PNAS RESEARCH ARTICLE | MICROBIOLOGY

Pulsatile basal gene expression as a fitness determinant in bacteria

K. Jain[a,1], R. Hauschild[a], O. O. Bochkareva[a,2], R. Roemhild[a], G. Tkačik[a], and C. C. Guet[a,1]

Edited by Andrew Murray, Harvard University, Cambridge, MA; received July 11, 2024; accepted February 19, 2025

Active regulation of gene expression, orchestrated by complex interactions of activators and repressors at promoters, controls the fate of organisms. In contrast, basal expression at uninduced promoters is considered to be a dynamically inert mode of nonfunctional "promoter leakiness," merely a byproduct of transcriptional regulation. Here, we investigate the basal expression mode of the *mar* operon, the main regulator of intrinsic multiple antibiotic resistance in *Escherichia coli*, and link its dynamic properties to the noncanonical, yet highly conserved start codon of *marR* across *Enterobacteriaceae*. Real-time, single-cell measurements across tens of generations reveal that basal expression consists of rare stochastic gene expression pulses, which maximize variability in wildtype and, surprisingly, transiently accelerate cellular elongation rates. Competition experiments show that basal expression confers fitness advantages to wildtype across several transitions between exponential and stationary growth by shortening lag times. The dynamically rich basal expression of the *mar* operon has likely been evolutionarily maintained for its role in growth homeostasis of Enterobacteria within the gut environment, thereby allowing other ancillary gene regulatory roles to evolve, e.g., control of costly-to-induce multidrug efflux pumps. Understanding the complex selection forces governing genetic systems involved in intrinsic multidrug resistance is crucial for effective public health measures.

gene regulation | basal expression | *mar* operon

Significance

Recent technological advances have enabled precise measurements of uninduced basal gene expression, also known as promoter leakiness, in individual bacteria. However, whether such basal expression mode is at all evolutionarily relevant remains unclear. We measure the basal expression of the *m*ultiple *a*ntibiotic *r*esistance operon and report its dynamic pulsatile behavior. We uncover a surprising role for this basal expression mode in general growth homeostasis of Enterobacteria that is intimately tied to their ecophysiology. We reveal how selective forces can shape gene expression modes of a global transcriptional regulator by differently trading off evolutionary costs and benefits. Understanding these trade-offs specifically for the *mar* operon as the main determinant of intrinsic multidrug resistance in Enterobacteria is essential for public health.

Basal gene expression, also known as promoter leakiness, is a characteristic of bacterial promoters that occurs in their OFF state due to the presence of a repressor or the absence of an activator. Unlike induced or constitutive expression in the ON state, basal expression is generally not thought of as a functional mode of gene expression, but rather as a lack thereof. While selection can tune various aspects of gene induction, it is unclear how it could act, if at all, on the basal expression mode. Given that promoter leakiness can be detrimental (1–3), it could be under negative selection. However, we wondered whether there are any alternative basal expression modes that could have regulatory functions in their own right and thus be positively selected for. Recent studies uncovered the existence of a much more dynamic, pulsatile basal expression mode for several bacterial genes. Such a dynamic basal mode can generate phenotypic diversity in a clonal population and has thus been rationalized as a possible bet-hedging mechanism (4–7). Essential to this explanation are two premises. The first is "frequency matching": Bet hedging conveys a long-term fitness benefit when it generates phenotypes in proportion to the frequencies of the environments for which these phenotypes are advantageous (8). The second is the existence of a growth rate "cost" for a pulse: Some (small) fraction of cells undergoing a pulse pay this cost upfront, in order to survive, or be more competitive, if a rare external stress should occur in that moment. While the bet-hedging explanation is attractive, the implied growth rate costs and benefits, as well as fitness effects more broadly, are rarely measured (4). This motivates a fundamental question: Are the two premises of bet hedging met or should one seek alternative explanations for the evolutionary maintenance of a pulsatile basal expression mode?

Here, we turn to the *marRAB* operon, initially discovered as the genetic determinant of *m*ultiple *a*ntibiotic *r*esistance, and a paradigmatic example of a highly complex bacterial regulatory circuit (9, 10). The repressor MarR and the activator MarA form a negative and a positive autoregulatory loop, respectively, and this unique topology of two interlocked loops jointly controls the *mar* function (11, 12). Salicylate induces the *marRAB* operon by derepressing MarR through a unique mechanism for inducers, by binding to the DNA binding domain of MarR and thus physically occluding the binding of repressor to the operators (13). While MarR is a local regulator of *marRAB* operon, MarA is a global regulator at the heart of one of the largest *Escherichia coli* regulons, encompassing over 30 genes, involved in multidrug efflux, pH regulation, outer membrane permeability, biofilm

Author affiliations: [a]Institute of Science and Technology Austria, Klosterneuburg 3400, Austria

Author contributions: K.J. and C.C.G. designed research; K.J., R.H., O.O.B., and R.R. performed research; K.J., R.H., O.O.B., R.R., and G.T. analyzed data; and K.J., G.T., and C.C.G. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Copyright © 2025 the Author(s). Published by PNAS. This article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

[1]To whom correspondence may be addressed. Email: kirtijn1@gmail.com or calin@ist.ac.at.

[2]Present address: Centre for Microbiology and Environmental Systems Science, Division of Computational System Biology, University of Vienna, Vienna 1030, Austria.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2413709122/-/DCSupplemental.

Published April 7, 2025.

Downloaded from https://www.pnas.org by IST AUSTRIA - LIBRARY; INSTITUTE OF SCIENCE & TECHNOLOGY on April 7, 2025 from IP address 81.223.14.217.

formation, and virulence (14–19). Modeling and experimental studies suggest that the *mar* interlocked regulatory loops could lead to pulsatile basal mode, resulting in phenotypic heterogeneity of expression that could support bet hedging (11, 20). Nevertheless, the characterization of the basal expression mode for the *mar* operon and its functional and fitness implications beyond bet hedging and antibiotic stress remains unexplored.

Conservation of GTG Start Codon in *marR* across *Enterobacteriaceae*. As MarR controls the repressed state of the *mar* operon and therefore its basal expression, the unusual presence of a weak GTG start codon in *marR* piqued our interest (21). Non-ATG start codons, i.e., GTG and TTG, initiate ~8% genes in *Gammaproteobacteria* (22), reducing the translation efficiency of these genes so that significantly lower expression is achieved than if genes used ATG. To determine whether the GTG start of *marR* is a historical contingency or the outcome of selection, we constructed the *marR* phylogenetic tree and determined GTG prevalence across *Gammaproteobacteria*. Among 889 representative genomes, *marR* homologs were found in ~300 species, distributed across 20 distinct bacterial families (Fig. 1*A*). Interestingly, we observed that *marR* belongs to the *marRAB* operon only in *Enterobacteriaceae*. Within all other bacterial families, *marR*-type transcription factors form operons with *emrAB*-type efflux pump genes (Fig. 1*A* and *SI Appendix*, Fig. S1).

The phylogenetic tree corroborates an evolutionary scenario in which *marRAB* operon evolved only once and was vertically inherited (Fig. 1*A* and *SI Appendix*, Fig. S2). Its formation in the ancestor

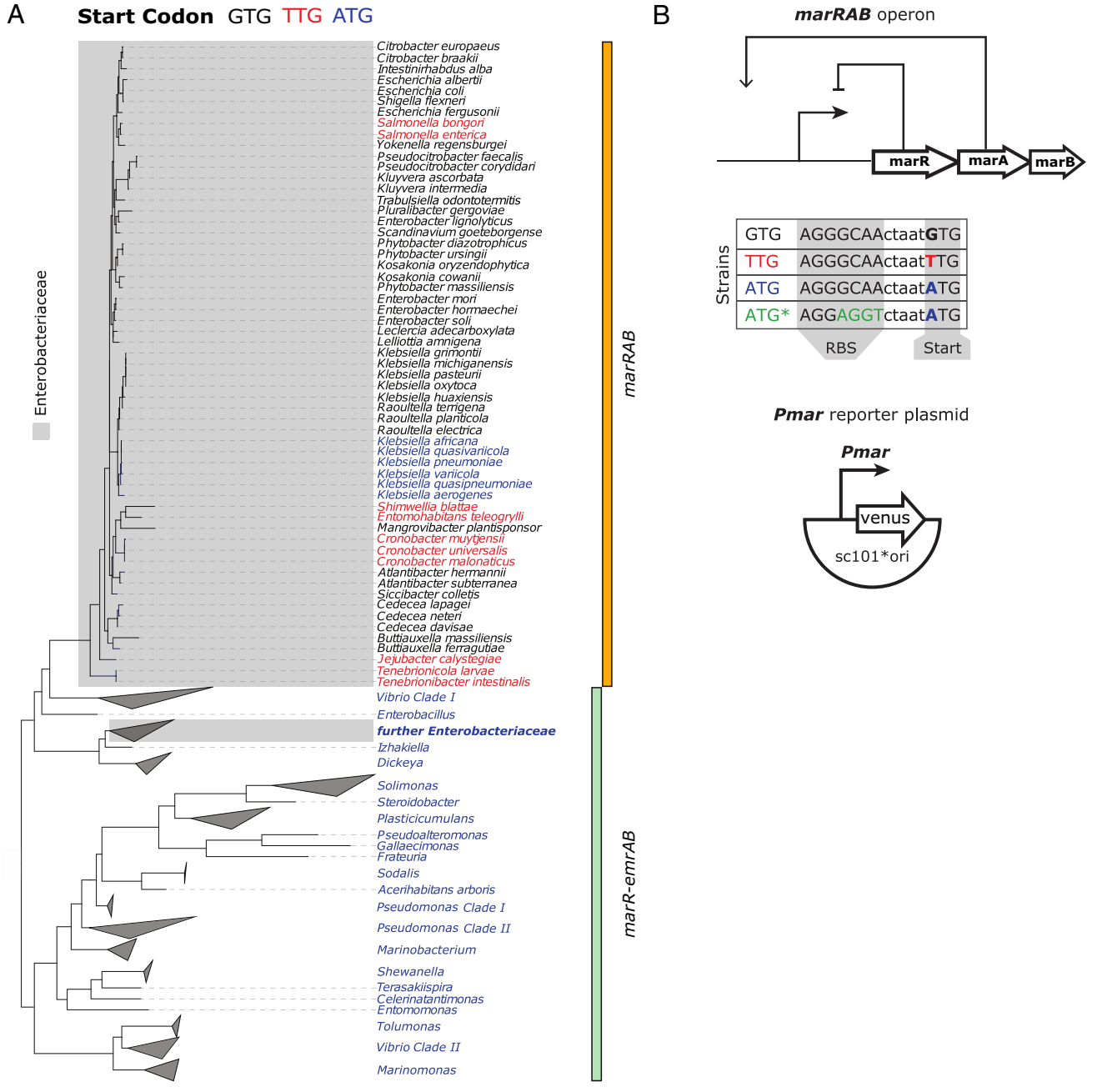

Fig. 1. Distribution of *marR* start codons across *Enterobacteriaceae*. (*A*) Maximum likelihood phylogenetic tree of *marR* in *Gammaproteobacteria*. The tree was constructed using other members of *marR* family (full tree in *SI Appendix*, Fig. S1). The *marRAB* operon is only present in *Enterobacteriaceae*, while in other bacterial families, a *marR*-type transcription factor was found in conjunction with the genes forming an *emrAB*-type efflux pump. (*B*) Schematic representation of *marRAB* operon (*Top*), strains constructed (*Middle*), and reporter plasmid (*Bottom*).

Downloaded from https://www.pnas.org by IST AUSTRIA - LIBRARY; INSTITUTE OF SCIENCE & TECHNOLOGY on April 7, 2025 from IP address 81.223.14.217.

of Enterobacteria coincides with a change of the *marR* start codon from the canonical ATG to the noncanonical GTG. While the *marR-emrAB* family has a strong ATG start codon, *marR* in *marRAB* operons uses the weaker GTG variant, with very few exceptions (*Cronobacter*, *Jejubacter, Pluralibacter, Salmonella, Shimwellia*, *Tenebrionicola*), where the putatively even weaker TTG is used. Furthermore, a switch from GTG to ATG occurred only in one *Klebsiella* clade. Few Enterobacteria harbor both, *marRAB* and *marR-emrAB* operons, (*Cedecea*, *Kosakonia*, *Phytobacter*, *Raoultella*), providing evidence for horizontal gene transfer of *marR-emrAB* into some *Enterobacteriaceae* (Fig. 1*A* and *SI Appendix*, Fig. S2). In addition to the start codon, the ribosome binding site (RBS) is also a determinant of translational efficiency. With a single exception, the RBS of *marR* in *marRAB* is fully conserved among Enterobacteria (*SI Appendix*, Fig. S3). Taken together, our phylogenetic analysis strongly suggests that the prevalent utilization of weak *marR* start codons across *marRAB* operons, in conjunction with a particular RBS variant, is selectively favored for a yet uncharacterized, but likely general, physiological role.

Pulsatile Basal Expression of *mar* Operon. To ask how the conserved GTG start codon affects *marRAB* function and fitness, we constructed scarless *marR* mutants with alternative start codons (ATG and TTG) in *E. coli*. In addition, in the ATG* mutant we combined the ATG start codon with a stronger RBS (Fig. 1*B*).

To investigate the basal *mar* expression mode in single cells, we measured the fluorescence output of a *Pmar-venus* promoter fusion using time-lapse microscopy in a microfluidic device (*SI Appendix*, Fig. S4*A*) (23, 24). We simultaneously monitored a chromosomal constitutive P_R-*mCherry* as a control. Average background-corrected *Pmar* expression depended significantly on the choice of start codon (Fig. 2*A*). TTG yielded threefold to fourfold higher *Pmar* expression

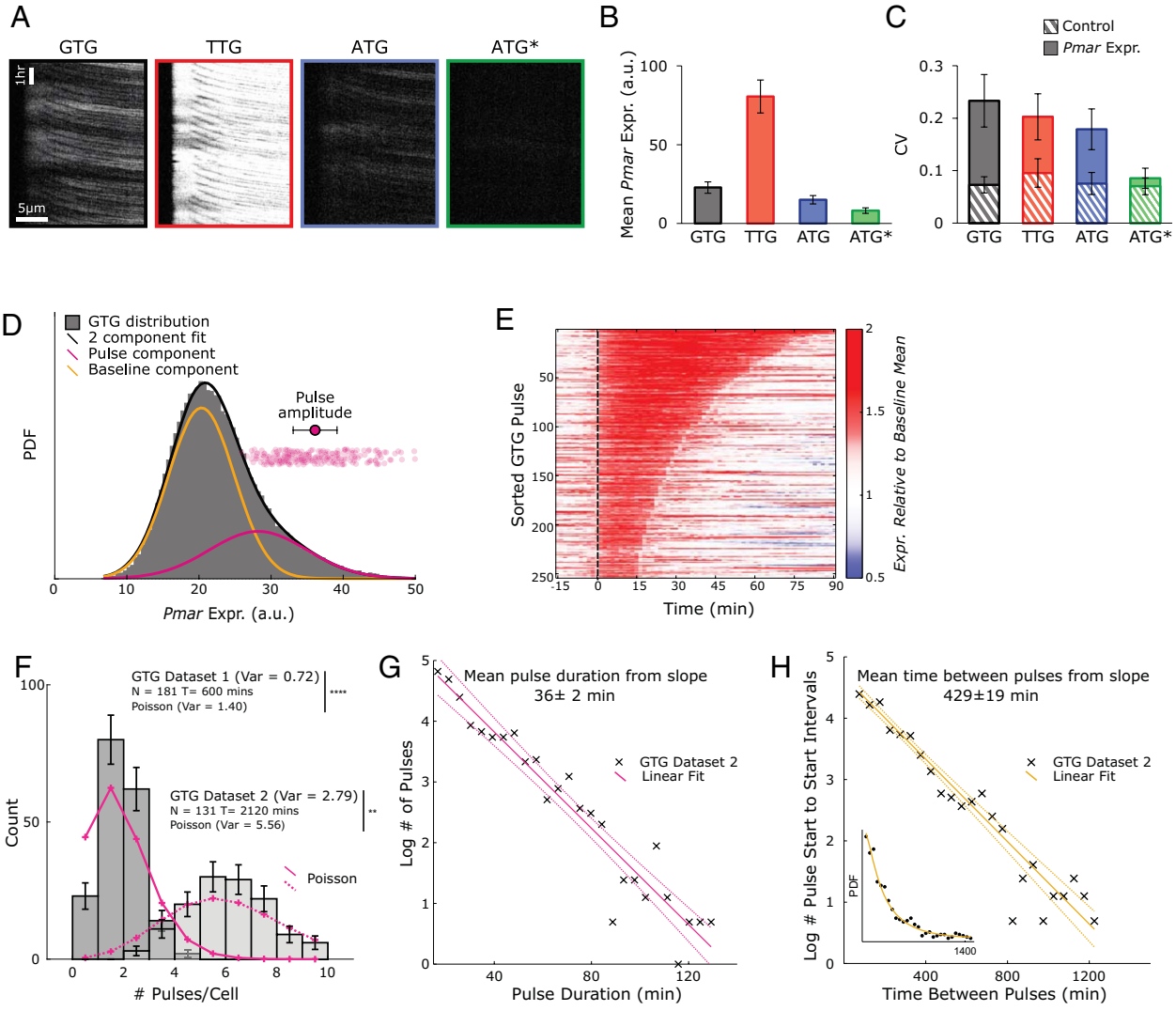

Fig. 2. Characterization of *Pmar-venus* basal expression dynamics. (*A*) Kymographs of wildtype (GTG) and mutants (TTG, ATG, and ATG*) showing *Pmar-venus* expression for a representative mother cell imaged in a microfluidic channel over 10 h. (*B*) Mean *Pmar-venus* expression in wildtype and mutants (Mean and SD over cells, $P < 10^{-4}$, rank-sum test). (*C*) CV of *Pmar-venus* expression (solid colors) in wildtype and mutants compared against CV of control (constitutive reporter P_R-*mCherry* expression, striped colors). *Pmar-venus* expression CV in wildtype is significantly different from mutants (mean and SD over cells, $P < 10^{-4}$, rank-sum test). (*Pmar-venus* expression CV is also significantly different from the corresponding P_R-*mCherry* control for each strain, $P < 10^{-4}$, rank-sum test). (*D*) Distribution of *Pmar-venus* expression is modeled as a sum of two Gaussian distributions, i.e., baseline and pulse component, for wildtype. Magenta circles represent the amplitudes of individual extracted pulses and magenta circle with an error bar is the mean amplitude ± SD over individual pulses. (*E*) Pulses for wildtype sorted by duration and pulse start is aligned to 0 min on the x-axis. (*F*) Distribution of the number of pulses per cell in wildtype in Dataset 1 (181 cells, 600 min) and Dataset 2 (131 cells, 2,120 min, the only difference between datasets is the length of the experiment; otherwise, experiments share the same stable behavior) showing significant deviation from Poisson. Error bars represent √(Count). (*G*) Log histogram of pulse durations. Mean pulse duration (in caption) was derived from the inverse of the slope of the linear fit. (*H*) Log histogram of time intervals between two consecutive pulses. *Inset* shows the corresponding PDF.

Downloaded from https://www.pnas.org by IST AUSTRIA - LIBRARY; INSTITUTE OF SCIENCE & TECHNOLOGY on April 7, 2025 from IP address 81.223.14.217.

than the wildtype (GTG), consistent with the expectation that TTG leads to weaker repressor translation. In contrast, the strong, canonical ATG start codon reduced *Pmar* basal expression below the wildtype (GTG) levels. The ATG* mutant that combines the canonical ATG start codon with a strong RBS, abolished most of *Pmar* expression and accessed a nearly complete OFF promoter state (Fig. 2*B*).

We next characterized the overall variability of the basal expression mode by computing the coefficient of variation (CV) of *Pmar-venus* fluorescence across 10 h of observation for each of the ~180 independent mother cells per genotype (Fig. 2*C*). The wildtype (GTG) showed maximum expression variability, followed by TTG (despite having higher mean expression than that of the wildtype), and then ATG. These three strains have at least twofold higher variability than constitutive controls. For ATG*, the CV was only slightly elevated relative to the control (Fig. 2*C* and *SI Appendix*, Fig. S4 *B* and *C*).

The observed high *Pmar* variability in single cells traces its origin to gene expression pulses: transient, stochastic, high-amplitude activations of transcription plainly visible in all strains (Movies S1–S5). To extract and statistically characterize these pulses (*SI Appendix*, Fig. S5), we first decomposed the observed *Pmar-venus* fluorescence distributions into a Gaussian mixture. The frequent lower-amplitude component corresponded to baseline fluctuating *Pmar* expression levels, whereas the rarer component corresponded to sporadic high-amplitude pulse-like excursions (Fig. 2*D* and *SI Appendix*, Fig. S6). This motivated an 85-percentile threshold (*Materials and Methods*) for extracting the pulses which could subsequently be aligned to their respective start times (Fig. 2*E* and *SI Appendix*, Fig. S7) and quantified.

We report a similar frequency of pulsing in the wildtype (GTG), TTG, and ATG strains, of one pulse per approximately 7 to 8 h. For ATG*, pulses appear to be much less frequent, but our detection may be biased by their low signal-to-noise ratio (Table 1). The pulse duration distribution was exponential for the wildtype (GTG), TTG, and ATG strains for which it could be reliably estimated, with similar average duration of ~33 to 37 min per pulse (Fig. 2*G*, Table 1, and *SI Appendix*, Fig. S8). The key difference between the strains lay in the overall *mar* expression that affects the baseline as well as pulse amplitudes (Table 1 and *SI Appendix*, Figs. S6*D* and S9*A*). This expression changed several-fold depending on the MarR start codon, implying that the MarR translation efficiency can tune *Pmar* expression by determining the promoter activity level outside and during the pulse. When expressed as fold-change increase over their respective baseline Gaussian components—which we refer to as pulse "signal-to-noise" ratio (SNR)—differences between strains were smaller but significant: Pulse amplitudes ranged from ~1.3 to 1.7, with the maximal SNR reached in the wildtype (GTG) (Table 1 and *SI Appendix*, Fig. S9*A*). This difference at the level of pulse characteristics is responsible for the maximal CV in *mar* expression reported for the wildtype (GTG). Finally, after z-scoring and accounting for the individual durations of the pulses, we find that pulses nearly collapse onto a universal shape, indicating that most of the variability across genotypes is accounted for by the statistics we extracted (*SI Appendix*, Fig. S9*B*).

Can pulsing be modeled as a stationary stochastic point process? If that were the case, pulse numbers should be Poisson-distributed over individual cells of the same genotype. We report strong and highly significant deviations from this expectation for the wildtype (GTG), TTG, and ATG strains but not for ATG* and controls (Fig. 2*F* and *SI Appendix*, Fig. S10), even though interpulse intervals are exponentially distributed as shown for wildtype (GTG) (Fig. 2*H*). Specifically, the observed pulse count distributions are underdispersed compared to Poisson, suggesting a more regular pulsing, possibly due to finite pulse duration or pulse–pulse correlations. Despite these quantitative deviations, individual pulses could be interpreted in the bet-hedging framework as stochastic switches into an alternative (high *marA*) phenotype once every ~14 to 16 generations and lasting for about one generation.

We next assessed the cost of *mar* basal mode expression. TTG cells had a significantly lower long-term elongation rate than wildtype (GTG) cells, whereas the elongation rates of ATG and ATG* were marginally higher than for the wildtype (GTG) (Fig. 3*A*). This corresponds to the ordering of mean *mar* expression levels across strains (Fig. 2*B*) and is consistent with the expectation that higher overall *mar* activity is costly. A detailed analysis, however, revealed a surprising finding. We compared the single-cell long-term elongation rates to the instantaneous elongation rates during different phases of the pulse. We expected the elongation rates to slow down around a pulse and subsequently return to the long-term average. In contrast, for all strains but ATG* we observed significantly increased elongation rates in the time window 0 to 20 min after we identify the pulse start (Fig. 3 *B* and *C*). The growth advantage could be caused by differences in pulse amplitudes, where different sets of *mar* regulon targets are engaged by different levels of MarA. This selective targeting is plausible, since genes in the *mar* regulon are known to respond continuously and with different sensitivities to MarA levels (25). Taken together, larger baseline *mar* expression has a cost, while a rare transient pulse confers an advantage. Therefore, selection may have to navigate this tradeoff in an environment-dependent way.

To verify that low and transient (as opposed to high and persistent) *mar* expression during the pulse is necessary for an elongation rate advantage, we exposed ~110 cells per strain in our microfluidic device to 2 mM salicylate to induce *Pmar* expression (Fig. 3*E*). Induction caused a prolonged increase in *marRAB* expression, with the largest, ~sixfold to sevenfold induction in the wildtype (GTG) and ATG, followed by TTG (fourfold to fivefold) and ATG* (~threefold) (Fig. 3*F*). In terms of absolute expression, these levels were substantially (2× to 4×, depending on the strain) above the pulse amplitudes in the basal mode (Table 1). Induction brought about a concomitant ~10% decrease in the elongation rate in all strains (Fig. 3*D*), consistent with previous reports and our expectation that prolonged and strong *marRAB* expression is detrimental likely because it engages additional, more costly-to-express *mar* regulon targets (25).

Table 1. Characterization of pulse expression, amplitude, duration, and frequency in wildtype (GTG), TTG, ATG, and ATG* strains

	GTG	TTG	ATG	ATG*
Baseline mean expression (a.u.)	20.35 ± 4.56	71.57 ± 13.53	13.49 ± 3.02	7.36 ± 1.04
Pulse mean SNR (amplitude/baseline mean)	1.69 ± 0.01	1.59 ± 0.01	1.55 ± 0.01	1.33 ± 0.02
Pulse amplitude relative to wildtype	1.00	3.39	0.61	0.30
Pulse mean duration (min)	35 ± 20	37 ± 23	33 ± 16	25 ± 11
Pulse frequency (per h)	0.14	0.12	0.15	0.04

Downloaded from https://www.pnas.org by IST AUSTRIA - LIBRARY; INSTITUTE OF SCIENCE & TECHNOLOGY on April 7, 2025 from IP address 81.223.14.217.

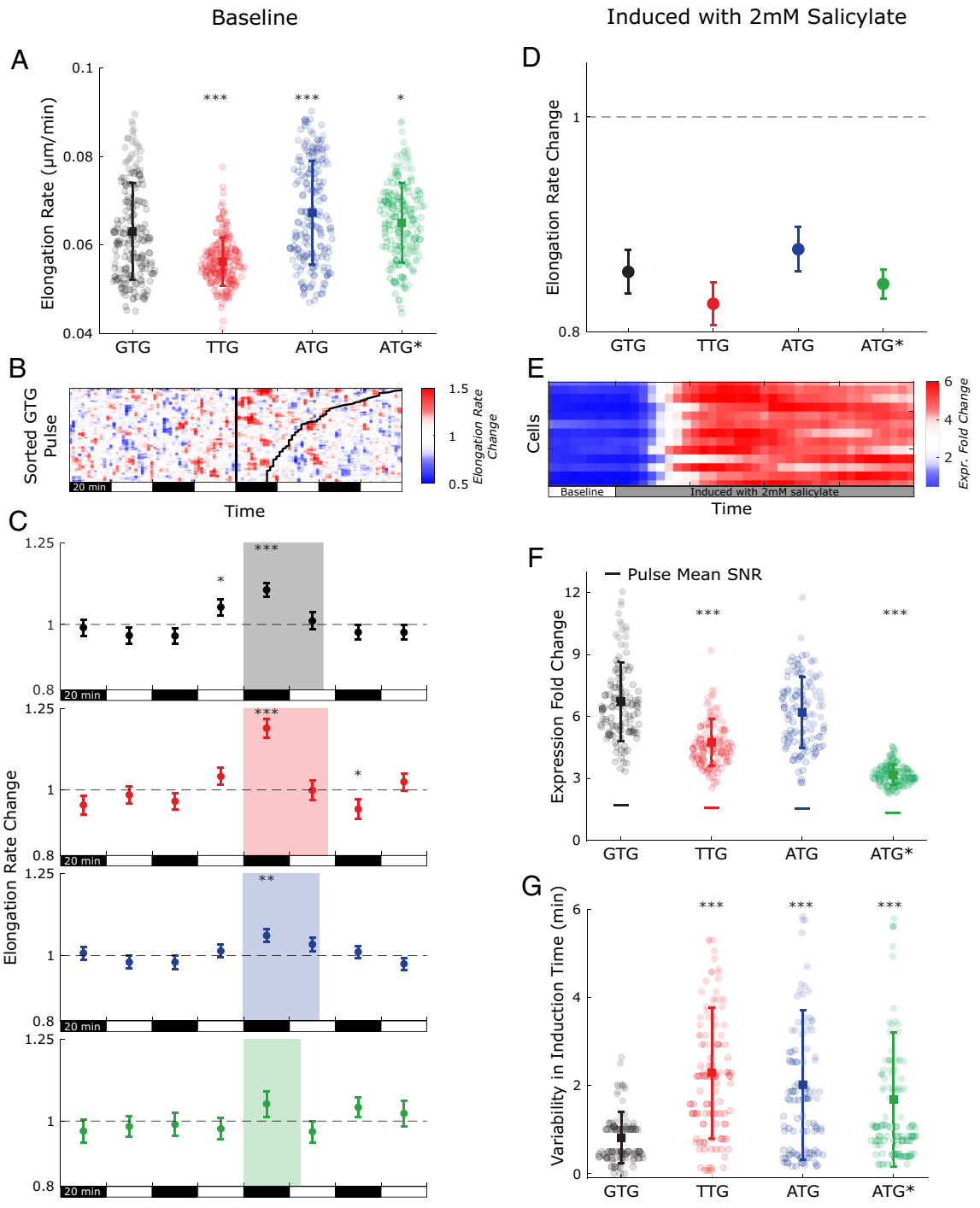

Fig. 3. Elongation rates and *Pmar-venus* expression quantification in baseline (*A*–*C*) and induced state (*D*–*G*). (*A*) Single-cell elongation rates in the baseline state [individual cell data (circles) with mean (square) and SD]. Wildtype elongation rate is significantly different from TTG & ATG ($P < 10^{-3}$, rank-sum test) and ATG* ($P < 0.05$, rank-sum test). (*B*) Raw data showing instantaneous elongation rate normalized to the long-term individual cell average and aligned to pulse start (vertical black line). Pulse duration is denoted in black as well. (*C*) Mean and SE of elongation rates (normalized to the long-term average elongation rate of the respective cell) in 20 min windows (80 min before and after the pulse start) for wildtype and mutants. The shaded regions illustrate the start of the pulse and the mean pulse duration in each strain. Stars indicate significant deviation from the long-term elongation rate ($P < 0.05$, *t* test). (*D*) Mean and SE of elongation rate change upon induction with 2 mM salicylate relative to their respective baselines shows similarly sized fold change relative to preinduction baseline and statistically significant reduction across strains ($P < 10^{-4}$, *t* test). (*E*) Representative raw data (*Pmar-venus* expression) showing the transition from baseline to 2 mM salicylate-induced state. (*F*) Individual cell data (circles), mean (square), and SD of fold change of *Pmar-venus* expression upon induction with 2 mM salicylate, the horizontal line shows a comparison with pulse mean SNR in the baseline state. Wildtype expression fold change upon induction is significantly different from TTG & ATG* ($P < 10^{-4}$, rank-sum test) and ATG (not significant at 0.05 significance level, rank-sum test). (*G*) Individual cell data (circles), mean (square), and SD of variability in induction time quantified by the time taken by expression to reach the inflection point for each cell. Wildtype is significantly different from all three mutants ($P < 10^{-4}$, rank-sum test).

The induction experiment revealed another significant difference between the wildtype (GTG) strain and the strain with the canonical start codon (ATG), which emerged when we analyzed the inflection times of individual cell induction curves (Fig. 3*G*). The wildtype (GTG), which exhibits maximal expression variability in the basal mode, surprisingly had the least variable induction curves and thus the most synchronized response to induction. In comparison, the mutants show a twofold reduction in synchronicity.

Downloaded from https://www.pnas.org by IST AUSTRIA - LIBRARY; INSTITUTE OF SCIENCE & TECHNOLOGY on April 7, 2025 from IP address 81.223.14.217.

Overall this suggests that the unique interlocked regulatory circuit with short-lived MarA results in precise induction to sudden stress (26), indicating that response speed and synchrony may be at a premium for the enterobacterial ecological niche.

Origin of Pulsatile Basal Expression in *mar* Promoter. Before further dissecting the fitness effects of *mar* in environments where *mar* expression mode transitions and timing could be relevant, we wanted to better characterize the differences between the basal pulsatile mode and the induced expression. To this end, we analyzed *Pmar* expression in single cells after induction with 2 mM salicylate, specifically to see whether expression pulses persist. The mean expression levels increased significantly relative to the uninduced state (Fig. 4*A*), with increases following the same rank order as the order of the basal expression, i.e., TTG > GTG > ATG > ATG*. Despite this increase in mean expression, variability significantly decreased from the uninduced state (Figs. 2*C* and 4*A*). Furthermore, during induction, no significant difference in CV was observed between different start codon mutants and wildtype. This strongly suggests that the increased CV of the basal mode originates from the pulsatile nature of gene expression; no such pulsatile characteristics could be detected in the induced data. Continuous expression, in contrast, results in significantly reduced growth rates in all strains, as shown in Fig. 3*D*. Taken together, these findings indicate that pulsatility is an inherent feature of the basal state, and that maximizing variability via pulsing leads to lower fitness costs.

An orthogonal insight into basal expression and its variability is afforded by looking at wildtype cells under decreasing concentrations of inducer (Fig. 4*B*). While the mean expression increased monotonically with increasing salicylate, expression CV was significantly and substantially higher in the basal mode compared to all of the induced levels, and the CV, when induced, showed no clear trend with the inducer concentration (Fig. 4*B*).

To elucidate the mechanism behind the pulsatile basal expression, we decided to explore Δ*marA* and Δ*marR* deletion mutants. Δ*marR* showed significantly higher *Pmar* mean expression and lower CV, i.e., it exhibited expression characteristics that are very similar to the fully induced wildtype (GTG) state. On the other hand, Δ*marA* showed similar *Pmar* expression to the wildtype (GTG) baseline, yet with the CV that was significantly lower (Fig. 4*C*). We analyzed pulse characteristics of Δ*marA* and Δ*marR,* analogous to Fig. 2, and found that *Pmar* expression in Δ*marR* exhibited no pulsatile features and appeared qualitatively consistent with noisy, constitutive-type gene expression (*SI Appendix*, Fig. S11 *A* and *B*). In contrast, in Δ*marA*, pulse-like fluctuations were observed, even though with smaller and more variable magnitudes relative to the baseline compared to the wildtype (*SI Appendix*, Fig. S11 *D* and *E* and Table S1). These observations suggest that pulses originate when the repressor, expressed mainly during the previous pulse, is sufficiently diluted by subsequent cell divisions. Intriguingly, this hypothesis correctly predicts the change in pulse frequency in TTG and ATG mutants given the change in *Pmar* expression during the pulse. In contrast, *marA* increases the pulse amplitudes and stabilizes their stereotypic form, but is not essential to their generation. Neither deletion mutant exhibited any growth advantage during the pulse-like fluctuations, as shown in *SI Appendix*, Fig. S11 *C* and *F*. Taken together, these analyses show that the origin of the pulses lies in the binding properties of MarR and that the overall expression variability can be modulated by the selection of different start codons, with GTG being optimal for the highest variability in basal expression.

Fitness Advantage for the Wildtype Basal Expression Mode across Growth Cycles. So far, data suggest a role of *mar* expression in physiology and growth homeostasis, which is in line with the subtle influence of *mar* in the transition from exponential to stationary phase and back (27, 28). We therefore decided to conduct pairwise competition experiments to assess the performance of the start codon mutants across the entire growth cycle. We competed the wildtype (GTG) vs. ATG, vs. TTG, vs. ATG*, vs. Δ*marA* strains, and vs. GTG itself (as a control), across four serial growth cycles over

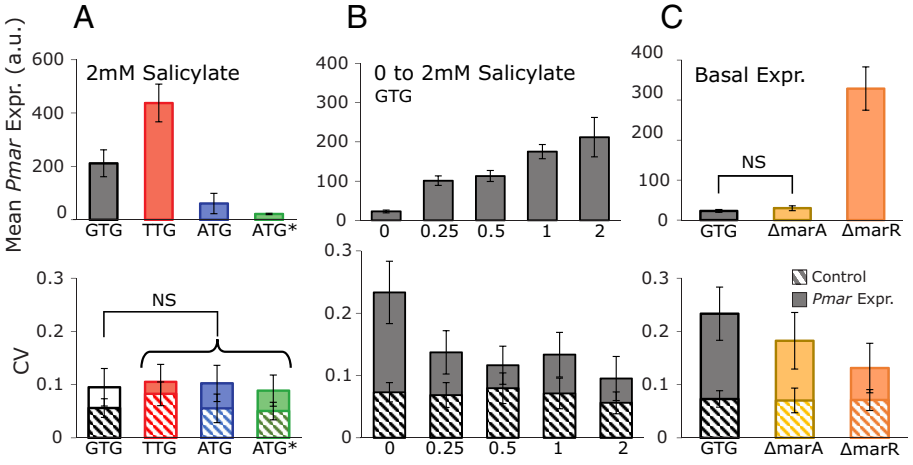

Fig. 4. Characterization of *Pmar-venus* expression in the presence of the inducer salicylate. (*A*) *Pmar-venus* expression in wildtype and mutants TTG, ATG, and ATG* with 2 mM salicylate (mean and SD over 112 cells, $P < 10^{-4}$, rank-sum tests for significant difference with wildtype) (*Top*). CV of *Pmar-venus* expression (solid colors) and CV of constitutive reporter P_R-*mCherry* expression (striped colors) with 2 mM salicylate. CV of wildtype is not significantly different from mutants at $P < 0.01$ level (mean and SD over cells, *t* test) (*Bottom*); however, it is significantly different from each respective control P_R-*mCherry* expression CV ($P < 10^{-4}$, rank-sum test). (*B*) Mean *Pmar-venus* expression in wildtype (GTG) in presence of 0 (181 cells), 0.25 mM (91 cells), 0.5 mM (97 cells), 1 mM (86 cells), and 2 mM (112 cells) salicylate; error bars are SD (*Top*). Uninduced wildtype expression is significantly different from all levels of induced expression ($P < 10^{-4}$, rank-sum test). CV of *Pmar-venus* expression in wildtype (solid colors) and CV of constitutive reporter P_R-*mCherry* expression (striped colors) (*Bottom*). *Pmar-venus* expression CV in uninduced wildtype is significantly different from induced *Pmar-venus* expression (mean and SD over cells, $P < 10^{-4}$, rank-sum test). (*C*) *Pmar-venus* expression in wildtype (181 cells), Δ*marA* (115 cells), and Δ*marR* (98 cells) mutants (mean and SD over cells, no significant difference for the mean Δ*marA* from wildtype, $P > 0.05$; and significant difference for Δ*marR,* $P < 10^{-4}$; rank-sum test) (*Top*). CV of *Pmar-venus* expression in wildtype, Δ*marA*, and Δ*marR* cells (solid colors) and CV of constitutive reporter P_R-*mCherry* expression (striped colors). CV of *Pmar-venus* expression in wildtype is significantly different ($P < 10^{-4}$, rank-sum test) from mutants and respective control P_R-*mCherry* expression (*Bottom*).

Downloaded from https://www.pnas.org by IST AUSTRIA - LIBRARY; INSTITUTE OF SCIENCE & TECHNOLOGY on April 7, 2025 from IP address 81.223.14.217.

4 d (together >40 generations) in LB media without any external inducers. The key question was how wildtype (GTG) compares against ATG, ATG*, (the two strains with more efficient MarR repressor translation and thus lower baseline *mar* expression), TTG (the strain with comparatively weaker MarR repression), and Δ*marA* (the strain with similar baseline expression but no MarA molecules present), when bacteria are forced to undergo repetitive transitions between exponential, stationary, and lag phases.

Starting from a 1:1 ratio, we saw the wildtype (GTG) increase to a ratio of 2:1 over the course of 53 generations in competition with the ATG strain; the same 2:1 ratio was reached in 35 generations in competition with ATG* and TTG. The effective selection coefficients were −0.013, −0.020, and −0.022 per generation for ATG, ATG*, and TTG respectively, relative to wildtype (GTG) (Fig. 5*A* and *SI Appendix*, Fig. S12*A*). Δ*marA* vs. GTG competition showed that absence of MarA confers a significant advantage. Δ*marA* has comparatively higher exponential growth rate as shown in the single cell elongation rates (Fig. 3*A* and *SI Appendix*, Fig. S11*G*). This is consistent with the cost of MarA expression during the exponential growth phase and highlights the importance of low baseline expression. To further quantify the *marA* expression cost, we repeated the competition experiment in the presence of 2 mM salicylate. Here, we observed that GTG loses in competition experiments to ATG, ATG*, and Δ*marA*, but wins against TTG (*SI Appendix*, Fig. S13). This pattern follows the order of *mar* expression except in the case of Δ*marA*, where the downstream physiology differs significantly, as there will be no regulation of any downstream targets by MarA.

The fitness advantage of the wildtype (GTG) in the absence of external inducers suggests a functional role of the conserved weaker GTG start codon for cell physiology during serial growth cycles, likely through its effects on the pulsatile basal mode of *mar* expression. To understand how the wildtype (GTG), which is at a disadvantage during exponential growth, outcompetes the start codon mutant strains, we measured how quickly various strains recovered from the lag phase and transitioned to exponential growth. We report a delay of around 8 to 12 min for ATG, ATG*, and Δ*marA* mutants vs. the wildtype (GTG) in LB medium that matched the conditions used in the competition experiment. The delay further increased by up to 25 to 50 min for nutrient-poor M9 media (Fig. 5*B* and *SI Appendix*, Fig. S12*B*). For TTG, as expected, we find that lag recovery happens marginally faster than the wildtype.

In summary, these results suggest that the wildtype (GTG) strain compensates for its slower exponential growth rate through an apparent shorter population lag time in the basal expression mode. It is instructive to consider a simple back-of-the-envelope calculation using realistic estimates from Fig. 5. If the long-term exponential growth rate of the wildtype (GTG) is ~5% slower than that of the ATG mutant but it emerges from lag phase with a 24 min advantage, then the ATG strain would require 8 h of uninterrupted growth to compensate for its delay and reach the same population size as the quicker-to-emerge but slower-to-grow wildtype strain.

An instructive way to visualize the various tradeoffs is to depict the growth rate, lag recovery time, and the overall selection coefficient of all our strains in a single plot (Fig. 5*C*), reminiscent of Pareto frontiers analyzed in multiobjective optimization (29). Wildtype (GTG) exhibits an optimal balance between growth rate and lag recovery. TTG shows faster lag recovery, but pays the cost with significantly slower exponential growth rates. ATG, and ATG* are worse off in lag recovery, but have comparatively faster elongation rates in the exponential phase. In contrast to start codon mutants and the wildtype that navigate the same growth/lag tradeoff, Δ*marA* is qualitatively different: This deletion mutant cannot generate pulsing-related heterogeneity in downstream MarA targets. Even though it has nearly the same growth rate and lag time as the ATG mutant, it nevertheless achieves a selection advantage in the lab competition experiment against the wildtype, pointing to a significant change in cell physiology of the deletion mutant. Nevertheless, this mutant will lack MarA-dependent functions, especially the pulse-generated variability in MarA downstream targets that could support bet hedging, and which

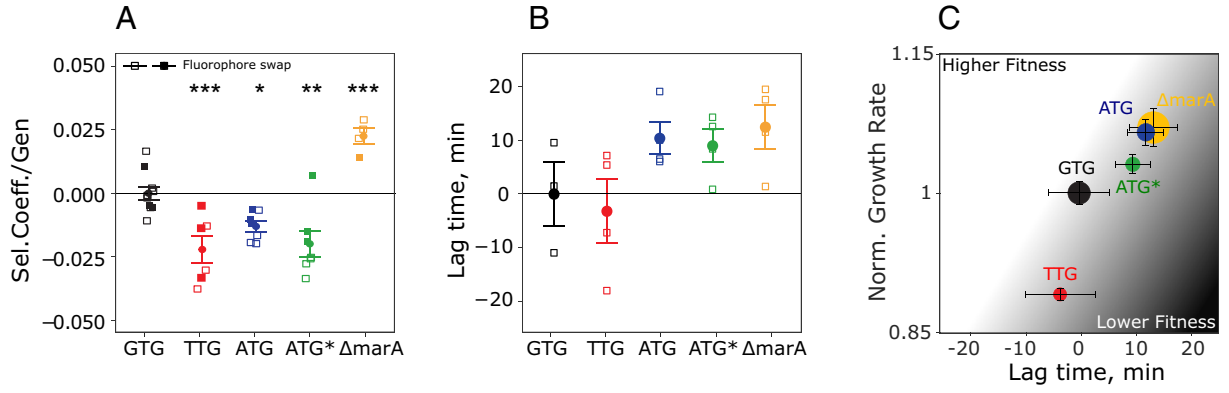

Fig. 5. Competition, regrowth dynamics of mutants and wildtype (GTG), and evolutionary tradeoffs. (*A*) Selection coefficient (square) was quantified from the slope of the log ratios of competitor over wildtype strain per generation (the circle represents mean and error bar represents SE over squares). TTG, ATG, and ATG* lost the competition against the wildtype as determined by negative selection coefficients that were significantly different from the wildtype–wildtype control competition (s = −0.02, $P < 10^{-3}$ for TTG; s = −0.01, $P < 0.05$ for ATG; s = −0.02, $P < 10^{-2}$ for ATG*, post hoc test and GLM). Δ*marA* wins the competition experiment as shown by positive selection coefficient (s = 0.03, $P < 10^{-3}$). mCherry (empty squares) and Venus fluorophores (filled squares) were used as markers to select for the respective strain background, and six to seven biological replicates across two independent experiments, including fluorophore swaps, were performed for each combination. (*B*) Time to regrow from stationary phase. Data points (square), mean (circle), and SE of lag time to initiate exponential growth when revived from 16 h overnight LB (fresh rich media) culture. Sample size was three to four biological replicates per strain, each calculated from eight technical replicates. ANOVA of lag time delays for the four experimental regrowth conditions [Unadjusted LB (fresh rich media) culture; OD Adjusted—LB (fresh rich media), 48 h LB culture (aged rich media), and 48 h old M9 glycerol culture (aged poor media), as shown in *SI Appendix*, Fig. S12) showed a significant effect of strain (F = 6.6, $P < 10^{-3}$, ANOVA) and a significant effect of regrowth condition (F = 7.4, $P < 10^{-3}$, ANOVA). (*C*) Wildtype and mutants navigate a tradeoff between the lag (x-axis; three to four biological replicates as shown in panel (*B*) and the growth rate (y-axis; ~180 biological replicates for each genotype, elongation rates from Fig. 3*A*). Size of the circle represents the selection coefficient in the competition experiment [calculated from six to seven biological replicates from panel (*A*)]; shown are mean and SE (error bars) over replicates.

Downloaded from https://www.pnas.org by IST AUSTRIA - LIBRARY; INSTITUTE OF SCIENCE & TECHNOLOGY on April 7, 2025 from IP address 81.223.14.217.

are likely relevant for the more complex Enterobacteria growth cycles characteristic of life in the wild. Such cycles would entail richer, multidimensional tradeoffs between phenotypes beyond just growth rate and recovery lag—perhaps including generating the variability in expression—and could rationalize the evolutionary maintenance of the MarA regulatory architecture.

Discussion

The process of turning genes ON and OFF is fundamental to life, and regulatory networks that control it have been the object of intense study in developmental, evolutionary, molecular, and systems biology. Nevertheless, the question of whether the properties of the basal expression in the OFF state of promoters can be selected for was rarely, if at all, considered. We showed that the *mar* operon basal expression mode is highly dynamic, consisting of pulsatile gene expression at the single cell level, and we measured the fitness consequences of such dynamics. To this end, we used synonymous codon mutations for the start codon of *marR*, which we uncovered to be evolutionarily conserved and tightly coupled to the interlocked regulatory architecture of the *marRAB* operon.

We find that the weaker, noncanonical GTG start codon acts as a regulatory knob ensuring the presence of sporadic transcription pulses that are much less pronounced or absent if *marR* and *marA* are translated with similar speed, as is the case in the various start codon mutants we measured. The basal transcription pulses disappear in the *marR* deletion mutant. Given the very tight (low nM) binding of MarR to its operators (30), this is consistent with a negative feedback-with-delay mechanism of pulse generation, where short-lived MarA has an auxiliary role of pulse amplification and sharpening, along with being the main downstream "effector" for the *mar* circuit. Thus, on top of this interlocked feedback architecture, the GTG start codon endows the wildtype basal expression mode with several unique characteristics: highest expression variability due to highest "signal-to-noise" ratio of individual pulses, transient growth advantage during a pulse, and synchronized expression during transition to the induced state. Altogether, these quantitative expression characteristics lead to a robust fitness advantage for the wildtype that becomes apparent in environments in which cellular physiology switches between exponential, lag phase, and stationary growth. Interestingly, knocking out *marA* had similar disadvantage in lag time as start codon mutants, but showed overall better fitness in cycling experiments, indicating an overall cost of sporadic expression of *marA* targets in the basal mode. Compared to natural habitats, our laboratory conditions are highly stereotyped and limited in environmental range: e.g., growth cycling in the wild occurs among very different food sources given the wide diversity of daily diets an animal has, under more temporal variability, and in the presence of a much wider diversity of microbiological ecologies. The maximal variability in basal expression of the wildtype has hence likely evolved to maximize fitness across these many different environments. More broadly, our findings point to an involvement of *marRAB* operon in general growth homeostasis across myriad environments, which widens the role of the *mar* regulon in the ecophysiology of Enterobacteria.

By studying specific mutants, such as start codon mutants that allow for a fine modulation of MarA levels rather than complete knockout of the *marA* gene, we can more precisely analyze the expression dynamics of the *mar* system without severely disrupting the *mar* regulatory network. This genetic approach enables us to explore how variations in MarA levels impact the overall system behavior, while maintaining the main functional interactions within the regulatory network of MarR and MarA at the *mar* promoter.

While the *marRAB* operon can be induced by various metabolic intermediates, no main physiological inducers have been characterized so far (31) and conspicuously its highest induction levels occur at lower temperatures (10), implying a potential role for free-living conditions (32, 33). Moreover, the presence of nearly identical *mar* systems across most Enterobacteria and the remarkable conservation of the noncanonical GTG start codon of *marR* suggest that the cellular processes controlled by the *mar* operon confer adaptations that are most likely linked to Enterobacteria physiology and their ecological niche (33). Essential to this ecological niche is that bacteria spend a significant fraction of their existence inside the guts of various animals. The quasi-regular intervals of feeding that are characteristic of the gut environment point toward a selective pressure for optimizing cycles of lag, exponential, and stationary phases. Our results support the idea that the *mar* pulsatile basal expression mode could be an essential building block of this physiological adaptation to the intestinal lifestyle, allowing the basal expression mode to be evolutionarily maintained.

A supporting observation for this hypothesis is the quantitative match between the multiple timescales related to the *mar* operon in *E. coli*: the typical time between two *mar* pulses, the timescale at which the shorter lag of wildtype would balance out its slower growth compared to *marR* start codon mutants, and the daily feeding cycles, which are all on the order of ~10 h. In the bet-hedging framework, this could represent a quantitative case of frequency matching: Yet instead of thinking of very rare stresses (such as antibiotic exposure), the system has evolved to match the frequency of environmental transitions typical for the ecological niche of Enterobacteria. Natural selective pressure to maintain the *mar* basal expression mode for such physiological *raison d'etre* would enable the same system to be co-opted for other hypothesized functions in its induced mode: to generate diversity necessary for bet-hedging against much rarer stresses, or to induce costly stress-resistance genes when needed, and thus allow MarA to become a global transcriptional regulator. In addition to cyclic nutrient availability, the gut ecology is also characterized by host defense mechanisms and antimicrobials secreted by other microbiota. Thus, co-opting regulation of physiological adaptation to cyclic nutrient availability and antibiotic resistance mechanisms would be consistent with the global regulatory role of *mar* (34).

The naming of the *mar* operon—an acronym for *m*ultiple *a*ntibiotic *r*esistance coined over 40 y ago by George and Levy (9, 35)—was based on the conferred broad antibiotic resistance phenotype in the induced state (36). By now, it is clear that the *mar* operon has not evolved "for" antibiotic resistance, which is at best an ancillary function—albeit one of fundamental importance for public health. To control it and counteract the looming multidrug resistance epidemic, our work demonstrates the acute need to understand the role of *marRAB* operon in bacterial physiology and growth homeostasis, in line with Seoane and Levy who argued early on for an alternative role of *mar* as a conveyor of "multiple adaptational response" (10). Beyond the *mar* operon and regulon, these results reveal that the choice of different start codons can be a highly effective genetic knob for fine-tuning the dynamics and variability of complex gene regulatory networks. Paying closer attention to individual conserved noncanonical start codons can thus reveal surprising insights into the underlying regulatory mechanisms.

Materials and Methods

Computational Genomics. Genomes of *Gammaproteobacteria* were downloaded from RefSeq database (37) using the PanACoTA pipeline (38), module "download" with filters: genome collections = "Reference" or "Representative";

Downloaded from https://www.pnas.org by IST AUSTRIA - LIBRARY; INSTITUTE OF SCIENCE & TECHNOLOGY on April 7, 2025 from IP address 81.223.14.217.

assembly level = "complete genomes" or "chromosome" or "scaffold." Then we used the PanACoTA module "annotate" to predict and functionally annotate CDS in the genomes.

marR genes were found using HMM for OG #1S2AX comprising *E.coli marR* (UniProt entry ID:P27245), from eggNOG5 database with e-value threshold 10^{-35} (39, 40). We additionally used HMMs for OG #1S26B and OG #1RPCJ which contain *slyA* (UniProt entry ID:P0A8W2) and *mprA* (UniProt entry ID:P0ACR9), correspondingly to include into the analysis more members of *marR* family in order to better resolve the gene tree (*SI Appendix*, Fig. S1). Species tree was constructed using concatenation of alignments of 39 ribosomal proteins: *rpsF, rplE, rpmA, rpsD, rplV, rpmF, rpsS, rpsH, rpmB, rplC, rpmC, rplN, rplX, rpsM, rplI, rpsR, rplT, rpsC, rpsE, rplO, rpsG, rplW, prmB, rplM, rpsO, rplK, rpmG, rplQ, rpmE, rplB, rpsP, rplJ, rpsJ, rpsQ, rplP, rpsA, rpsB, rpsI, rplD* (*SI Appendix*, Fig. S2).

To find ribosomal binding sites (RBSs) and correct putative errors in gene start prediction, *marR* upstream regions of 100 bp length were extracted and aligned taking into account the known RBS of *marR* in *E. coli* (21). Four downstream genes were used for the annotation of the *marR* genomic context.

The alignments of protein sequences and upstream regions were made using Muscle v. 3.8.31 (41). The gene trees were constructed using IQtree v.1.6.12 with 1,000 bootstrap runs (42). Visualization and annotation of the trees was done using Itol server (43). The logos of upstream regions were created using WebLogo 3 web-based application (44).

Strains and Media. All experiments were performed using the derivate of *E. coli* K-12 MG1655 strain, with incubations at 37 °C and aeration.

Lysogeny Broth (LB) media were used in all experiments except when noted. For plates, 1.5% agar was added to LB. For microfluidics experiment, 0.01% Tween 20 was used to prevent attachment of bacterial cells to the PDMS device. Media and antibiotics were from Sigma, Sylgard for making PDMS was from Dow Chemicals. List of strains and primers used in this study are listed in *SI Appendix*, Tables S2 and S3.

Strain construction. The DIRex method was used to generate scar-free point mutations for changing the start codon and RBS for *marR* (45). Briefly, the method uses a single λ Red recombineering step (pSIM5-Tet temperature sensitive) and a semistable AcatsacA intermediate. The desired changes were introduced through custom-made oligos with a homology to the target region. The AcatsacA cassette codes for three genes which help in selection and counterselection steps a) *cat* leading to chloramphenicol resistance (use 12.5 mg/L chloramphenicol to select for AcatsacA formation), b) *amilCP*, present as dual inverted copies, causing AcatsacA+ colonies to be of blue color helping in selection, and c) *sacB*, sucrose sensitivity gene (use 5% sucrose to select for self-excision), helps in counterselection generating scar-free mutants. We use a constitutive chromosomal P_R-*mCherry* reporter as control (23). *Pmar-venus* reporter is on a low copy plasmid (24).

Microfluidics Set Up and Imaging. We used the same microfluidic chip as previously used in our lab (23). The length of the growth channel is 23 μm and the width of these growth channels ranged from 1.2 μm to 1.4 μm and the height is approximately 1.1 μm (*SI Appendix*, Fig. S4*A*). To make mother machine devices from Epoxy replica, we used PDMS in a ratio of 10:1 (Sylgard and curing agent) and mixed and degassed in Thinky Machine (THINKY ARE-250) for 2 min each. Further degassing was done after pouring on the epoxy replica using a desiccator. Curing of PDMS was done overnight in an incubator at 80 °C. Next, the PDMS device was peeled out carefully from the epoxy replica and holes were punched using an electropolished 18ga needle. The device was cleaned with scotch tape and the coverslip (24 mm × 50 mm, thickness 0.17 mm ± 0.005) was cleaned with isopropanol. Device was then bonded using the plasma bonding technique (Harrick PDC-002 plasma cleaner, medium power for 1 min, for both PDMS and coverslip) followed by gently placing on the coverslip. After bonding, it was kept on a hot plate (~80 °C) for 1 h.

Before starting the experiment, the device was wetted with 0.01% Tween 20 for a couple of minutes followed by blowing out. This step also ensures that the bonding is leak-proof. Next, a pellet from exponential grown cells (overnight culture of the desired genotype in LB plus Tween 0.01% is grown and subcultured 1:1,000 and grown for around 4 h and centrifuged at 4,000 ×*g* for 3 min) was loaded using a pipette. After confirming the loading of cells by checking under the scope, media flow was connected with polyethylene tubing (BTPE -50Instech). Image acquisition settings were kept identical throughout all experiments (Exposure time for *mCherry*: 200 ms and *venus*: 300 ms) with an image interval of 90 s. Images were acquired with an Olympus IX83 inverted fluorescence microscope, a 100× NA 1.45 objective, with a custom-made autofocus, and a Hamamatsu Orca Flash4.0v2 camera (46).

Image Analysis.

Pre segmentation. Images of the channel areas within the microfluidic chip were cropped and background and shading corrected (46). The background correction was done to correct for both autofluorescence of PDMS and the Venus background signal (mean Venus signal of mCherry cell without the Venus fluorophore). The choice of fluorophores was based on the least possible bleed through emission between mCherry and Venus filters. Our background correction takes care of the background Venus signal's impact on CV quantification.

Segmentation. Bacteria segmentation was carried out using Cellpose (47). A custom model was trained on a dataset of over 2,000 hand-labeled cells selected from a diverse range of expression levels and morphologies.

Tracking. A customized Matlab script was employed for tracking, generating lineage trees, and conducting further analyses. In the initial stage, the microfluidic channels were automatically detected, and cells located outside the channels were eliminated. A heuristic method was applied to address missing cell detection and correct undersegmentation errors. Subsequently, each channel underwent individual tracking: The link cost function that establishes connections between cell detections at consecutive time points to form tracks, took into account the specific characteristics of the mother machine. This was achieved by assigning a higher link cost to reverse movement, instances where cells swapped positions within the channel, and a reduced overlap in cell segmentation compared to the segmentation at the previous time point. Cell divisions were identified when two cells overlapped with the same segmentation from the preceding frame. Empty channels were omitted for clarity.

Growth rate. We defined and quantified the growth rate and the promoter activity as done by Kim et al. (6, 48). The growth rate g is calculated as the logarithm of the ratio of the area of the cell immediately prior the cell division A_d and the area at the initial time point following the last cell division A_0: $g = \log\left(\frac{A_d}{A_0}\right)/\Delta t$.

As we did not observe any long-term change in the average cell size, and we set $\left(\frac{A_d}{A_0}\right) = 2$.

According to this definition, it can be inferred that the growth rate remains constant between individual cell divisions.

Instantaneous elongation rate and cell division events. The instantaneous elongation rate R is calculated as $R = \frac{dL}{dt}$. We derived the length L from a linear fit of the cell segmentation area, rather than an ellipse fit to the segmented cell, as this method was more robust against segmentation errors. To smooth out the growth rate, we applied a running average over a 15-min window. Cell division events were identified by employing criteria that recognize when the segmented area of a cell approximately halves during division. Cell division events were excluded from the growth rate analysis.

Pulse detection and analysis. The distribution of combined raw fluorescence intensities is accurately modeled as a sum of two distinct Gaussian distributions. This modeling provides a natural cutoff value, effectively differentiating baseline expression fluctuations from stochastic pulse-like expressions. Using this cutoff value, we perform a z-transform on each cell's raw fluorescence intensities, utilizing the mean and SD of the baseline expression. Pulses are identified by applying a threshold to the z-score (*SI Appendix*, Fig. S5). This method is employed for both fluorescence derived from the *Pmar-venus* promoter fusion and the chromosomal constitutive P_R-*mCherry* expression. The purpose for setting the threshold is to detect pulses which are not due to growth rate fluctuations and hence we use the *mCherry* expression data to decide on the threshold. Additionally, we conducted a stochastic simulation replicating the characteristics of the constitutive expression. Analysis of the pulse length histogram in all three cases supports a cutoff for determining the duration of genuine gene expression pulses: Pulses shorter than 15 min are deemed random baseline fluctuations, whereas longer pulses, indicative of actual gene expression, are compiled for further examination. A linear regression on the inverse slope of the histogram of true pulses provides an average pulse length comparable to the directly calculated mean pulse duration. We also fitted Poisson distributions to the number of pulses per cell.

Downloaded from https://www.pnas.org by IST AUSTRIA - LIBRARY; INSTITUTE OF SCIENCE & TECHNOLOGY on April 7, 2025 from IP address 81.223.14.217.

For visualization, pulses are aligned temporally, setting the start time of each pulse to t = 0 and sorting them by duration. This approach allows for the calculation of the average intensity over time from *Pmar-venus* by normalizing each pulse's intensity to its peak value. However, the transient pulse intensity is affected by two factors: the transient nature of each pulse and the distribution of pulse duration. To isolate the impact of varying pulse duration, we normalize the duration of each pulse to one. Consequently, the ensemble average of all duration-normalized pulses reveals the stereotypical shape of a pulse.

Inflection point. In order to quantify the synchronicity of induction, the inflection point of the fluorescence I was determined by finding the peak of $\frac{d\bar{I}}{dt}$ with $\bar{I}$ being the 10 min moving average of I. The induction magnitude M, (=fold change of I) is given by the ratio of the average I at a fixed time interval before and after the inflection point (*SI Appendix*, Fig. S14):

$$M = \overline{I(t+\Delta)} \,/\, \overline{I(t-\Delta)}.$$

Statistical tests. We performed rank-sum tests for pairwise comparison of wild-type (GTG) with mutants for *Pmar-venus* and P_R-*mCherry* expression. For elongation rate during pulse and elongation rate upon induction, we performed *t* tests.

Competition Experiment. We used two different marker methods: fluorophores (mCherry and Venus) (Fig. 5*A* and *SI Appendix*, Fig. S13) and resistance (chloramphenicol) (*SI Appendix*, Fig. S12*A*). Each genotype was grown overnight in four replicates for each marker separately (with 0 or 2 mM salicylate). Competitions were set up as head-to-head competitions of two genotypes, with dye-swap controls, where each marker was used for half of the replicates. The optical density (OD) was measured and pairwise cultures [wildtype (GTG): wildtype (GTG), wildtype (GTG):ATG, wildtype (GTG):ATG*, wildtype (GTG):TTG, wildtype (GTG):*ΔmarA*] were mixed in a 1:1 ratio, diluted 1,000× as to permit for 10 generations of growth and incubated at 37°C for 24 h. The 1,000×-fold dilution was repeated for three consecutive days so that the genotypes were in direct competition for a total of 40 generations. In parallel, the cultures were plated on LB agar for CFU quantification of the genotypes. Plates were imaged using a custom-build fluorescence macroscope and fluorescent colonies were counted using ImageJ (46). When using the resistance markers, cells were plated on both LB agar and LB agar with chloramphenicol. Selection coefficients were calculated as the slope of the linear model fit to the log ratio of genotypes (using natural log) over generations of competitive growth. We corrected for fitness costs of the markers by subtracting the baseline selection coefficients of control competitions where the same genetic background was competed against itself with different markers. We then used a generalized linear model with genotype as fixed factor and experiment as random factor and a post hoc test with user-defined contrasts to statistically test for significant selection between genotype pairs and estimate the selection coefficient. These analyses were done using the statistical software *R* and the packages *multcomp* and *nlme*.

Population Measurements of Growth Rate and Lag Phase. Population growth rate and growth lag were measured by 1:1,000 dilution of overnight cultures into 0.3 mL of fresh media in Honeycomb plates and measuring OD during exponential regrowth at 37°C with vigorous shaking every 4 min using the Bioscreen C plate reader (OY Growth Curves, Helsinki, Finland; Ref. FP-1100-C) system for at least 6 h. Exponential growth rates were estimated as doublings per hour using the slope of the linear model fit to the plot of log-transformed OD over time in hours during the exponential growth phase. Accuracy of exponential growth was assessed using R^2 values of the fit, which was >0.94 for all growth curves (average 0.98). Lag phase was estimated as the time to restart exponential growth of OD, for which we used the time at which OD reached above the detection threshold of OD = 0.004. Time delay of mutant strains over the wildtype strain was calculated by subtracting the average lag time of the wildtype from the lag times of the mutants. To allow for estimation of lag phase, growth curves were started with equal ODs. ODs of overnight cultures were normalized using OD immediately before inoculation into prepared Honeycomb plates. Without this correction, start growth time is dependent on starting cell density (49). We used ANOVA to test for significant differences in growth parameters between strains and four different growth conditions, where "LB, fresh" refers to regrowth in LB following dilution of a 16 h overnight culture in LB (with and without OD adjustment), "LB aged" refers to regrowth in LB following dilution of a 48 h culture in LB and "M9 glycerol aged" refers to regrowth in M9 minimal medium with 0.2% glycerol as the sole carbon source following dilution of a 48 h culture in that medium. We also quantified CFU per OD for 16 h LB overnight cultures to compare the viability of the 16 h lag estimation data in LB media (Fig. 5*B*) and we did not observe any significant difference in viability across strains (*SI Appendix*, Fig. S15).

Data, Materials, and Software Availability. Raw data have been deposited in IST DataRep and are publicly available at 10.15479/AT:ISTA:19294 (50). All other data are included in the manuscript and/or supporting information.

ACKNOWLEDGMENTS. K.J. thanks B. Wu, I. Tomanek, K. Tomasek for detailed discussions on the manuscript, all other members from the Guet laboratory for valuable feedback, R. Chait, & Imaging and Optics Facility, Institute of Science and Technology Austria for helping with microscopy, Dr. Sudha Rao and Dr. Raja Mugasimangalam, Genotypic Technology India for allowing time off to address the revisions. K.J. acknowledges Institute of Science and Technology fellowship IC1006FELL02, R.H. was supported in part by Chan Zuckerberg Initiative and Donor Advised-Fund grant 2020-225401 (10.37921/120055ratwvi), O.O.B. acknowledges Fonds Zur Förderung der Wissenschaftlichen Forschung (FWF) Grant ESP253-B, R.R. acknowledges FWF Grant 10.55776/ESP219, C.C.G. acknowledges FWF I5127-B.

1. A. L. Koch, The protein burden of lac operon products. *J. Mol. Evol.* **19**, 455–462 (1983).
2. D. M. Stoebel, A. M. Dean, D. E. Dykhuizen, The cost of expression of *Escherichia coli* lac operon proteins is in the process, not in the products. *Genetics* **178**, 1653–1660 (2008).
3. J. M. L. Ho, C. A. Miller, S. E. Parks, J. R. Mattia, M. R. Bennett, A suppressor tRNA-mediated feedforward loop eliminates leaky gene expression in bacteria. *Nucleic Acids Res.* **49**, e25 (2021).
4. O. Patange *et al.*, *Escherichia coli* can survive stress by noisy growth modulation. *Nat. Commun.* **9**, 5333 (2018).
5. E. Nadezhdin, N. Murphy, N. Dalchau, A. Phillips, J. C. W. Locke, Stochastic pulsing of gene expression enables the generation of spatial patterns in Bacillus subtilis biofilms. *Nat. Commun.* **11**, 950 (2020).
6. J. M. Kim, M. Garcia-Alcala, E. Balleza, P. Cluzel, Stochastic transcriptional pulses orchestrate flagellar biosynthesis in *Escherichia coli*. *Sci. Adv.* **6**, eaax0947 (2020).
7. N. M. V. Sampaio, C. M. Blassick, V. Andreani, J.-B. Lugagne, M. J. Dunlop, Dynamic gene expression and growth underlie cell-to-cell heterogeneity in *Escherichia coli* stress response. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2115032119 (2022).
8. E. Kussell, S. Leibler, Phenotypic diversity, population growth, and information in fluctuating environments. *Science* **309**, 2075–2078 (2005).
9. A. M. George, S. B. Levy, Amplifiable resistance to tetracycline, chloramphenicol, and other antibiotics in *Escherichia coli*: Involvement of a non-plasmid-determined efflux of tetracycline. *J. Bacteriol.* **155**, 531–540 (1983).
10. A. S. Seoane, S. B. Levy, Characterization of MarR, the repressor of the multiple antibiotic resistance (mar) operon in *Escherichia coli*. *J. Bacteriol.* **177**, 3414–3419 (1995).
11. J. Garcia-Bernardo, M. J. Dunlop, Tunable stochastic pulsing in the *Escherichia coli* multiple antibiotic resistance network from interlinked positive and negative feedback loops. *PLoS Comput. Biol.* **9**, e1003229 (2013).
12. M. K. Prajapat, K. Jain, S. Saini, Control of MarRAB operon in *Escherichia coli* via autoactivation and autorepression. *Biophys. J.* **109**, 1497–1508 (2015).
13. S. P. Cohen, S. B. Levy, J. Foulds, J. L. Rosner, Salicylate induction of antibiotic resistance in Escherichia coli: Activation of the mar operon and a mar-independent pathway. *J. Bacteriol.* **175**, 7856–7862 (1993).
14. T. M. Barbosa, S. B. Levy, Differential expression of over 60 chromosomal genes in *Escherichia coli* by constitutive expression of MarA. *J. Bacteriol.* **182**, 3467–3474 (2000).
15. P. J. Pomposiello, M. H. J. Bennik, B. Demple, Genome-Wide Transcriptional profiling of the *Escherichia coli* responses to superoxide stress and sodium salicylate. *J. Bacteriol.* **183**, 3890–3902 (2001).
16. R. G. Martin, J. L. Rosner, Genomics of the marA/soxS/rob regulon of *Escherichia coli*: Identification of directly activated promoters by application of molecular genetics and informatics to microarray data. *Mol. Microbiol.* **44**, 1611–1624 (2002).
17. P. Casaz *et al.*, MarA, SoxS and Rob function as virulence factors in an *Escherichia coli* murine model of ascending pyelonephritis. *Microbiology* **152**, 3643–3650 (2006).
18. P. Sharma *et al.*, The multiple antibiotic resistance operon of enteric bacteria controls DNA repair and outer membrane integrity. *Nat. Commun.* **8**, 1444 (2017).
19. L. M. Chubiz, The mar, sox, and rob systems. *EcoSal Plus* **11**, eesp00102022 (2023), 10.1128/ecosalplus.esp-0010-2022.
20. I. El Meouche, Y. Siu, M. J. Dunlop, Stochastic expression of a multiple antibiotic resistance activator confers transient resistance in single cells. *Sci. Rep.* **6**, 19538 (2016).
21. R. G. Martin, J. L. Rosner, Transcriptional and translational regulation of the marRAB multiple antibiotic resistance operon in *Escherichia coli*. *Mol. Microbiol.* **53**, 183–191 (2004).
22. F. Belinky, I. B. Rogozin, E. V. Koonin, Selection on start codons in prokaryotes and potential compensatory nucleotide substitutions. *Sci. Rep.* **7**, 12422 (2017).
23. T. Bergmiller *et al.*, Biased partitioning of the multidrug efflux pump AcrAB-TolC underlies long-lived phenotypic heterogeneity. *Science* **356**, 311–315 (2017).
24. C. C. Guet, L. Bruneaux, P. Oikonomou, M. Aldana, P. Cluzel, Monitoring lineages of growing and dividing bacteria reveals an inducible memory of mar operon expression. *Front. Microbiol.* **14**, 1049255 (2023).

Downloaded from https://www.pnas.org by IST AUSTRIA - LIBRARY; INSTITUTE OF SCIENCE & TECHNOLOGY on April 7, 2025 from IP address 81.223.14.217.

25. R. G. Martin, E. S. Bartlett, J. L. Rosner, M. E. Wall, Activation of the *Escherichia coli* marA/soxS/rob regulon in response to transcriptional activator concentration. *J. Mol. Biol.* **380**, 278–284 (2008).
26. K. L. Griffith, I. M. Shah, R. E. Wolf, Proteolytic degradation of Escherichia coli transcription activators SoxS and MarA as the mechanism for reversing the induction of the superoxide (SoxRS) and multiple antibiotic resistance (Mar) regulons. *Mol. Microbiol.* **51**, 1801–1816 (2004).
27. G. P. Price, A. C. St John, Purification and analysis of expression of the stationary phase-inducible Slp lipoprotein in Escherichia coli: Role of the Mar system. *FEMS Microbiol. Lett.* **193**, 51–56 (2000).
28. K. J. Allen, D. Lepp, R. C. McKellar, M. W. Griffiths, Targeted microarray analysis of stationary phase Escherichia coli O157: H7 subjected to disparate nutrient conditions: Transcriptome analysis of starved E. coli O157: H7. *J. Appl. Microbiol.* **109**, 2118–2127 (2010).
29. O. Shoval *et al.*, Evolutionary trade-offs, pareto optimality, and the geometry of phenotype space. *Science* **336**, 1157–1160 (2012).
30. R. G. Martin, J. L. Rosner, Binding of purified multiple antibiotic-resistance repressor protein (MarR) to mar operator sequences. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 5456–5460 (1995).
31. L. M. Chubiz, C. V. Rao, Aromatic acid metabolites of *Escherichia coli* K-12 can induce the marRAB operon. *J. Bacteriol.* **192**, 4786–4789 (2010).
32. J. L. Rosner, J. L. Slonczewski, Dual regulation of inaA by the multiple antibiotic resistance (mar) and superoxide (soxRS) stress response systems of Escherichia coli. *J. Bacteriol.* **176**, 6262–6269 (1994).
33. M. A. Savageau, Escherichia coli habitats, cell Types, and molecular mechanisms of gene control. *Am. Nat.* **122**, 732–744 (1983).
34. J. F. Brooks *et al.*, The microbiota coordinates diurnal rhythms in innate immunity with the circadian clock. *Cell* **184**, 4154–4167.e12 (2021).
35. A. M. George, S. B. Levy, Gene in the major cotransduction gap of the Escherichia coli K-12 linkage map required for the expression of chromosomal resistance to tetracycline and other antibiotics. *J. Bacteriol.* **155**, 541–548 (1983).
36. J. L. Rosner, Nonheritable resistance to chloramphenicol and other antibiotics induced by salicylates and other chemotactic repellents in Escherichia coli K-12. *Proc. Natl. Acad. Sci. U.S.A.* **82**, 8771–8774 (1985).
37. N. A. O'Leary *et al.*, Reference sequence (RefSeq) database at NCBI: Current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* **44**, D733–D745 (2016).
38. A. Perrin, E. P. C. Rocha, PanACoTA: A modular tool for massive microbial comparative genomics. *NAR Genom. Bioinform.* **3**, lqaa106 (2021).
39. UniProt Consortium, UniProt: The universal protein knowledgebase in 2023. *Nucleic Acids Res.* **51**, D523–D531 (2023).
40. J. Huerta-Cepas *et al.*, eggNOG 5.0: A hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res.* **47**, D309–D314 (2019).
41. R. C. Edgar, MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
42. L.-T. Nguyen, H. A. Schmidt, A. von Haeseler, B. Q. Minh, IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015).
43. I. Letunic, P. Bork, Interactive Tree Of Life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **49**, W293–W296 (2021).
44. G. E. Crooks, G. Hon, J.-M. Chandonia, S. E. Brenner, WebLogo: A sequence logo generator. *Genome Res.* **14**, 1188–1190 (2004).
45. J. Näsvall, Direct and Inverted Repeat stimulated excision (DIRex): Simple, single-step, and scar-free mutagenesis of bacterial genes. *PLoS One* **12**, e0184126 (2017).
46. R. Chait, J. Ruess, T. Bergmiller, G. Tkačik, C. C. Guet, Shaping bacterial population behavior through computer-interfaced control of individual cells. *Nat. Commun.* **8**, 1535 (2017).
47. C. Stringer, T. Wang, M. Michaelos, M. Pachitariu, Cellpose: A generalist algorithm for cellular segmentation. *Nat. Methods* **18**, 100–106 (2021).
48. L. Potvin-Trottier, N. D. Lord, G. Vinnicombe, J. Paulsson, Synchronous long-term oscillations in a synthetic gene circuit. *Nature* **538**, 514–517 (2016).
49. R. Hazan, Y.-A. Que, D. Maura, L. G. Rahme, A method for high throughput determination of viable bacteria cell counts in 96-well plates. *BMC Microbiol.* **12**, 259 (2012).
50. K. Jain *et al.*, Data from "Pulsatile basal gene expression as a fitness determinant in bacteria." ISTA. https://doi.org/10.15479/AT:ISTA:19294. Deposited 4 March 2025.

Downloaded from https://www.pnas.org by IST AUSTRIA - LIBRARY; INSTITUTE OF SCIENCE & TECHNOLOGY on April 7, 2025 from IP address 81.223.14.217.