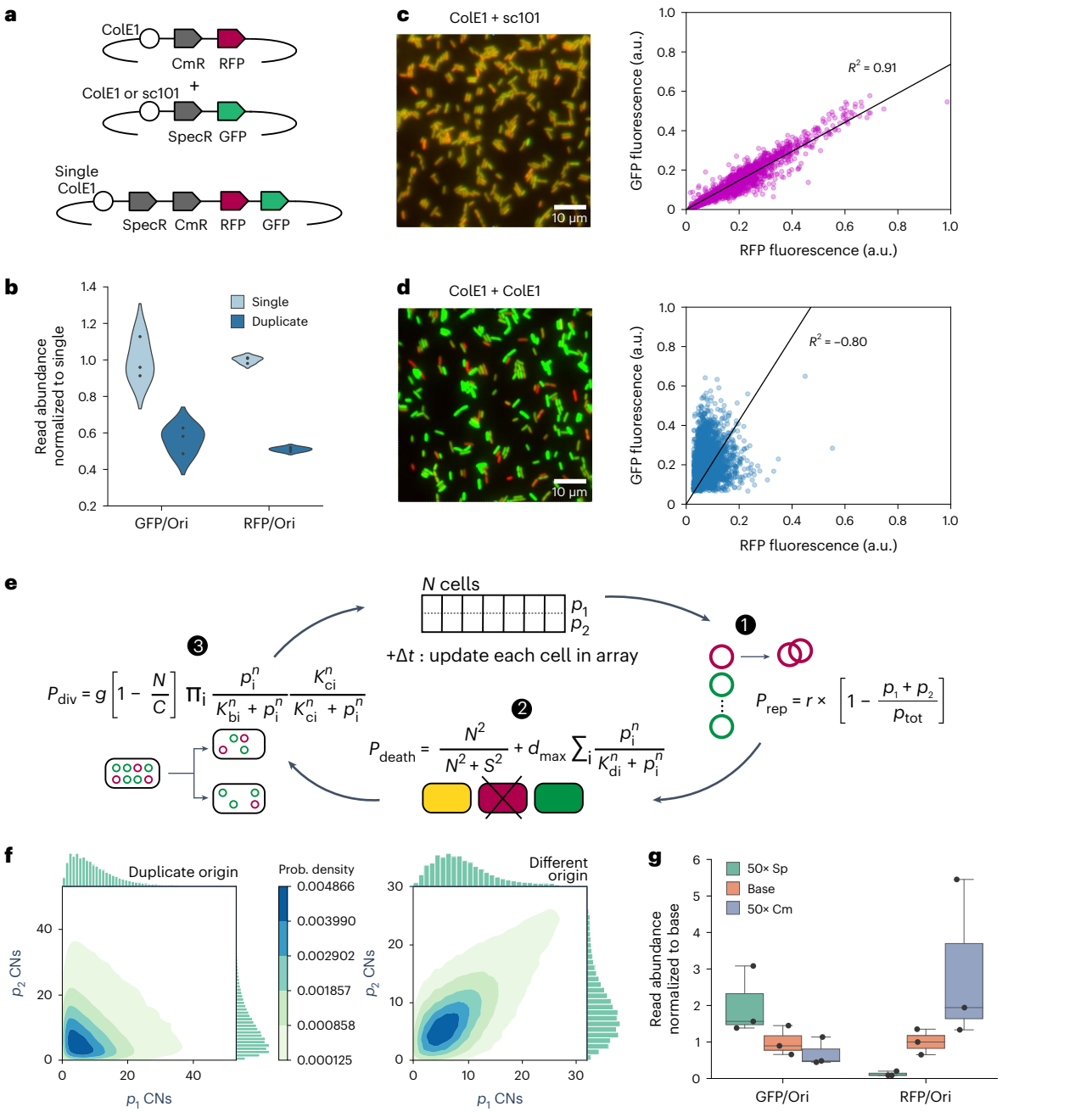


Fig. 2 | Two-plasmid systems with duplicate origins show increased heterogeneity and adaptability. a, Diagram of the plasmids used to look at CN behaviour with altered origin combinations. **b**, Relative CNs of fluorescent gene targets as a fraction of total origin CN for a duplicate-origin two-plasmid population versus a single-plasmid population, as determined by qPCR ($n = 3$ biological replicates). **c**, Fluorescence micrograph of cells with different-origin plasmids (left) and a scatter plot showing the single-cell GFP/RFP distribution for the population (right) ($n = 2{,}090$ cells from three cultures). **d**, Fluorescence micrograph of cells with duplicate ColE1-origin plasmids (left) and a scatter plot showing the single-cell GFP/RFP distribution for the population (right) ($n = 2{,}629$ cells from three cultures). R^2 values describe a linear fit of the data, constrained to a zero intercept. **e**, Schematic of a probabilistic model in which cells undergo plasmid replication, cell death and cell division based on individual cell plasmid numbers. **f**, Results from a model simulation showing the CN distribution for duplicate-origin pairs and different-origin pairs ($n_{Sim} = 50$). This simulation was run with $g = 1$, $d = 0.5$ and the default parameters shown in Supplementary Table 4. **g**, Relative PCN fractions for GFP and RFP plasmids within the duplicate strain when grown with varying antibiotic concentrations, normalized to 0.5× the base antibiotic condition ($n = 3$ biological replicates). The boxplots show the minimum, 25th percentile, median, 75th percentile and maximum values. Cm, chloramphenicol; Ori, origin; Prob., probability; Sp, spectinomycin.

40 traps, with GFP-dominant subpopulations rapidly taking over (Supplementary Video 1). It is possible that traps that did not resume growth did not retain subpopulations of GFP-biased individuals following lactose incubation and were thereby unable to adapt to the arabinose condition. This phenomenon probably occurs in overburdened small populations without enough heterogeneity to survive the secondary environment. Larger populations, such as those analysed for bulk fluorescence in Fig. 3b showed consistent adaptation.

Naturally, each operon is found in the genome at one copy and is therefore probably expressed optimally at low CNs. We hypothesized that cells with a higher CN of the essential plasmid had considerable burden when the operon was turned on compared with cells with lower

CNs of the essential plasmid and high CNs of the non-essential and transcriptionally off plasmid. Here a non-essential plasmid is one that is not needed for cell growth in a specified media type. For example, when the Dup_Ori_Met strain is grown in arabinose, it requires at least one copy of the arabinose operon to survive, but selects for cells with high lacOp PCNs. Since the lacOp plasmid genes are turned off when lactose is not present, this plasmid acts as a buffer to drive down the CN of the more burdensome araOp plasmid. We call this feature of coupled plasmid systems plasmid buffering.

To validate that a low CN of the araOp plasmid improves growth in arabinose media, we first built a version of our ColE1 araOp plasmid with direct anhydrotetracycline (aTc)-inducible CNs (Extended Data Fig. 2b)[10]. This plasmid has a TetO operator in its pUC origin, a single-base pair variant of ColE1 and repressed replication in the absence of aTc. When transformed into *E. coli* MG1655 Z1 (a constitutive tetR strain), induction of increased CN led to a dose-dependent decrease in the growth rate, suggesting that high CNs produce a large burden that is strongly enhanced in arabinose media (Fig. 3d). To directly measure burden from the same araOP plasmid used in the Dup_Ori_Met strain, we also co-transformed a burden-reporting construct that expresses GFP under the *htpG1* stress promoter into our strains[25]. The original ColE1 AraOP plasmid within JS006 cells reported a higher burden in arabinose media than the same araOP on a low CN origin (Extended Data Fig. 2c). Lastly, we visualized the single-cell distribution and growth of the Dup_Ori_Met strain within a mother machine device[26] when pre-conditioned and grown in arabinose media (Methods). The fluorescence and division frequency of cell lineages indicated that the vast majority of cells were green shifted and the fastest dividers were optimized to a small range of moderately high GFP/RFP fluorescence values (Extended Data Fig. 2d). Together, these results strongly suggest that the Dup_Ori_Met strain uses plasmid buffering to balance the burden of an essential yet costly carbon-metabolizing plasmid.

In synthetic biology applications, circuit expression is often not optimized for a given environment. For the Dup_Ori_Met strain, expression of the native lactose operon was too costly. To explore whether weakened lacOP variants result in altered CN adjustment, we mutated the pLac or lacI regions of the plasmid to decrease expression[27]. After co-transformation with araOP, the resulting Dup_Ori_Met strains were grown in various concentrations of lactose (Fig. 3e). Weakened lacOP expression led to increased GFP/RFP, demonstrating decreased cost relative to the original variant. Across all strains, as the lactose concentration decreased, GFP/RFP increased, suggesting a greater benefit for more copies of the lacOP plasmid. Noticeably, the pLac mutant showed a non-monotonic response at intermediate lactose concentrations, potentially due to secondary effects of mutation on promoter bistability.

Despite conventional wisdom, our results suggest that plasmid buffering could stabilize duplicate-origin plasmids. Specifically, we hypothesized that Dup_Ori_Met grown on arabinose without antibiotics would retain the lacOP plasmid due to its role in buffering araOP CN. When grown in glucose Dup_Ori_Met lost the lacOP plasmid faster than the Diff_Ori_Met control, as was classically expected, but when grown on arabinose Dup_Ori_Met had significantly higher lacOP plasmid retention than the control (Fig. 3f). We did not see increased stability for the araOP plasmid when Dup_Ori_Met was cultured in lactose, potentially due to the lower cost and preference for lactose in the carbon hierarchy[28].

To explore conditions for which plasmid buffering enhances stability, we used our model to simulate plasmid distribution with varied plasmid-encoded costs. Cost is described by a Hill function impacting cell division, where K_c is the PCN for half-maximal growth reduction (Methods and Supplementary Table 4). A larger K_c value thereby corresponds to a lower plasmid-encoded cost. We extracted the retention time, or the time when less than 5% of the population contained any non-essential plasmid, as a function of each plasmid's K_c. Our results confirmed a region of enhanced stability for duplicate-origin strains when the essential plasmid cost is high (low K_c1) and the non-essential plasmid cost is moderate to low (Fig. 3g). For different-origin systems, non-essential plasmid stability is mainly a function of its own cost. To test the predictive power of our model, we experimentally moved across parameter space through the addition of isopropyl β-D-1-thiogalactopyranoside (IPTG)—a metabolically inactive inducer of lacOP. The addition of 100 µM IPTG to arabinose-passaged cultures resulted in increased loss of the lacOP plasmid in both Dup_Ori_Met and Diff_Ori_Met strains (Fig. 3h). Our model suggests that IPTG induces high lacOP burden in arabinose media, lowering its stability in Dup_Ori_Met and Diff_Ori_Met strains. The addition of IPTG to lactose-passaged cultures showed a small decrease in Dup_Ori_Met stability and an increase in Diff_Ori_Met stability. Based on our model, these stability changes can only be described by IPTG decreasing the lacOP cost in lactose media. Although not intuitive, this finding agrees with previous work showing that IPTG competition for lacY permease can decrease the metabolic burden in the presence of high lactose[22].

Our model serves as a useful tool with which to interpret experimental results that probe the cost–benefit relationships of two metabolic operons. It also suggests that the CN response decreases in magnitude but increases in speed with lower essential plasmid cost, highlighting the impact of population growth on response time scales (Fig. 3i). In these parameter sweeps, we proportionally decreased g and d rates relative to r, to prevent plasmid dilution from rapid cell turnover at very low costs.

Fig. 3 | Duplicate-origin plasmids allow burden-based adaptation to different nutrient environments and improve plasmid stability. a, Circuit diagram for metabolic operon plasmids, with constitutive GFP on the lactose operon-expressing plasmid and RFP on the arabinose operon-expressing plasmid. In the Diff_Ori_Met strain, the arabinose operon plasmid has the sc101 origin. **b**, Plate reader fluorescence for Dup_Ori_Met, Diff_Ori_Met and Dup_Ori_Met_MG1655 strains grown on different carbon sources. The error bars represent means ± s.d. of $n = 3$ cultures. **c**, Representative fluorescence micrographs for the Dup_Ori_Met strain grown under microfluidics. The carbon source was switched from lactose to arabinose at $t = 10.5$ h. **d**, Growth rate estimates for the ipUC-Ara strain whose CN is directly induced with aTc ($n = 3$ cultures). The boxplots show the minimum, 25th percentile, median, 75th percentile and maximum values. **e**, Plate reader fluorescence for different versions of the Dup_Ori_Met strain with mutated *lac* operons under gradient concentrations of lactose ($n = 3$ cultures). The error bars represent means ± s.d. **f**, Plasmid retention data for the Dup_Ori_Met (D) and Diff_Ori_Met (C) strains following 5 d of antibiotic-free passaging in arabinose (A) or glucose (G) ($n = 3$ biological replicates). A two-sided t-test with independent samples and Bonferroni correction yielded P values of $^{***}9.384 \times 10^{-4}$ and $^{**}1.563 \times 10^{-3}$ and t statistics of 1.048×10^{1} and −9.181, respectively, for the duplicate versus control comparison with arabinose and glucose. **g**, Simulation of non-essential plasmid retention for duplicate-origin and different-origin plasmids. The colour map corresponds to the number of time steps with the plasmid still retained out of 1,000. The simulation was run with $g = 1$, $d = 0.5$ and the default parameters shown in Supplementary Table 4. **h**, Plasmid retention data as in **f**, but with the addition of 100 µM IPTG and growth in arabinose (A) or lactose (L) for 3 d ($n = 3$ biological replicates). A two-sided t-test with independent samples and Bonferroni correction yielded P values of $^{*}1.966 \times 10^{-2}$, $^{****}4.071 \times 10^{-5}$, $^{***}2.155 \times 10^{-4}$ and $^{*}2.063 \times 10^{-2}$ and t statistics of 7.498, 6.005×10^{1}, 1.818×10^{1} and −5.550, respectively, for the $IPTG^{+}$ versus $IPTG^{-}$ comparison in the duplicate strain with arabinose, the control strain with arabinose, the duplicate strain with lactose and the control strain with lactose. **i**, Parameter sweeps of the effect of changing plasmid-encoded metabolic cost (K_{c1}) on CN change magnitude and speed. These simulations were run with $g = 0.5$, $d = 0.25$ and the default parameters shown in Supplementary Table 4. Const., constitutive; WT, wild type.

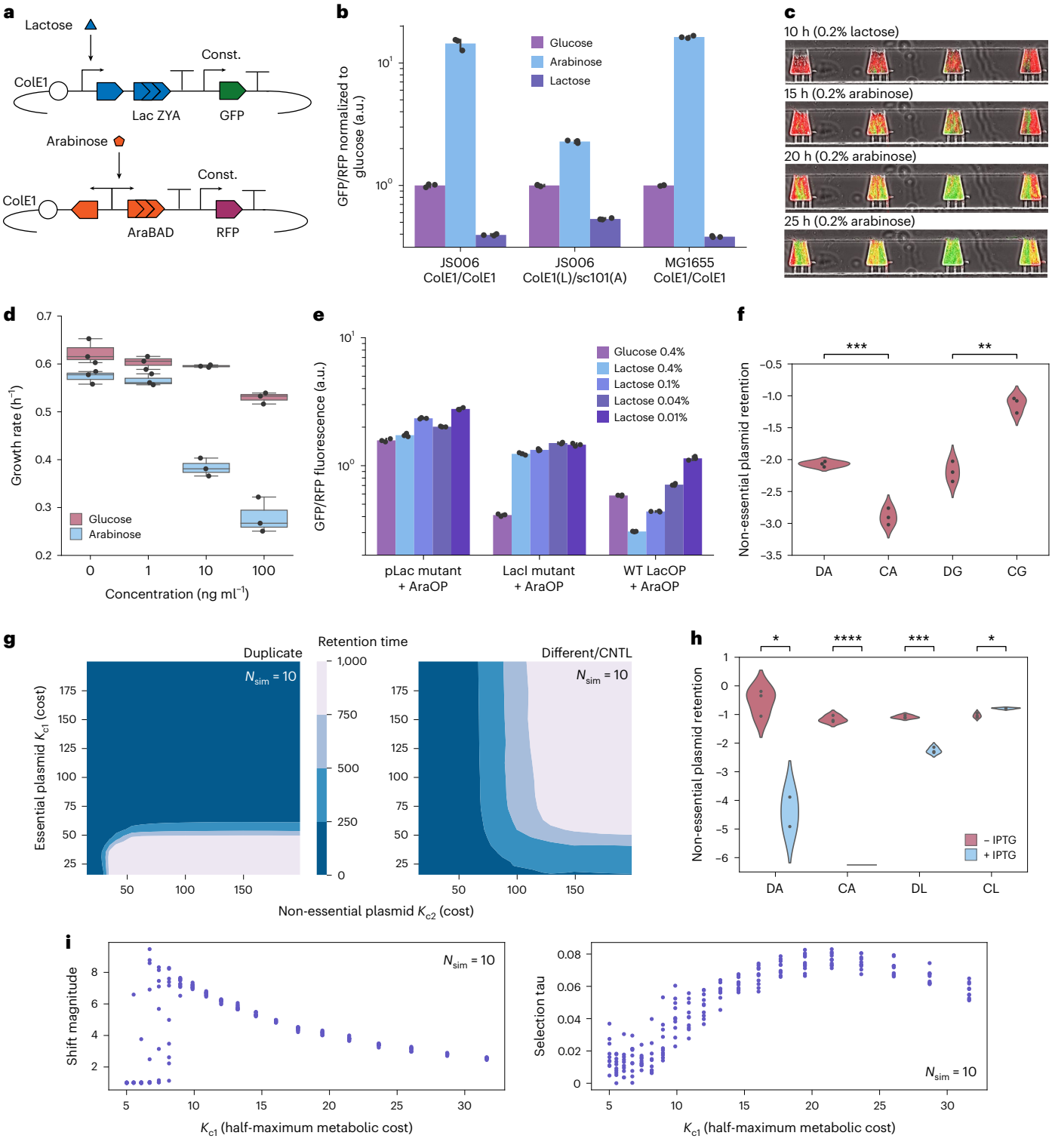
https://doi.org/10.1038/s41564-024-01706-w

In summary, we showed that coupled carbon-metabolizing plasmids optimize their CN distribution based on environmentally induced burden and can have increased stability over compatible origins due to plasmid buffering.

Minimal sister plasmid allows reversible expression tuning

Next, we hypothesized that a low-cost sister plasmid coupled to an arbitrary synthetic gene circuit could promote CN buffering to sensitively match environmental need.

To test this idea, we co-transformed a minimal ColE1-origin plasmid with a ColE1 plasmid carrying acyl-homoserine lactone (AHL)-inducible kanamycin resistance (KanR) and constitutive GFP to make the strain Dup_Ori_Kan (Fig. 4a). The cost–benefit relationship of this plasmid can be directly varied bidirectionally by inducing KanR expression with AHL and manipulating kanamycin concentrations in the media. We then grew Dup_Ori_Kan in media containing a 1× to 15× kanamycin working concentration and 0–100 nM AHL. We saw that as kanamycin increased in the media, the KanR plasmid's CN also increased, as represented by increased GFP divided by optical density (OD). Conversely, as the expression of KanR was increased via AHL, the CN decreased to compensate for excess KanR expression (Fig. 4b). After passaging into 1× kanamycin media, we then incubated



https://doi.org/10.1038/s41564-024-01706-w

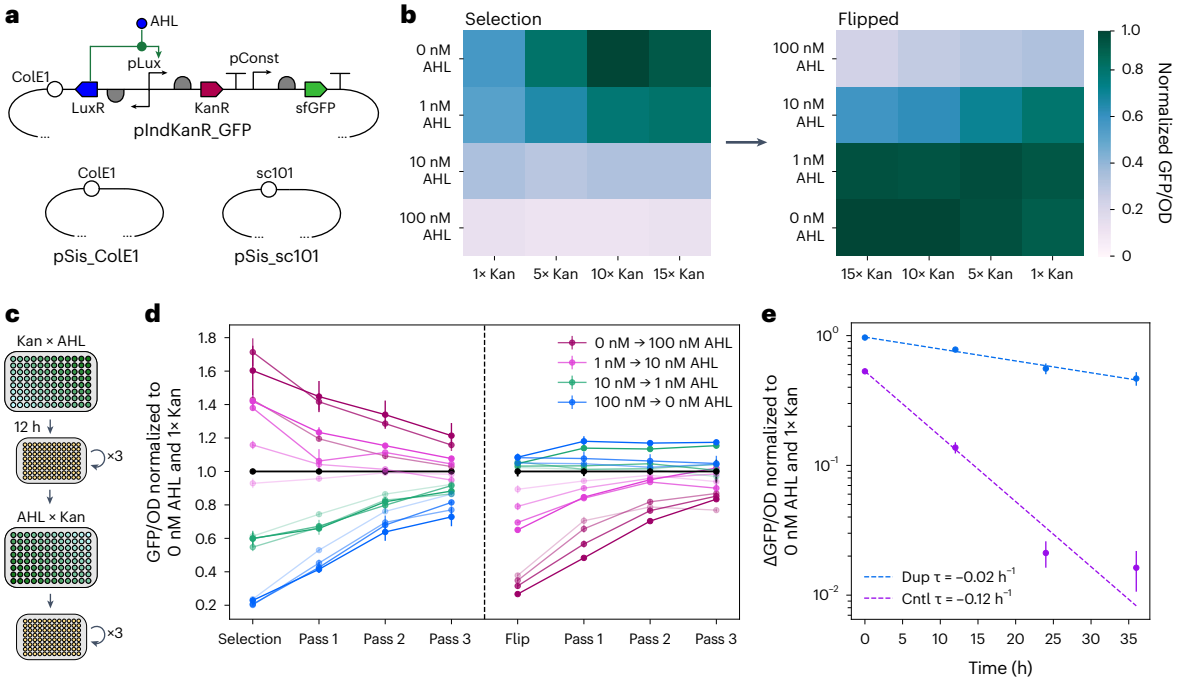


Fig. 4 | Incompatible sister plasmid buffering enables population adaptation and promotes long-term memory. a, Circuit diagram of the AHL-inducible KanR plasmid and accompanying minimal sister plasmids. The Dup_Ori_Kan strain contains the ColE1 sister plasmid, whereas the Diff_Ori_Kan strain carries the sc101 sister plasmid. **b**, Left, heatmap showing plate reader fluorescence data for the Dup_Ori_Kan strain under different concentrations of AHL and kanamycin (Kan). Each square represents the mean of three separate cultures. Right, heatmap showing the fluorescence of cells following passage into flipped inducer conditions. **c**, Diagram depicting the experimental workflow for **b** and **d**. **d**, Plate reader fluorescence measurements for Dup_Ori_Kan in all AHL × Kan combinations (first selection), followed by measurements during three non-selective passages, secondary flipped selection and three more non-selective passages. The line opacity represents the environmental kanamycin concentration with the highest opacity, corresponding to 15× kanamycin. The error bars represent means ± s.d. ($n = 3$ biological replicates). **e**, Representative plot of memory time constants for the Dup_Ori_Kan and Diff_Ori_Kan strains during passaging in non-selective media. The fluorescence measurements for three separate cultures are fit to an exponential curve to estimate *tau*. The error bars represent means ± s.d. OD, optical density.

the cultures in the reverse conditions of the initial selection. For example, the same cells initially incubated in 15× kanamycin and 0 nM AHL were now diluted into 1× kanamycin and 100 nM AHL. The resulting GFP/OD measurements matched those expected, showing reversibility and secondary adaptation of Dup_Ori_Kan. When this experiment was repeated for a control strain with an sc101-origin sister plasmid (Diff_Ori_Kan), we saw decreased-magnitude responses of GFP/OD to different selection conditions (Extended Data Fig. 3a). Single ColE1 plasmids still exhibit inherent CN heterogeneity, which can lead to partial environmental response within control strains[29,30]. We predicted, however, that the Dup_Ori_Kan strain would exhibit longer CN memory than the Diff_Ori_Kan strain upon removal of high selective pressure.

In the aforementioned experiment, GFP/OD was also measured over three non-selective passaging time points after each selection. Each passage was a 1:100 dilution into 1× kanamycin and 0 nM AHL media (Fig. 4c). Rather than immediately returning to baseline once selection was removed, the Dup_Ori_Kan population retained memory for many generations (Fig. 4d). The memory time scale for Dup_Ori_Kan fluorescence change (0.02 h^{-1}) was considerably longer than for Diff_Ori_Kan (0.12 h^{-1}), where any CN adjustment was rapidly lost upon removal of selection (Fig. 4e). Noticeably, the secondary selection of the Dup_Ori_Kan strain in 15× kanamycin and 0 nM AHL showed a decreased GFP/OD response as compared with the initial selection for this condition (Fig. 4d). Additionally, the size difference between the KanR plasmid and sister plasmid could lead to altered replication rates and thereby asymmetric time scales necessary for adaptation in each direction (Extended Data Fig. 3b).

In summary, a shared-origin sister plasmid enables the reversible CN tuning of a coupled synthetic construct based on its fitness under a specific environment and can enhance population memory.

Duplicate origins from varied classes promote memory

To further explore memory dynamics in an easily tunable duplicate-origin system with rapid selection, we coupled two inducible lysis plasmids (Dup_Ori_Lys). Each plasmid transcribes E lysis from phage ϕX174 in response to AHL or arabinose and expresses GFP or RFP, respectively (Fig. 5a).

After demonstrating CN adaptation of Dup_Ori_Lys against lysis pressure in plate reader experiments (Extended Data Fig. 4a), we then tested the strain in a previously described microfluidic chip along a gradient of arabinose and AHL concentrations (Fig. 5b, Supplementary Video 2 and Extended Data Fig. 4b). The device houses monolayer cell traps of different sizes that are fed by eight discrete media conditions formed from mixing two inlet sources[31]. In continuous culture, the Dup_Ori_Lys strain again shifted CN against lysis pressure, with GFP- and RFP-dominant cells occupying high arabinose (0.017%) and AHL (100 nM) environments, respectively (Fig. 5b). We noticed that larger traps showed faster and more complete adaptation to inducers due to increased population size, and validated these results by simulation (Extended Data Fig. 4c). The control strain (Diff_Ori_Lys) with the pAra_lyse plasmid on an sc101 origin did not noticeably shift CN along gradient arabinose and AHL conditions (Extended Data Fig. 4d).

We had found that the pAra_Lyse plasmid by itself had substantial leaky lysis when uninduced (Extended Data Fig. 5a). To investigate the sensitivity of CN adaptation to lysis strength, we developed a library of pAra_Lyse plasmids with differing ribosome binding sites driving E lysis (Fig. 5c). Library members were induced with 0.02% arabinose to assess the lysis strength before co-transformation with the original pAHL_Lyse plasmid (Extended Data Fig. 5b and Methods). The GFP/RFP ratios of the eight strains tested clearly showed that decreased arabinose lysis strength leads to increased relative RFP (Fig. 5d and Extended Data Fig. 5c).

https://doi.org/10.1038/s41564-024-01706-w

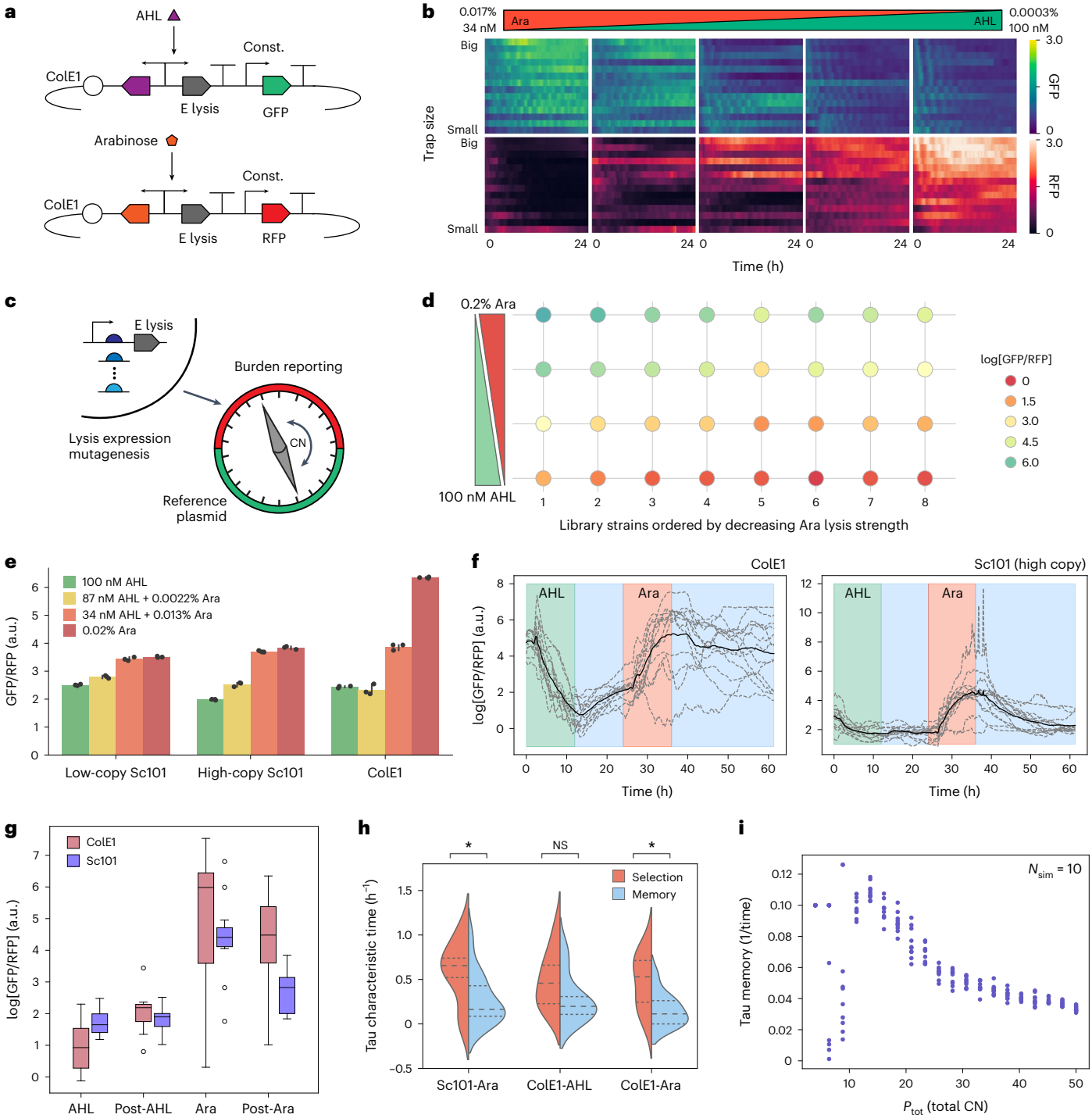


Fig. 5 | Inducible cell lysis triggers population CN shifting in duplicate-origin systems from commonly used replication classes. a, Circuit diagram for inducible lysis plasmids with constitutive GFP on the AHL-inducible lysis plasmid and RFP on the arabinose-inducible lysis plasmid. In the Diff_Ori_Lys strain, the arabinose lysis plasmid has the sc101 origin of replication. **b**, Heatmap showing mean fluorescence over time for the Dup_Ori_Lys strain in a microfluidic device with different inducer concentrations and trap sizes ($n = 14$ traps per media condition). **c**, Diagram depicting the construction of the Dup_Ori_Lys strain library with variant arabinose lysis strength for subsequent burden-reporting experiments. **d**, Mean plate reader fluorescence of the Dup_Ori_Lys strain library, ordered by arabinose lysis strength across four inducer conditions ($n = 3$ cultures). **e**, Plate reader fluorescence for three Dup_Ori_Lys strains, each with a different-origin set, across four inducer conditions ($n = 3$ cultures). The error bars represent means ± s.d. **f**, Log ratio fluorescence over time for two Dup_Ori_Lys strains with different-origin sets, cultured in parallel within a microfluidic device with four induction windows: 100 nM AHL, 0.2% glucose, 0.02% arabinose and glucose again ($n = 11/14$ traps). **g**, Mean fluorescence of each Dup_Ori_Lys strain at the end of each induction window for $n = 11$ (ColE1) or 14 (Sc101) individual traps. The boxplots show the minimum, 25th percentile, median, 75th percentile and maximum values, removing outliers outside of the interquartile range by more than 1.5×. **h**, Selection and memory time constants extracted from microfluidic experiments, as explained in the Methods. The horizontal lines within violin plots show median and quartile values. A two-sided paired t-test with Bonferroni correction yielded P values of $^{*}2.454 \times 10^{-2}$, 2.940×10^{-1} and $^{*}1.716 \times 10^{-2}$ and t statistics of 3.117, 1.846 and 3.603, respectively, for selection versus memory in Sc101-Ara, ColE1-AHL and ColE1-Ara. **i**, Parameter sweep showing the effect of total CN on memory time scales in simulation. Lysis strength is rescaled so as not to change with P_{tot} in this parameter sweep. This simulation was run with $g = 0.5$, $d = 0.25$ and the default parameters shown in Supplementary Table 4. Ara, arabinose; NS, not significant.

Next, we sought to generalize our findings to plasmid origins from other replication classes. Common plasmid vectors used for engineering are generally based on either RNA inhibition, such as ColE1, or protein-based priming and dimerization, such as sc101 (refs. 32–35). We constructed two additional Dup_Ori_Lys strains using sc101 origins of replication: a low-copy variant plus the M78I high-copy variant used to generate Diff_Ori strains in the previous experiments. These alternative Dup_Ori_Lys strains also shifted their plasmid ratio as expected, minimizing lysis-induced plasmids across four inducer environments (Fig. 5e). The original ColE1 Dup_Ori_Lys strain had the largest fluorescence shift magnitude, followed by the high-copy sc101 variant, validating simulation results showing that a higher total CN leads to larger selection magnitudes (Extended Data Fig. 6a).

Finally, we visualized the three Dup_Ori_Lys strains under changing environments using a multi-strain microfluidic chip to characterize selection and memory dynamics in parallel[24]. Using 12 h induction windows, cells were grown under five media conditions: glucose, 100 nM AHL, glucose, 0.02% arabinose and then glucose again (Fig. 5f and Supplementary Video 3). Antibiotics were used throughout the experiment to ensure that one copy of each plasmid was always maintained. The Dup_Ori_Lys strains responded to both selection windows, showing reversible adaptation. Following non-selective growth for 12 h, even in continuous culture, ColE1 and high-copy Sc101 populations retained their CN memory of the previous selection event, as measured by GFP/RFP levels. Low-copy sc101 populations followed similar response curves, but with diminished response (Extended Data Fig. 6b). GFP/RFP shifts showed consistent response to AHL induction and highly variable responses to arabinose induction across traps, potentially due to the noisiness of the arabinose promoter (Fig. 5g)[36]. We extracted the characteristic time scales for selection and memory for each strain and condition by fitting each trap's fluorescence ratio time trace to the predicted equation for CN ratio, as described in the Methods (Fig. 5h). Memory time scales were slower than those for selection, demonstrating generational memory for CN state. Lastly, simulations provided potential design guidance, suggesting that memory improves with larger CNs (Fig. 5i).

In summary, duplicate-origin plasmids across replication classes enable fast adaptation to engineered lysis burden in continuous culture experiments. These strains are tunable by lysis expression strength and display population-state memory over many generations.

Discussion

Synthetic biologists strive for tight control of cellular phenotypes, often seeing variability as a negative feature. However, natural populations often exploit heterogeneity to survive unpredictable environments. In this work, we take inspiration from nature to create heterogeneous populations capable of environmental adaptation, burden minimization and population memory by coupling two plasmids of the same origin type.

To guide our conceptual understanding, we developed a simple probabilistic model for PCN distribution. Since the 1970s, many models have been published describing the dynamics of plasmid loss between incompatible plasmids[37–40]; however, few consider plasmid fitness. One recent study on the persistence of conjugating plasmids incorporated experimentally derived fitness parameters into the model framework[41]. In our model, the probability of cell division and death are functions of PCN and its environmentally encoded cost. Future work to improve the model could include alternative strategies for modelling host–circuit coupling[42] or direct estimation of parameters from experimental data. Despite these limitations, the model recapitulates key properties of multicopy, incompatible plasmids, such as heterogeneity and burden buffering, that we find experimentally.

Previous work has shown that the expression level of metabolic proteins evolves towards an optimum when cultured under a specific nutrient environment[43]. Here we show optimization of protein expression to the environment by CN adjustment rather than mutation. We introduce the concept of plasmid buffering, where an essential plasmid is given a tunable CN by a secondary non-essential plasmid with the same origin. Work by the Ellis laboratory and others on burden-responsive feedback control highlights the importance of considering the effects of circuit burden on cellular fitness[25,44]. Our study shows how plasmids with shared replication mechanisms allow cells to minimize the burden caused by protein expression that is mismatched with the culture environment. In addition, we show that plasmid buffering can also increase the stability of a non-essential plasmid by giving it an indirect benefit to the cell. Current research on improving genetic stability includes decreasing host mutational ability, overlapping sequences of essential and non-essential genes and recoding translation, among many other strategies[45]. Here we are able to enhance genetic stability using duplicate-origin systems, against conventional wisdom.

Recent research into understanding gene evolution has also highlighted the potential benefits of CN heterogeneity. Tomanek et al.[46] investigated a natural case of bacterial gene CN adaptation termed amplification-mediated gene expression tuning. Although the mechanism for generating CN variation differs, their study shares the core idea of stochastic gene amplification promoting population adaptation. Interestingly, another case of CN adaptation has been found in the extrachromosomal DNA of drug-resistant cancer cells[47]. These instances of naturally occurring CN variation lead us to ask whether plasmid variants with shared origins are also used in nature to cope with environmental change. Although not yet shown in natural populations, the potential long-term advantages of multicopy plasmids with different alleles in surviving fluctuating antibiotic conditions has been demonstrated[48].

In this work, we intentionally created synthetic constructs with shared plasmid origins to build environmentally adaptive populations. The use of plasmid coupling in synthetic biology, although not prevalent, was a key component of earlier work by Tang and Liu[14] on environmental recording and inspired a new study on adaptive decision-making in microbial populations[49] in which the authors use controlled, unidirectional adaptation of the CN ratio in response to chemical patterns to direct learning in a sensor library. As in our work, they harnessed CN flexibility to generate adaptive memory of environmental exposure, but with the motivation of using iterative learning to master games such as tic–tac–toe. Here we demonstrate bidirectional population memory following environmental selection for multiple plasmid types, particularly high-copy origins. Researchers have created population-state memory in synthetic systems through the use of bistable circuit motifs[50,51], and more recently the development of a methylation-based epigenetic system within *E. coli*[52]. In this work, plasmid inheritance drives generational memory of population CN distribution. Studies on the history dependence of microbial populations have cited chromatin state, protein inheritance, metabolic history and strong heterogeneity as potential mechanisms for memory of past environmental exposures[53]. We take inspiration from these natural mechanisms of population adaptation to demonstrate the broad use of duplicate-origin plasmids as a tool with which to improve the performance of synthetic microbial populations.

Methods

Strains and plasmids

E. coli MG1655 (NCBI U00096.3) was used as the host strain for the majority of experiments, except where noted. For experiments involving the lactose and arabinose operon strains, *E. coli* strain JS006 (ref. 23) was used. This strain is derived from the *E. coli* Keio knockout collection parent strain BW25113 and has the following relevant mutations: Δ(araD–araB)567, ΔlacZ4787(::rrnB-3), ΔlacI and ΔaraC. For experiments involving inducible pUC CN plasmids, the *E. coli* K-12 MG1655 Z1 strain was used, which constitutively expresses *lacI* and *tetR* from the genome.

All plasmids were constructed by Gibson assembly using PCR-amplified fragments of previously constructed plasmids from our group. The full *E. coli* lactose operon was amplified from the MG1655 strain genome. The full *E. coli* arabinose operon was also amplified from the MG1665 strain genome. The inducible lysis library was created using a mutagenic primer targeting the ribosome binding site in front of its lysis gene. Strain and plasmid contents are further annotated in Supplementary Tables 1 and 2.

Before any experiment, all cultures were seeded from a glycerol stock containing cells that were grown once in Luria–Bertani (LB) broth following transformation. Strains were grown in the presence of antibiotics for both plasmids in all cases except Dup_Ori_Met and Diff_Ori_Met strains grown in lactose or arabinose media. Additionally, plasmid loss experiments required the passaging of cultures with no antibiotic.

Plate reader experiments for estimating average CN in batch culture using fluorescence

For plate reader experiments assessing strain growth and bulk fluorescence, saturated overnight cultures of a given strain grown from glycerol stocks were diluted 1:100 into 200 µl fresh media and cultured in 96-well flat-bottom plates at 37 °C with orbital shaking in a Tecan Infinite M200 PRO. Optical density and fluorescence measurements were taken every 5 min and the cultures were grown to saturation. GFP and RFP measurements were normalized by dividing by the culture OD at each time point. For use in comparative plots between conditions, final GFP/OD and RFP/OD values were determined from the point at which the OD was maximized or stopped increasing—a signal of the early stationary phase.

Agarose pad slide preparation for single-cell imaging and analysis

To image live *E. coli* cells following growth under relevant media conditions, overnight cultures were diluted tenfold in sterile water before 2 µl was pipetted onto a glass coverslip. Cells were then sandwiched by placing an agar pad (1.5% agarose) on top to create a flat focal plane. Agar pads were made the depth of one coverslip using the method described by Skinner et al.[54]. To analyse single-cell fluorescence from images, a custom MATLAB script was written that segmented single cells before recording their average fluorescence intensity. The segmentation workflow included image filtering, edge detection followed by object filling and lastly individual cell classification and assignment using size and fluorescence thresholds. This script, along with other relevant code, is available at https://doi.org/10.6084/m9.figshare.24751908 (ref. 55).

Quantitative PCR for the determination of average CN

For direct quantification of PCN, saturated cell cultures were analysed by qPCR. Following growth under relevant media conditions, cells were incubated at 95 °C for 10 min before freezing at −20 °C. Lysed cultures were then diluted 100-fold for use as final templates. Each qPCR reaction consisted of 10 µl SYBR Green qPCR Master Mix (MedChemExpress), 4 µl water, 5 µl diluted template and 0.5 µl of each primer. Primers were designed to amplify approximately 100 base pair regions within the GFP gene, RFP gene, plasmid origin and *E. coli* chromosome (Supplementary Table 3). Reactions were run in triplicate using MicroAmp Fast Optical 96-well plates within the Applied Biosystems QuantStudio 3 machine for comparative cycle threshold (Ct) measurements. The resulting measurements were normalized to chromosomal control reactions to obtain the relative read abundance ($F = 2^{-[\mathrm{Ct(Target)} - \mathrm{Ct(Chromosome)}]}$). Of note, due to the use of relative measurements for each primer pair, abundance comparisons must be made only within the same target group. Sources of sample variability include differences in cell number and distribution between subpopulations within each reaction well. In addition, exact gene placement relative to the plasmid origin within different plasmid constructs can introduce small biases in read abundance.

Microfluidics and microscopy for carbon-source operon strains and lysis strains in multi-strain devices

For time-lapse microscopy experiments with carbon-source operon strains, another previously developed device was used[24]. This device allows for observation of multiple strains within individual traps that are grown with the same inlet media source without cross-contamination. Carbon-source experiments consisted of 24 h induction periods with various supplemented M9 media and no antibiotics. Lysis memory experiments consisted of 12 h induction periods between selective inducer media (100 nM AHL or 0.02% arabinose) and non-selective media (0.2% glucose), with the constant presence of stabilizing antibiotics. Images were taken every 10 min with brightfield, red fluorescence and green fluorescence channels.

Mother machine analysis of the carbon-source operon strain in arabinose

To investigate the fitness-mediated adaptation of CN to arabinose media at the single-cell level, we loaded our Dup_Ori_Met strain into a previously described mother machine device for visualization of separated cell lineages[26]. Due to the slow time scales of carbon-source adaptation and the poor overall growth of this strain in arabinose media, the culture was first pre-conditioned in M9 arabinose supplemented with 0.5 mg ml^{-1} bovine serum albumin and 0.075% Tween before visualization under the same conditions for roughly 12 h. As our current imaging setup is incompatible with automated cell segmentation and tracking by machine-learning algorithms, we used macro-assisted hand segmentation of cells to extract fluorescence, size and location data over time. The final results showed the end-point lineage size from each initially loaded cell and its corresponding fluorescence.

Serial passaging and plating experiments to estimate the plasmid loss rate for carbon-source operon strains

Carbon-metabolizing operon strains were grown overnight in 3 ml LB media with chloramphenicol and spectinomycin directly from glycerol stocks. The next day, the overnight culture was washed twice in M9 minimal media and used to inoculate new cultures at a 1:100 dilution in M9 media that was supplemented with Wolfe's Vitamin Solution, MOPS (Teknova, M2101) for trace minerals, and either 0.2% glucose, arabinose or lactose (with no antibiotics). The cultures were grown for ~24 h and passaged 1:100 into fresh, supplemented M9 minimal media that contained the same carbon source that the strain had been grown in on the previous day. After five passages, samples of the different cultures were plated on LB agar plates containing either no antibiotics, chloramphenicol or spectinomycin to determine colony forming unit (c.f.u.) numbers. Specifically, 3 µl droplets of tenfold serial dilutions of the culture were spotted on the agar plates and left to dry before culturing the agar plates at 37 °C overnight. Three biological replicates for each condition and strain were plated in technical triplicates. The following day, c.f.u. numbers in the culture were determined by counting the number of colonies that formed for the serial dilution that led to between 10 and 30 colonies on average. The c.f.u. numbers in the undiluted cultures were then determined by multiplying the number of counted colonies by the dilution factor. Finally, c.f.u. numbers in antibiotics were divided by the total c.f.u. numbers found on just LB to determine plasmid retention. Plasmid loss experiments including the addition of 100 µM IPTG were plated after three passages for the appropriate retention range.

Serial passaging and memory quantification of inducible kanamycin strains in plate readers

For memory experiments with inducible kanamycin strains, overnight cultures were diluted 1:100 in selective media containing AHL (0–100 nM) and kanamycin (1–15×) within a 96-well plate. After 12 h of growth, cells were diluted into media containing 1× kanamycin and 0 nM AHL in a fresh 96-well plate. Three non-selective passages of 12 h each

were carried out before incubation of cells under flipped selective media conditions. This selective phase exposed each culture to the opposite condition to the initial selective phase; for example, cultures first incubated at 100 nM AHL and 1× kanamycin were now exposed to 0 nM AHL and 15× kanamycin. Following this secondary selective incubation, three more non-selective passages of 12 h each were performed. GFP/OD measurements were then plotted over time and each non-selective memory phase was fit to an exponential curve to determine the memory time constant (return rate) following the GFP response.

Microfluidics and microscopy for inducible lysis strains in a concentration gradient device

For time-lapse microscopy experiments with the inducible cell lysis strains within discrete inducer gradients, a previously developed microfluidic device was used[31]. This device consisted of separate arrays of small cell traps maintained at mixed inducer concentrations depending on the inducer concentrations in two upstream inlet channels of the device. Lysis experiments were always run with antibiotics.

Image analysis following inducible lysis strain time-lapse microscopy experiments

To quantify population fluorescence from microfluidics experiments, we developed a FIJI macro. This macro created regions of interest around each individual trap, with thresholded fluorescence in each channel to select only healthy cells. It then merged these regions of interest and measured the mean intensity for each time slice. One value each for the threshold minimum and maximum was set for an entire experiment to avoid non-cell background noise and also exclude unhealthy filamentous cells. Post-processing of the results also included background value subtraction for each trap and time slice.

Fitting of selection and memory time constants to dynamic lysis experiments

In our simulation results, we are able to directly use CN over time to estimate selection and memory time constants through exponential curve fits. Our microscopy data, however, use GFP/RFP fluorescence as a robust metric for CN change, regardless of cell state. To extract time constants from fluorescence ratio data, the results from each incubation window were fit to the below equation, where R_i is the initial GFP/RFP value and A and B are unknown constants. This equation is derived from dividing two exponential functions representing p_1 and p_2 CN change.

$$\frac{([R_i - \mathrm{A}]e^{-\tau t} + \mathrm{A})}{([1 - \mathrm{B}]e^{-\tau t} + \mathrm{B})} \tag{1}$$

PCN model

$$P_{\mathrm{rep}} = r\left(1 - \frac{p(t)}{p_{\mathrm{tot}}}\right) \tag{2}$$

$$P_{\mathrm{death}} = \frac{N^2}{N^2 + S^2} + d_{\max}\left(\frac{p_1^n}{p_1^n + K_{d1}^n}\right) + d_{\max}\left(\frac{p_2^n}{p_2^n + K_{d2}^n}\right) \tag{3}$$

$$P_{\mathrm{division}} = g_{\max}\left(1 - \frac{N}{C}\right) \times \left(\frac{p_1^n}{K_{\mathrm{b1}}^n + p_1^n}\right)\left(\frac{K_{\mathrm{c1}}^n}{K_{\mathrm{c1}}^n + p_1^n}\right) \times \left(\frac{p_2^n}{K_{\mathrm{b2}}^n + p_2^n}\right)\left(\frac{K_{\mathrm{c2}}^n}{K_{\mathrm{c2}}^n + p_2^n}\right) \tag{4}$$

To study how the PCN distribution of a population changes over time, we created a probabilistic model that takes into account the cost versus benefit associated with a given plasmid. The model consists of cells containing two distinct plasmids that can have either the same or orthogonal mechanisms of replication and CN maintenance. In each generation, the following steps are carried out: (1) plasmids are replicated; (2) some members of the population die (due to encoded plasmid or effects of population saturation); (3) cells have the chance to divide based on plasmid-associated cost–benefit functions; and (4) for dividing cells, plasmids are binomially partitioned into two daughter cells. We assume that the plasmid replication probability is given by equation (2), where r is the maximum replication probability, $p(t)$ is the current PCN and p_{tot} is the maximum PCN. For incompatible plasmids, $p(t)$ represents the total CN of both plasmid types and p_{tot} is their shared maximum CN, whereas for compatible plasmids, each plasmid has its own P_{rep} equation. The number of plasmids that replicate in a given cycle is a random number from the binomial distribution with probability P_{rep} and $p(t)$ trials.

Equation (3) describes the probability of a given cell dying based on its current PCNs. P_{death} represents the death probability due to general saturation effects as the population size approaches the carrying capacity C. The parameter S represents the population size at which the death probability due to culture saturation is half maximal. The two subsequent terms represent death probability due to lysis gene expression from plasmids p_1 and p_2, respectively. Death probability due to plasmid gene expression is assumed to follow the Hill equation where K_{d} represents the PCN that leads to half-maximal death probability and $d_{\max}$ represents the maximum death probability.

Division probability is calculated in a similar manner to death probability, as shown in equation (4). In this equation, we assume logistic growth of the bacterial population where $g_{\max}$ is the maximum division probability and C is the carrying capacity. Each plasmid (p_1 and p_2) is assumed to have a potential benefit and cost that can increase or decrease the division probability. The Hill equation is used to model plasmid cost–benefit effects where for each plasmid K_{b} represents the PCN that leads to half-maximal benefit and K_{c} represents the PCN leading to half-maximal cost. A cell divides if a number of one is randomly chosen from a binomial distribution with probability P_{division} and one trial. If a cell is chosen to divide, its plasmids are partitioned into two daughter cells binomially.

Reporting summary

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

Data availability

Source data are provided with this paper. Source data are also available via Figshare at https://doi.org/10.6084/m9.figshare.24751896 (ref. 56). Any additional clarifications or information required to reanalyse the data reported in this paper are available from the corresponding author upon request. Annotated plasmid sequences used in this study are available in Supplementary Data 1 and via Figshare at https://doi.org/10.6084/m9.figshare.24751911 (ref. 57).

Code availability

The original modelling code to repeat the simulations in this study is available in consolidated Python notebooks in Supplementary Code 1 and via Figshare at https://doi.org/10.6084/m9.figshare.24751908 (ref. 55).

References

1. Del Vecchio, D., Qian, Y., Murray, R. M. & Sontag, E. D. Future systems and control research in synthetic biology. *Annu. Rev. Control* **45**, 5–17 (2018).
2. Reyes Ruiz, L. M., Williams, C. L. & Tamayo, R. Enhancing bacterial survival through phenotypic heterogeneity. *PLoS Pathog.* **16**, e1008439 (2020).
3. Grote, J., Krysciak, D. & Streit, W. R. Phenotypic heterogeneity, a phenomenon that may explain why quorum sensing does not always result in truly homogenous cell behavior. *Appl. Environ. Microbiol.* **81**, 5280–5289 (2015).
4. Armbruster, C. R. et al. Heterogeneity in surface sensing suggests a division of labor in *Pseudomonas aeruginosa* populations. *eLife* **8**, e45084 (2019).

5. Grimbergen, A. J., Siebring, J., Solopova, A. & Kuipers, O. P. Microbial bet-hedging: the power of being different. *Curr. Opin. Microbiol.* **25**, 67–72 (2015).
6. Aronson, M. S., Ricci-Tam, C., Zhu, X. & Sgro, A. E. Exploiting noise to engineer adaptability in synthetic multicellular systems. *Curr. Opin. Biomed. Eng.* **16**, 52–60 (2020).
7. Dar, R. D. & Weiss, R. Perspective: engineering noise in biological systems towards predictive stochastic design. *APL Bioeng.* **2**, 020901 (2018).
8. Zhu, L., Zhu, Y., Zhang, Y. & Li, Y. Engineering the robustness of industrial microbes through synthetic biology. *Trends Microbiol.* **20**, 94–101 (2012).
9. Kittleson, J. T., Cheung, S. & Anderson, J. Rapid optimization of gene dosage in *E. coli* using dial strains. *J. Biol. Eng.* **5**, 10 (2011).
10. Rouches, M. V., Xu, Y., Cortes, L. B. G. & Lambert, G. A plasmid system with tunable copy number. *Nat. Commun.* **13**, 3908 (2022).
11. Joshi, S. H.-N., Yong, C. & Gyorgy, A. Inducible plasmid copy number control for synthetic biology in commonly used *E. coli* strains. *Nat. Commun.* **13**, 6691 (2022).
12. Li, C., Zou, Y., Jiang, T., Zhang, J. & Yan, Y. Harnessing plasmid replication mechanism to enable dynamic control of gene copy in bacteria. *Metab. Eng.* **70**, 67–78 (2022).
13. Baumgart, L., Mather, W. & Hasty, J. Synchronized DNA cycling across a bacterial population. *Nat. Genet.* **49**, 1282–1285 (2017).
14. Tang, W. & Liu, D. R. Rewritable multi-event analog recording in bacterial and mammalian cells. *Science* **360**, eaap8992 (2018).
15. Del Solar, G. & Espinosa, M. Plasmid copy number control: an ever-growing story. *Mol. Microbiol.* **37**, 492–500 (2000).
16. Camps, M. Modulation of ColE1-like plasmid replication for recombinant gene expression. *Recent Pat. DNA Gene Seq.* **4**, 58–73 (2010).
17. Velappan, N., Sblattero, D., Chasteen, L., Pavlik, P. & Bradbury, A. R. Plasmid incompatibility: more compatible than previously thought? *Protein Eng. Des. Sel.* **20**, 309–313 (2007).
18. Lau, B. T., Malkus, P. & Paulsson, J. New quantitative methods for measuring plasmid loss rates reveal unexpected stability. *Plasmid* **70**, 353–361 (2013).
19. Thompson, M. G. et al. Isolation and characterization of novel mutations in the pSC101 origin that increase copy number. *Sci. Rep.* **8**, 1590 (2018).
20. Elowitz, M. B., Levine, A. J., Siggia, E. D. & Swain, P. S. Stochastic gene expression in a single cell. *Science* **297**, 1183–1186 (2002).
21. Prakash, S., Racovita, A., Petrucci, T., Galizi, R. & Jaramillo, A. qsanger: Quantification of genetic variants in bacterial cultures by sanger sequencing. *BioDesign Res.* **5**, 0007 (2023).
22. Eames, M. & Kortemme, T. Cost–benefit tradeoffs in engineered *lac* operons. *Science* **336**, 911–915 (2012).
23. Stricker, J. et al. A fast, robust and tunable synthetic gene oscillator. *Nature* **456**, 516–519 (2008).
24. Lezia, A., Csicsery, N. & Hasty, J. Design, mutate, screen: multiplexed creation and arrayed screening of synchronized genetic clocks. *Cell Syst.* **13**, 365–375 (2022).
25. Ceroni, F. et al. Burden-driven feedback control of gene expression. *Nat. Methods* **15**, 387–393 (2018).
26. Thiermann, R. et al. Tools and methods for high-throughput single-cell imaging with the mother machine. *eLife* **12**, RP88463 (2024).
27. Penumetcha, P. et al. Improving the *lac* system for synthetic biology. *Bios* **81**, 7–15 (2010).
28. Wang, X., Xia, K., Yang, X. & Tang, C. Growth strategy of microbes on mixed carbon sources. *Nat. Commun.* **10**, 1279 (2019).
29. Shao, B. et al. Single-cell measurement of plasmid copy number and promoter activity. *Nat. Commun.* **12**, 1475 (2021).
30. Jahn, M., Vorpahl, C., Hübschmann, T., Harms, H. & Müller, S. Copy number variability of expression plasmids determined by cell sorting and droplet digital PCR. *Microb. Cell Fact.* **15**, 211 (2016).
31. Miano, A., Liao, M. J. & Hasty, J. Inducible cell-to-cell signaling for tunable dynamics in microbial communities. *Nat. Commun.* **11**, 1193 (2020).
32. Cesareni, G., Helmer-Citterich, M. & Castagnoli, L. Control of ColE1 plasmid replication by antisense RNA. *Trends Genet.* **7**, 230–235 (1991).
33. Selzer, G., Som, T., Itoh, T. & Tomizawa, J. The origin of replication of plasmid p15A and comparative studies on the nucleotide sequences around the origin of related plasmids. *Cell* **32**, 119–129 (1983).
34. Chattoraj, D. K. Control of plasmid DNA replication by iterons: no longer paradoxical. *Mol. Microbiol.* **37**, 467–476 (2000).
35. Rakowski, S. A. & Filutowicz, M. Plasmid R6K replication control. *Plasmid* **69**, 231–242 (2013).
36. Megerle, J. A., Fritz, G., Gerland, U., Jung, K. & Rädler, J. O. Timing and dynamics of single cell gene expression in the arabinose utilization system. *Biophys. J.* **95**, 2103–2115 (2008).
37. Novick, R. P. Plasmid incompatibility. *Microbiol. Rev.* **51**, 381–395 (1987).
38. Novick, R. P. et al. Uniform nomenclature for bacterial plasmids: a proposal. *Bacteriol. Rev.* **40**, 168–189 (1976).
39. Hernández-Beltrán, J. C. R., San Millán, A., Fuentes-Hernández, A. & Peña-Miller, R. Mathematical models of plasmid population dynamics. *Front. Microbiol.* **12**, 606396 (2021).
40. Ishii, K., Hashimoto-Gotoh, T. & Matsubara, K. Random replication and random assortment model for plasmid incompatibility in bacteria. *Plasmid* **1**, 435–445 (1978).
41. Alonso-del Valle, A. et al. Variability of plasmid fitness effects contributes to plasmid persistence in bacterial communities. *Nat. Commun.* **12**, 2653 (2021).
42. Nikolados, E.-M., Weiße, A. Y. & Oyarzún, D. A. Prediction of cellular burden with host–circuit models. *Methods Mol. Biol.* **2229**, 267–291 (2021).
43. Dekel, E. & Alon, U. Optimality and evolutionary tuning of the expression level of a protein. *Nature* **436**, 588–592 (2005).
44. Ceroni, F., Algar, R., Stan, G.-B. & Ellis, T. Quantifying cellular capacity identifies gene expression designs with reduced burden. *Nat. Methods* **12**, 415–418 (2015).
45. Kumar, S. & Hasty, J. Stability, robustness, and containment: preparing synthetic biology for real-world deployment. *Curr. Opinion Biotechnol.* **79**, 102880 (2023).
46. Tomanek, I. et al. Gene amplification as a form of population-level gene expression regulation. *Nat. Ecol. Evol.* **4**, 612–625 (2020).
47. Lange, J. T. et al. The evolutionary dynamics of extrachromosomal DNA in human cancers. *Nat. Genet.* **54**, 1527–1533 (2022).
48. Rodriguez-Beltran, J. et al. Multicopy plasmids allow bacteria to escape from fitness trade-offs during evolutionary innovation. *Nat. Ecol. Evol.* **2**, 873–881 (2018).
49. Racovita, A. et al. Engineered gene circuits with reinforcement learning allow bacteria to master gameplaying. Preprint at *bioRxiv* https://doi.org/10.1101/2022.04.22.489191 (2022).
50. Gardner, T. S., Cantor, C. R. & Collins, J. J. Construction of a genetic toggle switch in *Escherichia coli*. *Nature* **403**, 339–342 (2000).
51. Urrios, A. et al. A synthetic multicellular memory device. *ACS Synth. Biol.* **5**, 862–873 (2016).
52. Maier, J. A., Möhrle, R. & Jeltsch, A. Design of synthetic epigenetic circuits featuring memory effects and reversible switching based on DNA methylation. *Nat. Commun.* **8**, 15336 (2017).
53. Vermeersch, L. et al. Do microbes have a memory? History-dependent behavior in the adaptation to variable environments. *Front. Microbiol.* **13**, 4052 (2022).

54. Skinner, S. O., Sepúlveda, L. A., Xu, H. & Golding, I. Measuring mRNA copy number in individual *Escherichia coli* cells using single-molecule fluorescent in situ hybridization. *Nat. Protoc.* **8**, 1100–1113 (2013).
55. Kumar, S. Code example files. *Figshare* https://doi.org/10.6084/m9.figshare.24751908.v1 (2024).
56. Kumar, S. Source data. *Figshare* https://doi.org/10.6084/m9.figshare.24751896.v1 (2024).
57. Kumar, S. Annotated plasmid sequences. *Figshare* https://doi.org/10.6084/m9.figshare.24751911.v1 (2024).

Acknowledgements

This work was supported by the National Institute of Biomedical Imaging and Bioengineering of the National Institutes of Health (grant number R01EB030134). A.L. and S.K. were supported in part by the National Science Foundation Graduate Research Fellowship Program under grant number DGE-2038238. S.K. was also supported by the National Institutes of Health-sponsored Quantitative Integrative Biology Training Grant (number 5T32GM127235). We thank N. Csicsery and R. O'Laughlin for critically reading the manuscript and providing valuable feedback. We also thank R. Thiermann and the Jun Lab for the donation of mother machine chips and sharing of protocols.

Author contributions

S.K., A.L. and J.H. conceptualized the project and wrote the paper. S.K. and A.L. designed, performed and analysed all of the experiments.

Competing interests

J.H. declares that he is a co-founder of GenCirq, which focuses on cancer therapeutics. He is on the Board of Directors and has equity in GenCirq. His spouse is employed part time by GenCirq for bookkeeping and employee support with human resources.

Additional information

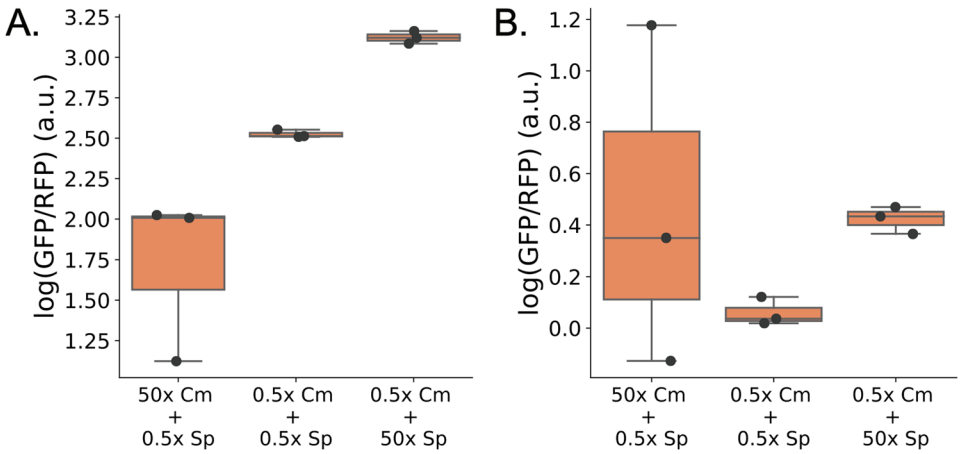
Extended data is available for this paper at https://doi.org/10.1038/s41564-024-01706-w.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41564-024-01706-w.

Correspondence and requests for materials should be addressed to Shalni Kumar.

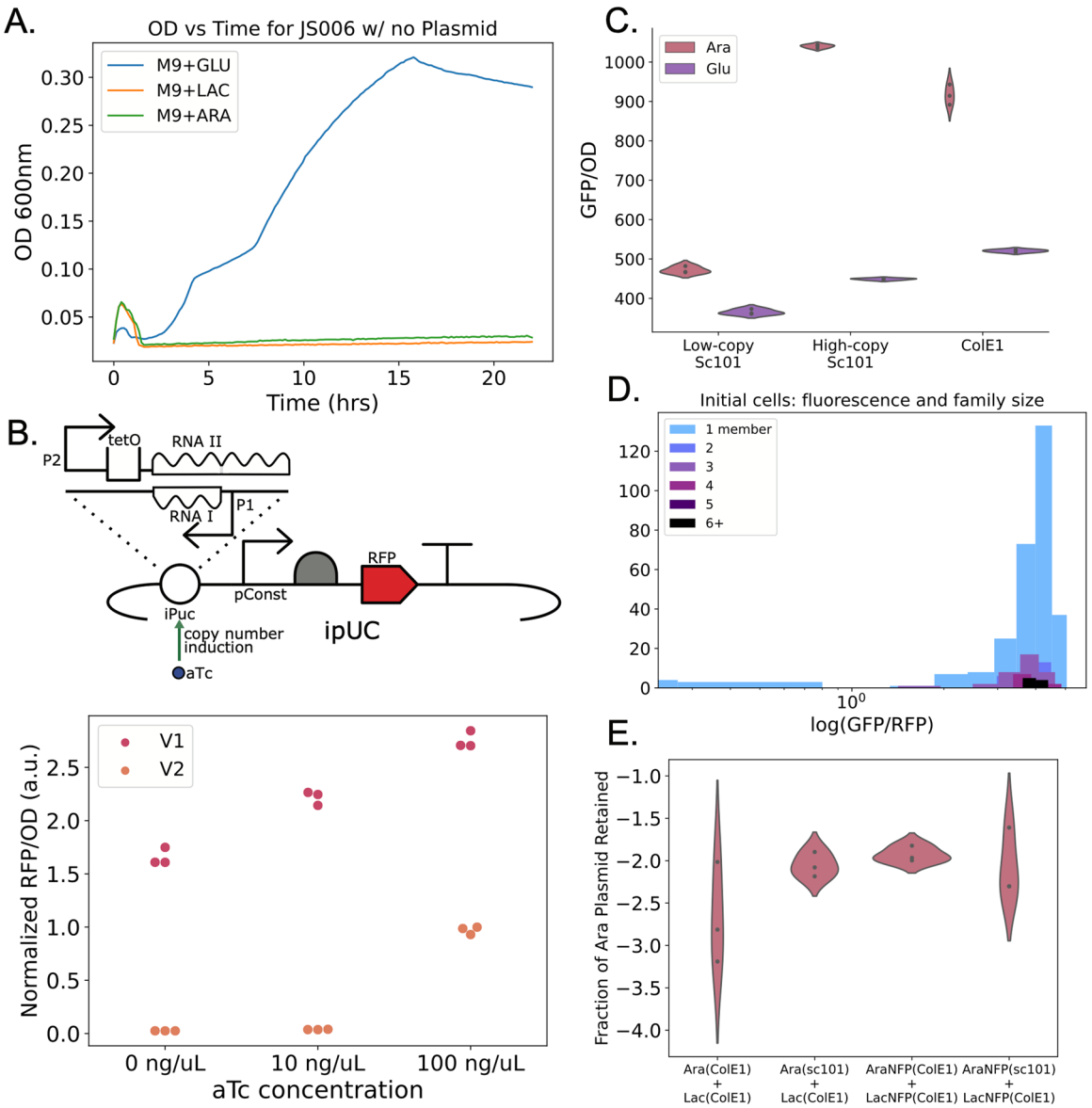
Peer review information *Nature Microbiology* thanks Alfonso Jaramillo and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

https://doi.org/10.1038/s41564-024-01706-w



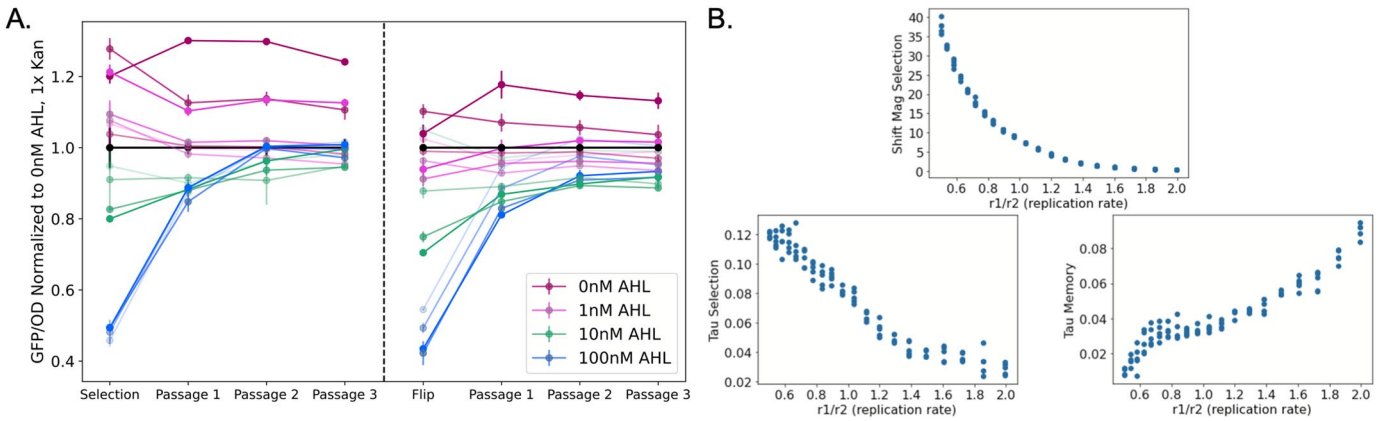
Extended Data Fig. 1 | Fluorescence shift of Dup_ori and Diff_ori strains under strong antibiotic. Relative GFP/RFP plate reader fluorescence for (**A**) Dup_ori and (**B**) Diff_ori strains following growth in various concentrations of antibiotic (N=3 biological replicates). Boxplots show the minimum, 25th percentile, median, 75th percentile, and maximum.

https://doi.org/10.1038/s41564-024-01706-w



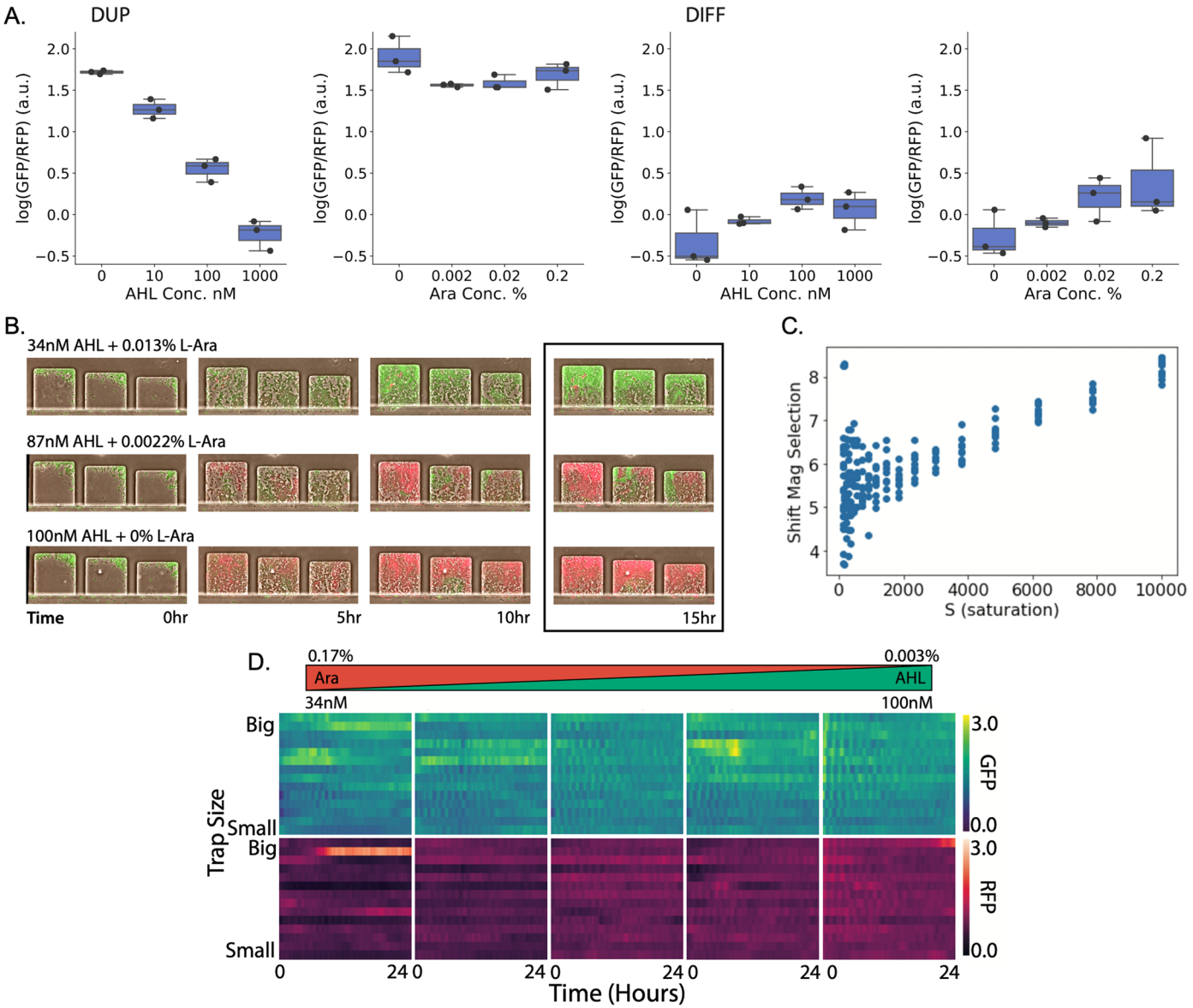
Extended Data Fig. 2 | High copy number of the arabinose operon induces cellular burden. (**A**) Strain JS006 is unable to grow on minimal media with arabinose or lactose as the only carbon source unless operon genes are supplied on plasmids. (**B**) Proof of function for inducible copy number plasmid, Top: Circuit diagram for testing aTc-inducible pUC copy number. Bottom: RFP fluorescence vs. aTc concentration for two different versions of the inducible pUC system with different tetO operator sequences. (N= 3 biological replicates) (**C**) GFP output of htpG1 burden reporter construct when co-transformed with AraOP plasmid variants in glucose or arabinose media (N=3 biological replicates). (**D**) Fluorescent distribution of single cells seeding a mother machine device and their subsequent lineage or family size following arabinose growth. Cells were preconditioned in M9 arabinose prior to loading. (N=376 cells) (**E**) Plasmid retention data for Dup_ori_met and Diff_ori_met strains grown on lactose minimal media. Comparison data for main Fig. 3f. NFP strains have fluorescent reporter removed (N=3 biological replicates).

https://doi.org/10.1038/s41564-024-01706-w



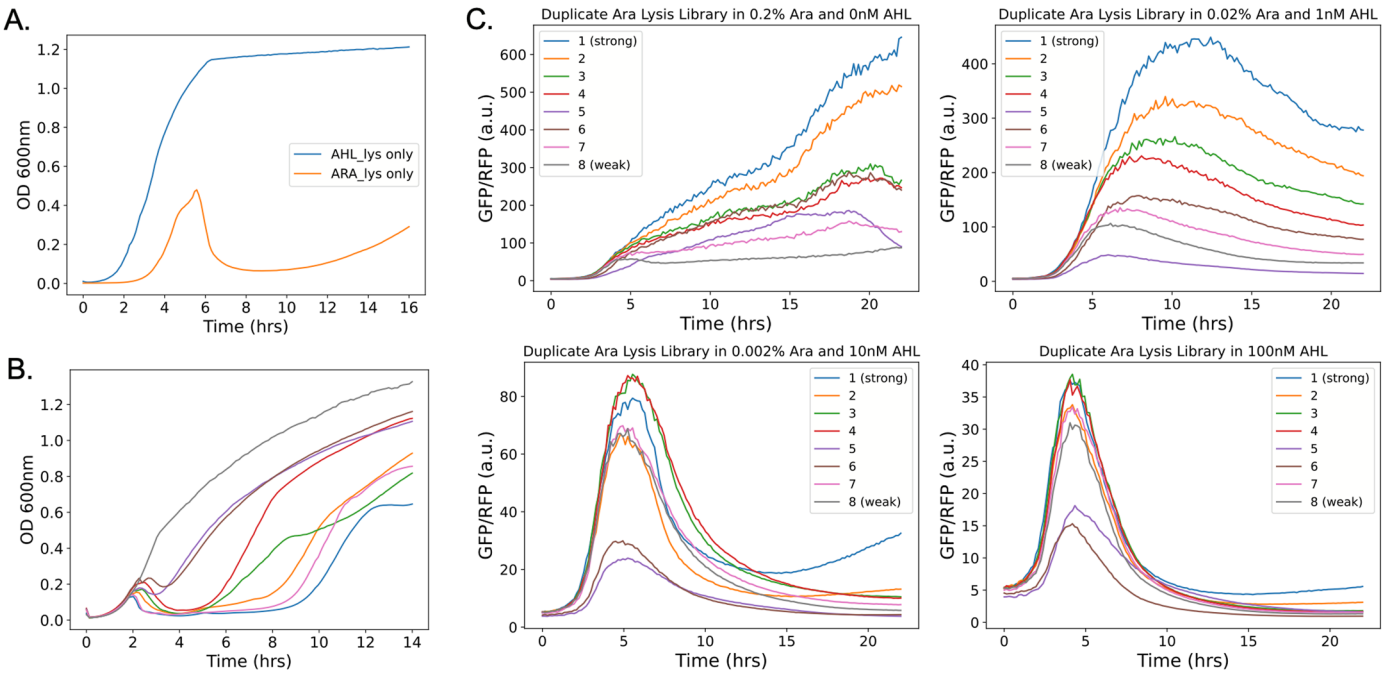
Extended Data Fig. 3 | Diff_ori_kan strain shows little memory and simulated plasmid size differences cause bias. (**A**) Plate reader fluorescence measurements for Diff_ori_kan in all inducer conditions over time during non-selective passaging with subsequent flipped induction. Comparison data for Fig. 4e. Error bars show mean +/- SD for three separate cultures. (**B**) Parameter sweep showing effect of plasmid replication rate bias on CN shift magnitude, response speed, and memory time.

https://doi.org/10.1038/s41564-024-01706-w



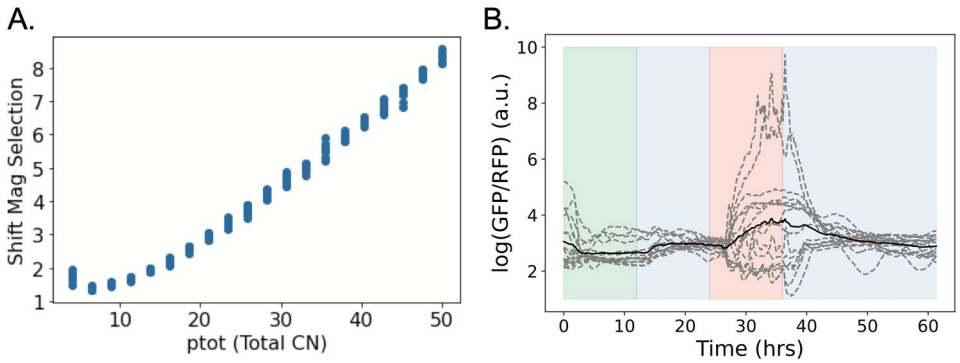
Extended Data Fig. 4 | Strong fluorescence shift of Dup_ori_lys strain compared to Diff_ori_lys and improvement with population size. (**A**) Relative GFP/RFP fluorescence for the Dup_ori_lys and Diff_ori_lys strains grown in batch culture in different concentrations of AHL and arabinose (N=3 biological replicates). Boxplots show the minimum, 25th percentile, median, 75th percentile, and maximum. (**B**) Representative microscope images from gradient microfluidic experiment with Dup_ori_lys strain. Images correspond to data shown in Fig. 5b. (**C**) Parameter sweep showing effect of population size on magnitude of CN shift under selection. (**D**) Heatmap showing GFP and RFP fluorescence levels over time for the Diff_ori_lys strain grown in various concentration of AHL and arabinose in a microfluidic device (N=14 traps). Comparison data for Fig. 5b.

https://doi.org/10.1038/s41564-024-01706-w



Extended Data Fig. 5 | Amelioration of Ara_lys strength and resulting fluorescence response gradient. (**A**) Growth curves for strains containing either the AHL_Lys plasmid or the Ara_Lys plasmid only without inducer. (**B**) Growth curves for 8 different Ara_lys strains with various lysis RBS strengths grown in LB with 0.02% arabinose. Average of three separate cultures is shown. Data related to Fig. 5c (**C**) Plate reader fluorescence data over time for the 8 different Dup_ori_lys library strains grown in different AHL and arabinose concentrations. Average of three separate cultures is shown. Data related to Fig. 5d.

https://doi.org/10.1038/s41564-024-01706-w



Extended Data Fig. 6 | Low copy number strains have reduced selection capacity. (**A**) Parameter sweep showing effect of total CN on magnitude of CN shift under selection. Kd lysis strength is not rescaled with copy number for these simulations (**B**) Log ratio fluorescence over time for the Dup_ori_lys strain with a low-copy sc101 origin cultured in parallel within a microfluidic device with four induction windows: 0.02% Arabinose, 0.2% Glucose, 100nM AHL, and Glucose again. (N=16 traps from 4 positions).

nature portfolio

Corresponding author(s): Shalni Kumar

Last updated by author(s): Apr 12, 2024

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	Original modeling code to repeat simulations in this study is available in consolidated python notebooks in the supplement and at https://doi.org/10.6084/m9.figshare.24751908.
Data analysis	All data analysis was completed using standard analysis techniques with simple python packages used for visualization. FIJI 2.1.0/1.54f was used for image analysis of population cell fluorescence, and a simple Matlab 2021b script was written for single cell image analysis. This script is also available at https://doi.org/10.6084/m9.figshare.24751908.

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

Source data for all displayed figures, including qpcr, plate reader, CFU, and processed microscopy results are provided along with this paper and through figshare at

the following doi: https://doi.org/10.6084/m9.figshare.24751896. Any additional clarifications or information required to reanalyze the data reported in this paper is also available upon request. Annotated plasmid sequences used in this study are available in the supplemental files and at https://doi.org/10.6084/m9.figshare.24751911.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

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

Life sciences study design

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

Sample size	No sample size calculations were performed. N=3 biological replicates were used for plate reader experiments due to low variability. qPCR and plasmid loss studies used three biological replicates with three technical replicates each to reduce technical variability. Microfluidics used larger samples sizes >10 due to high variability from small bacterial populations. Finally single cell experiments used very large sample sizes >1000 to generate distributions. Statistics performed on plasmid loss studies showed sufficient power from n=3.
Data exclusions	Data was only excluded if microfluidics experiments suffered from mechanical issues affecting overall cell health or experimental validity unrelated to experimental condition. qPCR data was also only excluded if suffering from high technical variability.
Replication	Plate reader experiments had high reproducibility between replicates and trends were successfully repeated in independent experiments at least twice for antibiotic, lysis, and carbon-based shifts. Reproducibility was also established through the use of multiple varied types of experimental tests (i.e. microfluidics, batch growth). The same trends seen in batch culture were established once more in microfluidics for lysis and carbon experiments. Microfluidics experiments also contained independently seeded traps per experiment. Reproducibility was only difficult in microfluidic runs of carbon-responsive strains due to poor growth overall, however the same results were showed repeatedly in batch culture and our microfluidic run had many independently seeded traps.
Randomization	Randomization was not done on samples. Each strain was placed in specific conditions relevant to the experiment and strains being compared would be tested in all the same conditions. Initial seeding of growth cultures would be from an unbiased stab from the same glycerol stock each time.
Blinding	Blinding was not performed in this study as all strains undergo identical measurement and analysis in the plate reader or microscope. When possible, strains were measured in parallel. Image analysis scripts also are run across all samples without adjustment.

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
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging