RESEARCH ARTICLE

Out-of-balance growth enables cost-free synthesis of the flagellum and other proteins in a single bacterium

Mayra Patricia Garcia Alcala[1], Josiah C Kratz[2], Philippe Cluzel[1]*

[1]Department of Molecular and Cellular Biology, Harvard John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, United States; [2]Computational Biology Department, School of Computer Science, Carnegie Mellon University, Pittsburgh, United States

eLife Assessment

This **important** study addresses a discrepancy between population-level growth laws and single-cell correlations. It shows, for flagellar and synthetic genes in *Escherichia coli*, that while gene expression of certain genes reduces population-average growth, expression levels positively correlate with growth at the single-cell level. The measurements are **convincing**, and the proposed mechanism—inheritance of growth factors such as ribosomes during asymmetric division—can explain this observation.

***For correspondence:** cluzel@mcb.harvard.edu

Competing interest: The authors declare that no competing interests exist.

Funding: See page 12

Preprint posted
21 January 2026
Sent for Review
02 February 2026
Reviewed preprint v1
01 April 2026
Reviewed preprint v2
05 August 2026
Reviewed preprint v3
08 October 2026
Version of Record
28 October 2026

Reviewing Editor: Ariel Amir, Weizmann Institute of Science, Israel

© Copyright Garcia Alcala *et al*. This article is distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use and redistribution provided that the original author and source are credited.

Abstract The cost of expressing unnecessary genes on growth has been explained at the population level through balanced growth assumptions. However, due to the high degree of stochasticity in growth and gene regulation, it remains unclear how classical growth laws derived from bulk measurements manifest at the level of individual cells. For example, flagellar gene expression, one of the most energy-intensive processes in *Escherichia coli*, is typically associated with reduced growth rates in bulk measurements. By contrast, we find at the single-cell level that a strikingly opposite behavior coexists: flagellar gene activity is associated with an increase in growth rate. We show that this apparent contradiction, reminiscent of Simpson's paradox, results from examining in each cell instantaneous growth and gene expression without indiscriminately aggregating the single-cell data like in bulk measurements. We attribute this unrecognized behavior to out-of-balance growth boosts that temporarily offset the burden of flagellum production, likely driven by a surplus of growth factors inherited across cell division. Furthermore, we show that this non-intuitive relationship between growth and gene expression at the single-cell level is general, as we observe similar effects with synthetic systems driving the expression of fluorescent proteins downstream of constitutive promoters. Finally, we use a computational model to show that the inheritance of parental growth factors can explain the apparent contradiction and constitutes a general mechanism for mitigating the short-term cost of nonessential and biosynthetically demanding genes in the single cell.

Introduction

The balance between the cost associated with the expression of genes and the demands for intracellular resources driven by growth is a tightly regulated process in bacteria, reflecting an evolutionary pressure to optimize fitness (***Dekel and Alon, 2005***; ***Dong et al., 1995***; ***Scott and Hwa, 2011***; ***Stoebel et al., 2008***). Recent studies have built on this principle to derive powerful, predictive growth laws (***Basan et al., 2015***; ***Scott et al., 2010***; ***Scott and Hwa, 2011***; ***Shachrai et al., 2010***). For example, at

the population level, these laws predict that expressing unnecessary proteins reduces the availability of ribosomes that would otherwise be destined for the synthesis of other proteins, including those involved in ribosomal function and other cellular processes (***Scott et al., 2010***). Consequently, the synthesis of unnecessary proteins is accompanied by a growth burden that increases with protein expression levels (***Dekel and Alon, 2005***; ***Dong et al., 1995***; ***Shachrai et al., 2010***). These laws have been derived under the assumption of balanced growth: a state where all cellular components, such as proteins, RNAs, DNA, and lipids, are synthesized at rates that are proportional to the constant growth rate of the cells measured at the population level (***Klumpp and Hwa, 2008***; ***Klumpp et al., 2009***). Balanced growth also ensures that cell growth has reached a steady-state and does not run into resource bottlenecks that could impair its physiology (***Dourado and Lercher, 2020***; ***Jun et al., 2018***). While balanced growth is a practical assumption for predicting bacterial growth at the population level, it does not take into account out-of-balance processes (***Bollenbach et al., 2009***; ***Korem Kohanim et al., 2018***; ***Kratz and Banerjee, 2023***), such as sudden and stochastic molecular events that inherently accompany growth and gene expression observed in single cells (***Biswas et al., 2024***; ***Goldman et al., 2023***; ***Jain et al., 2025***; ***Jones and Uphoff, 2021***; ***Kim et al., 2020***; ***Park et al., 2018***; ***Patange et al., 2018***; ***Sampaio et al., 2022***), even when environmental conditions are well-controlled and steady. Therefore, extracting the laws that govern the interplay between stochastic growth variations and protein synthesis is crucial for predicting and understanding the behavior and constraints of the single cell.

In recent years, studies in *Escherichia coli* have reported that flagellar genes, responsible for the synthesis of one of the most energy-intensive bacterial organelles, exhibit large stochastic expression pulses (***Kim et al., 2020***). The assembly of this complex organelle requires more than 30 different proteins that mostly constitute the basal body of a rotary motor, as well as a 20,000 subunit-long filament (***Chevance and Hughes, 2008***; ***Sassi et al., 2020***; ***Figure 1—figure supplement 1A***). During flagella synthesis, the cell directs a sizable portion of protein synthesis resources to this process (***Erickson et al., 2017***; ***Mori et al., 2021***; ***Schavemaker and Lynch, 2022***). While it has been demonstrated from bulk measurements that during balanced growth, the expression of this locomotion apparatus is regulated to compensate on average for changes in cell size across distinct growth conditions (***Honda et al., 2022***; ***Ni et al., 2020***; ***Ni et al., 2017***), far less is known about out-of-balance processes and their effect on gene expression, such as stochastic fluctuations of growth observed at the single-cell level. In view of the large protein synthesis demand that accompanies the stochastic transcriptional expression of flagellar genes, here we investigate the interplay between such gene activity and growth fluctuations within a single bacterium. Subsequently, we extend our analysis to that of synthetic expression systems and develop a computational model to validate the generality of our observations.

Collectively, production of flagellar proteins represents a sizable proteome investment that may consume resources otherwise available for growth, presenting a potential resource allocation tradeoff which the cell must navigate (***Ni et al., 2020***; ***Yi and Dean, 2016***). The activity of flagellar genes follows a well-characterized transcriptional cascade (***Chevance and Hughes, 2008***; ***Fitzgerald et al., 2014***). Briefly, it begins with Class-1 genes that encode the master regulator FlhDC, which drives the transcriptional activation of Class-2 genes. These genes are involved in the synthesis of the flagellar basal body, the hook, and the alternative sigma factor, FliA. This sigma factor, in turn, activates Class-3 genes that encode the filament and signaling chemotaxis proteins (***Figure 1—figure supplement 1A***). Recently, the transcriptional activity of Class-2 and Class-3 flagellar genes has been found to be well-delimited in time with alternating and distinct *on* and *off* states, each spanning several cellular divisions (***Kim et al., 2020***). Operationally, this temporal pattern makes the system experimentally accessible for dissecting the relationship between growth fluctuations and gene activity in living individual cells.

Results

We tracked single cells as they grew and divided inside a microfluidic device called a 'Mother Machine' (***Moffitt et al., 2012***; ***Wang et al., 2010***), which allows bacteria to grow exponentially under controlled conditions for extended periods of time (***Figure 1—figure supplement 1B***). This device confines cells within narrow channels, with one end exposed to a continuous flow of fresh medium while the other end is sealed. The temporal activity of flagellar promoters and the growth of individual

cells are simultaneously monitored using fluorescence and phase-contrast microscopy ('Materials and methods'). This experimental platform enables us to capture time-lapse images and extract various physiological parameters such as cell size, elongation rate, and fluorescence signals from reporter proteins with 5-minute resolution over ~30 generations (***Figure 1—figure supplement 1C***). One key feature of this approach is its ability to isolate intrinsic fluctuations within the cellular system from those typically caused by uncontrolled environmental variations.

As previously observed by some of us, in the commonly used wild-type *E. coli* MG1655 strain, Class-2 and Class-3 promoters exhibit sparse stochastic pulses, whereas Class-1 displays low but sustained activity that resembles that of a constitutive promoter. The question then becomes how the growth rate varies during distinct periods of active or inactive flagellar gene activity. Here we use the elongation rate of cell size as monitored in the Mother Machine channels, as a measure of growth rate (***Figure 1—figure supplement 2***). All experiments were performed in an MG1655 background harboring a *motA* mutation that disables swimming motility but does not impair flagellar assembly (***Kojima and Blair, 2001***; ***Morimoto et al., 2010***), thereby isolating the regulatory and growth-related dynamics of the flagellar system (see 'Materials and methods', ***Figure 1—figure supplement 3***). We monitored the activity of the Class-2 promoter, *fliFp*, using the fluorescent reporter SCFP3A as it also encapsulates the dynamics of Class-3, as previously described (***Kim et al., 2020***). Practically, to capture the relationship between instantaneous variations in elongation rate and those of Class-2 activity, we extracted these two quantities from time series at each time point using a time window Δt=5-minute time window (***Figure 1***, left). These pairs were then sorted and grouped into bins based on their activity levels (***Figure 1—figure supplement 4***). In addition to the strain with wild-type flagellar expression, we tested three distinct strains whose master regulator, FlhDC, is driven by unregulated synthetic promoters (P1, P2, P3; ***Davis et al., 2011***; ***Kim et al., 2020***) with increasing strength, resulting in progressively higher transcriptional activity of downstream flagellar genes.

The well-known population result is recovered when we fit the whole bulk of the instantaneous single-cell data across the four different strains (***Figure 1***, right, dashed line, ***Figure 1—figure supplement 5***), indicating that as the population of cells increases expression of unnecessary proteins, their growth rate decreases. Surprisingly, when investigating growth relationships for individual cells within each strain, we found that bins with higher instantaneous transcriptional activity were also associated with higher elongation rates (***Figure 1***, right, ***Figure 1—figure supplement 5***). This result is unexpected under the canonical growth laws and directly contrasts with population-level observations, for which elevated flagellar gene activity correlates with reduced growth rates (***Figure 1—figure supplement 6***, ***Honda et al., 2022***; ***Ni et al., 2020***; ***Ni et al., 2017***; see ***Figure 1—figure supplement 7*** for an alternative growth rate proxy). The coexistence of these opposing trends, positive within strains but negative across strains, suggests that a phenomenon such as the Simpson's paradox is at play.

Intuitively, the Simpson's paradox, a classic statistical phenomenon, can arise when distinct groups (i.e., individual strain) occupy different regions of a shared X–Y parameter plane (here, expression-growth). For example, each individual group might exhibit a positive internal trend between X and Y, but their respective data clusters are down shifted relative to one another. However, when we analyze the data across groups, this internal covariation within each group is masked. As a result, the aggregate analysis reflects only the relationship between these shifted group averages, completely reversing the underlying trend observed within the individual groups.

In our study, although single-cell and population analyses rely on instantaneous single-cell measurements, they differ fundamentally in how the data are aggregated, a distinction from which the observed paradox arises. In the population-level analysis (***Figure 1***, dashed line), each point represents the mean instantaneous growth rate and mean promoter activity averaged across each distinct strain. This aggregation collapses co-fluctuations of growth and activity within cells and compares averages across genetic backgrounds, thus reflecting differences across strains governed by balanced growth constraints. In the binned analysis, each data point reflects paired fluctuations of growth and gene expression measured within the same cells, preserving their natural co-variation at a given time point. Hence, the binned analysis reveals dynamic coupling between growth and expression within single cells, whereas the aggregated data captures only the average trade-off observed across strains.

We performed the same analysis using the Class-1 promoter activity. In this case, the within-strain relationship between instantaneous promoter activity and elongation rate remained positive, but the trend was noticeably weaker than Class-2 (***Figure 1—figure supplement 8***). We interpret this

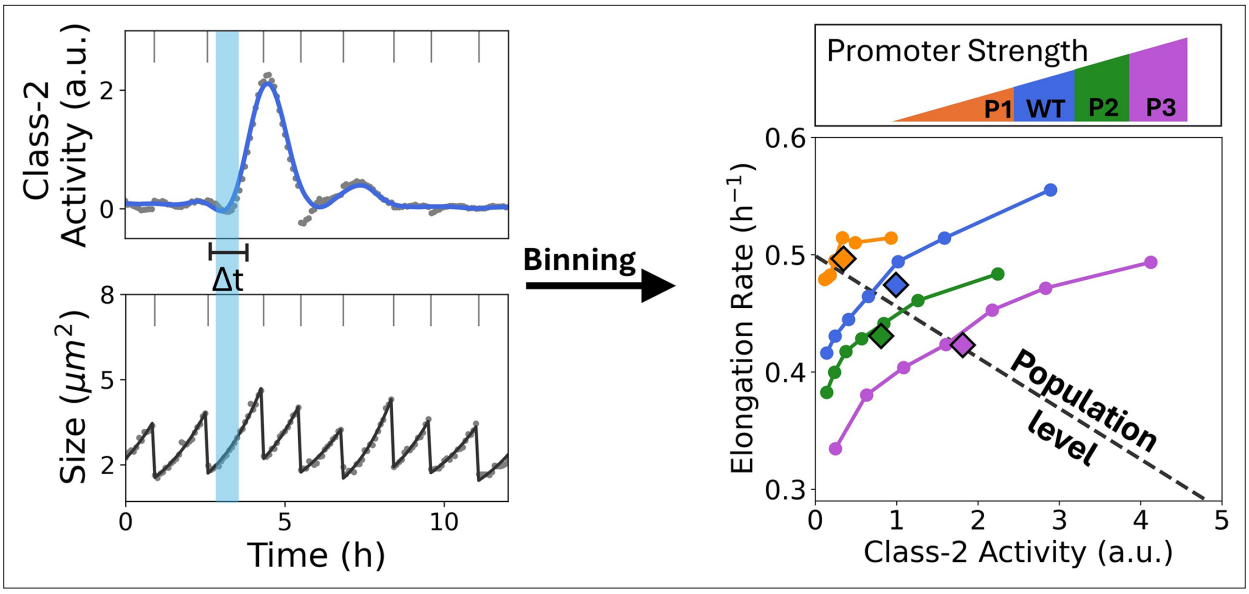


Figure 1. Relationship between instantaneous growth rate and flagellar gene activity in single cells versus population. Left: time traces of Class-2 activity and Size (S) measured from the same mother cell across divisions (vertical ticks). We defined Class-2 activity as the time derivative of total fluorescence normalized by size, and the elongation rate as $ER\left(t\right)=\left(1/S\right)\left(\Delta St/\Delta t\right)$. Direct measurements of activity and size are shown as dots, whereas lines show the exponential fit for the size and the smoothed activity (see 'Methods and materials'). First, we used Class-2 activity to sort data points with Δt=5 minutes and then bin the pair activity-elongation rate (see ***Figure 1—figure supplement 4***). Finally, we average the data in each bin and display the result in the right panel. Because the averaged bins are built from pairs of elongation rate values using a 5-minute time window, this binning allows us to determine the relationship between instantaneous flagellar gene activity and growth rate in single cells. Right: instantaneous Class-2 activity versus elongation rate for different strains with varying mean levels of flagellar expression. In such strains, the native Class-1 promoter was replaced by synthetic promoters with different strengths, P1, P2, and P3 (color coded in top panel), P3 being threefold stronger than P1 (***Davis et al., 2011***). The dashed line shows the linear fit to the population average of the pair Class-2 activity-elongation rate across strains, which illustrates the well-established relationship with a negative slope (cost) between growth rate and increasing mean flagellar gene activity across strains (***Ni et al., 2020***). By contrast, single-cell binned data within each distinct strain shows a positive slope between instantaneous growth rate and flagellar activity, demonstrating that cells with higher flagellar gene activity are also the ones growing faster. Only the data exhibiting Class-2 activity (excluding *off*-states) are considered in this plot; see ***Figure 1—figure supplement 4D*** for all states. The number of time traces is n=69, 107, 96, 67 for the strains with the promoter P1, WT, P2, and P3, respectively, all with a duration of ~47 hours.

The online version of this article includes the following figure supplement(s) for figure 1:

Figure supplement 1. Simultaneous monitoring of growth and pulses of flagellar gene activity in single bacteria.

Figure supplement 2. Time series of Class-2 activity and cell elongation rate in flagellar reporter strains.

Figure supplement 3. Determining motility of the synthetic flagellar series in soft agar plates and microfluidic device.

Figure supplement 4. Elongation rate vs Class-2 promoter activity scatter plot and binning.

Figure supplement 5. Statistical analysis of single-cell relationships between elongation rate and Class-2 activity.

Figure supplement 6. Bulk measurements show a decrease in growth rate when flagellar proteins are overexpressed in *E. coli*.

Figure supplement 7. Instantaneous growth rate and flagellar Class-2 promoter activity.

Figure supplement 8. Class-1 activity and elongation rate correlation.

reduction as a consequence of Class-1 being noisy and that only a smaller fraction of Class-1 activity yields a productive Class-2. This behavior was explained in ***Kim et al., 2020***. Thus, fluctuations in Class-2 activity more directly report changes in global translational capacity than Class-1 and therefore show a stronger coupling to growth.

We reasoned that if this single-cell effect reflects dynamic coupling between growth and Class-2 activity, it should also be observable along individual time series and not only in binned data. To examine this point further, we took advantage of the fact that flagellar genes show pulses of activity, which allow us to clearly track how growth rate changes during significant fluctuations in gene activity. We isolated time series segments of Class-2 activity featuring two consecutive pulses, rescaled their total duration to a unit interval (arbitrary units), and computed the average across a large collection of such pulse-to-pulse excerpts (***Figure 2***, top; 'Materials and methods').

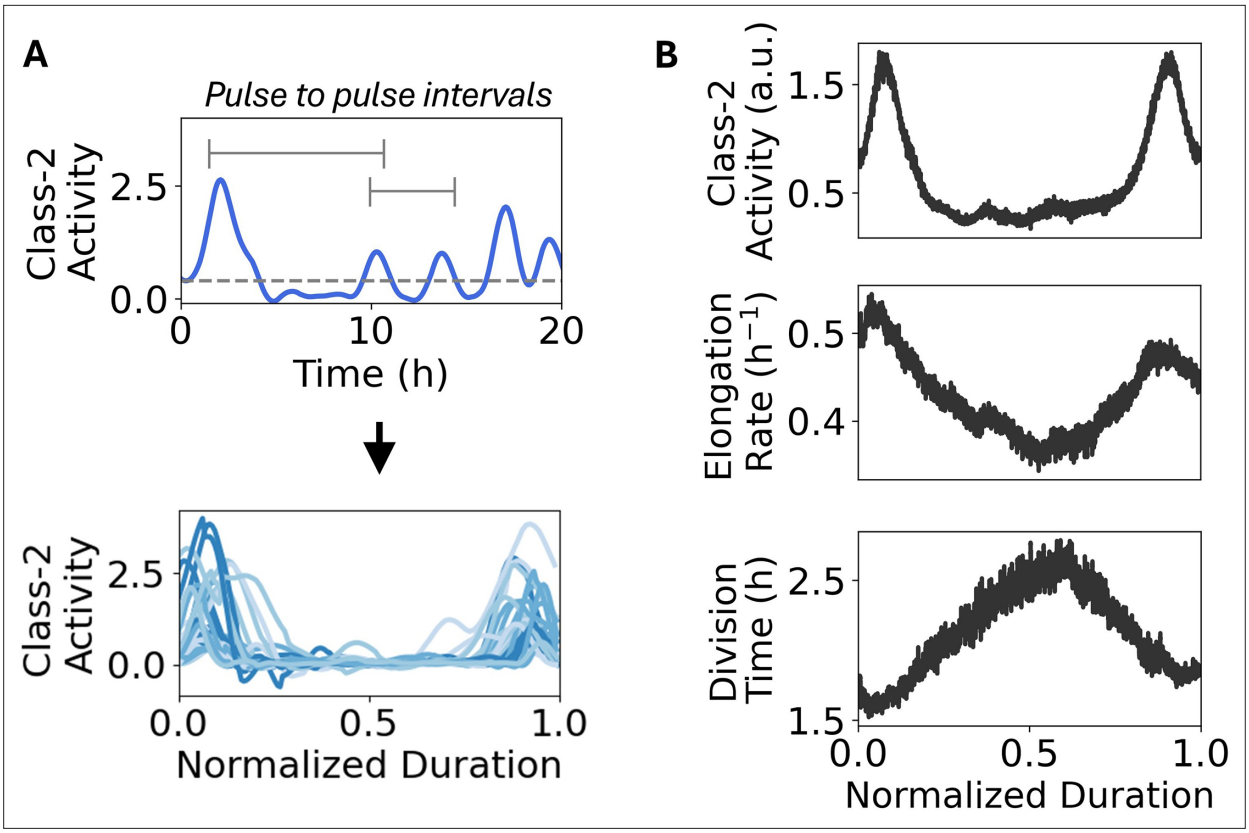


Figure 2. Cells grow faster during pulses. (**A**) The Class-2 activity time series were divided into intervals that consisted of a pulse of activity followed by an *off* state and another pulse. We selected all the intervals with similar durations (within a range of no more than 6 hours). The pulses were defined by the period for which the activity is above an empirical threshold (activity = 50 arbitrary units). Then, the intervals of similar duration were normalized such that they spanned from 0 to 1, in order to average them. (**B**) Top: averaged pulse-to-pulse intervals (real time rescaled to 1) of the Class-2 activity. Middle and bottom: elongation rate and division time associated with the same time intervals as the top panel. The averages were calculated by applying a sliding window mean over the overlapping data consisting of n=103 intervals.

The online version of this article includes the following figure supplement(s) for figure 2:

Figure supplement 1. Class-2 pulse-to-pulse and elongation rate analysis.

Figure supplement 2. Cross-correlation between elongation rate and Class-2 activity in strains of varying flagellar activity.

Figure supplement 3. Class-3 activity and elongation rate correlation.

Figure supplement 4. Cross-correlation analysis of elongation rate and fluorescence for various promoters and strains.

Figure supplement 5. Cross-correlation analysis of elongation rate and promoter activity for various promoters and strains.

Then, we jointly plotted the time series of elongation rate and division time that are directly associated with the pulse-to-pulse excerpts. We found that the aggregated time traces indicate that increases in flagellar gene activity coincide with higher elongation rates, while division time varies in the opposite phase (***Figure 2***, middle and bottom). This relationship remains consistent across pulse-to-pulse intervals of different lengths (see ***Figure 2—figure supplement 1***).

To further examine whether there is a phase delay between elongation rate and Class-2 activity, we calculated the cross-correlation function of the full time series (about 20 divisions), to quantify how strongly elongation rate and activity are related when the activity signal is shifted in time (by a lag) relative to elongation rate. Thus, positive delays indicate that elongation rate lags behind promoter activity, whereas negative delays indicate that elongation rate precedes promoter activity. We found that the fluctuations in elongation rate temporally precede those of flagellar gene activity, and that they are positively correlated (***Figure 2—figure supplement 2***, see 'Materials and methods'). This result was unexpected, since one might intuitively assume, based on population measurements, that increased flagellar gene activity would cause a cost that reduces the elongation rate. If that were the

case, we would expect to see a negative correlation where changes in gene activity preceded those of elongation rate, rather than the positive correlation and temporal order we observed.

We repeated the correlation analysis using a reporter for the Class-3 promoter, *fliCp*, which drives expression of the energetically costly flagellin protein. We again observed a positive correlation (***Figure 2—figure supplement 3***), with a slightly larger shift between elongation rate and activity for Class-3 compared to Class-2, consistent with Class-2 activity preceding and driving Class-3.

We also analyzed the three strains whose synthetic constitutive promoters (P1, P2, P3) control the expression of the flagellum master regulator and found similar temporal delay (***Figure 2—figure supplement 2***). Since the constitutive promoters P1, P2, and P3 are synthetic, they do not respond to any input specific to flagellar regulation; rather they behave more like unregulated, constitutive promoters (***Figure 2—figure supplements 4 and 5***). This result suggests that the observed temporal shift in flagellar gene activity may be driven by global intracellular factors rather than flagellum-specific ones. Indeed, the signature observed in the correlation between elongation rate and Class-2 activity is consistent with the scenario described by ***Kiviet et al., 2014***, in which the dominant noise source fluctuations are global cellular processes.

Recent work has observed heterogeneity in ribosome density across populations of *E. coli* cells (***Chai et al., 2014***; ***Chao et al., 2024***; ***Papagiannakis et al., 2025***; ***Pavlou et al., 2025***). While ***Pavlou et al., 2025*** found no immediate association between ribosome abundance and unequal division (***Pavlou et al., 2025***), several super-resolution studies suggest that ribosome heterogeneity arises from asymmetric partitioning during cell division (***Chai et al., 2014***; ***Chao et al., 2024***; ***van Heerden et al., 2024***). They demonstrated such asymmetry stems from large clusters of ribosomes that are asymmetrically distributed at cell division, leading to differences in growth after division. Such a pronounced effect would not be possible if each ribosome was independently partitioned according to a binomial distribution, which, due to their large number, would result in a highly symmetrical distribution between daughter cells (***Biswas et al., 2024***; ***Tanouchi et al., 2015***). With these results in mind and given that ribosome synthesis is an upstream process relative to flagellar gene expression, we proceeded to analyze the symmetry between cells following cell division as well as the heritability of flagellar gene activity and growth rate fluctuations from one generation to the next.

First, we asked whether after each division the two new daughter cells differ systematically in growth and flagellar gene activity. We compared the daughter with the new pole and the daughter that inherited the old pole (cell adjacent to the closed end of the channel inherits the old pole) (***Figure 3A***). We found a strong bias: the daughter with new pole has a greater elongation rate than the daughter with the old pole. This bias is even stronger when the daughter with the new pole also has a greater promoter activity than its sister (***Figure 3—figure supplement 1***).

To quantify this bias, we computed, at each division, the ratio of elongation rate and Class-1 and Class-2 promoter activity between new-pole and old-pole daughters. All ratios are clearly above 1 (***Figure 3A***), showing that new-pole daughters tend to grow faster and exhibit higher flagellar gene activity than old-pole daughters. This observed bias is consistent with previous work showing that old-pole regions lose gene expression activity (***Shi et al., 2020***), and that aging at the old pole enhances phenotypic heterogeneity (***Proenca et al., 2024***). More importantly, it agrees with observations that ribosomes are denser in daughters with new poles, where higher ribosome levels are associated with higher elongation rates (***Chao et al., 2024***). Together, these observations support the hypothesis that differences in growth rate between sister cells are governed by differences in growth factor distribution at division.

We then focused exclusively on cells that exhibited changes in flagellar gene activity across division (from mother to daughter) and measured the corresponding changes in elongation rate. We first selected the cells that transitioned from *off* to *on* activity across division (***Figure 3B***, upper right) and compared their elongation rates (***Figure 3B***, yellow). Our results indicate that the increase in activity across divisions is associated with an increase in the elongation rate that is greater than for cells with no transition (*off–off*, dotted line). By contrast, cells showing a transition from *on* to *off* activity (***Figure 3B***, green) fall below the dotted line.

Together, these observations show that cell division systematically generates asymmetry in both growth and flagellar gene activity and that transitions in flagellar activity across division are accompanied by concordant changes in growth rate, indicating that both traits are modulated by short-timescale global fluctuations.

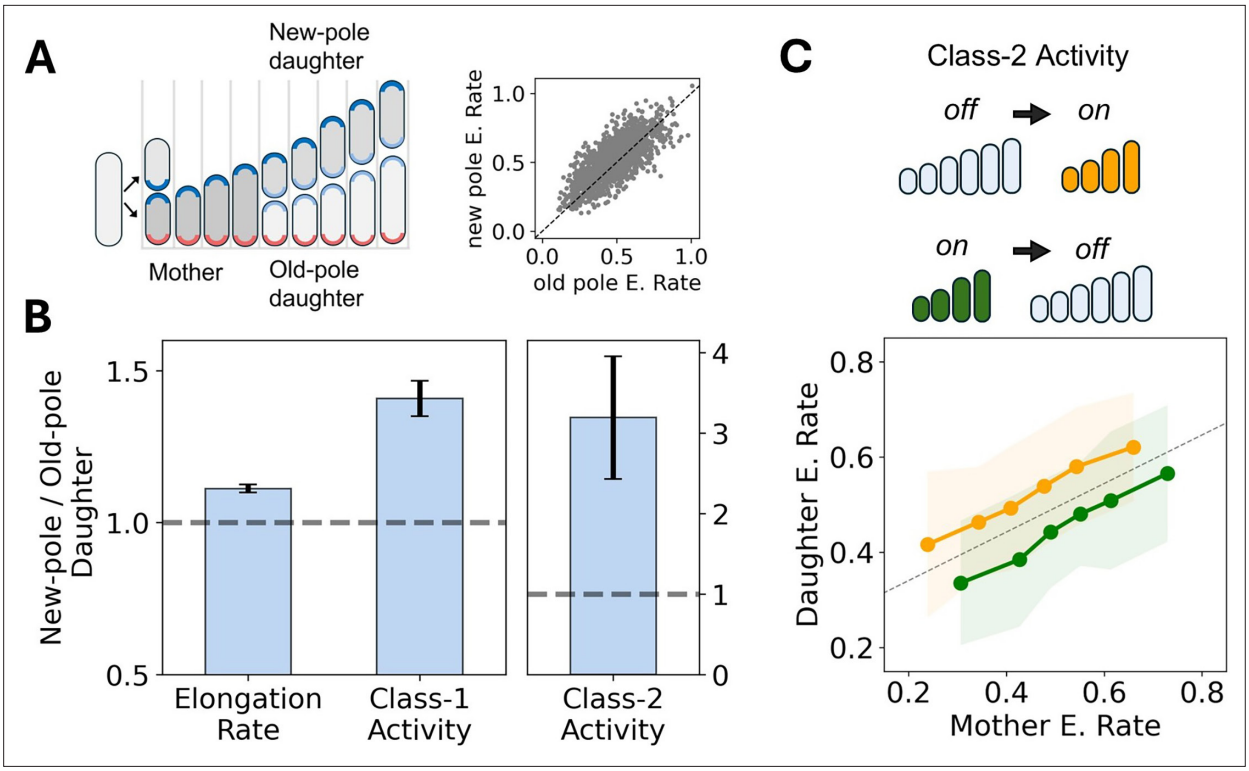


Figure 3. Short-timescale shifts in Class-2 activity correlate with immediate changes in growth. (**A**) Left: schematic of a dividing mother cell producing two daughters, one inheriting the new pole (red pole) and the other retaining the old pole (blue pole). Scatter plot of elongation rate in sister cells (new-pole/old-pole daughters). (**B**) Bars show the mean ratio of new-pole to old-pole daughters for elongation rate, Class-1 activity, and Class-2 activity. The gray dotted line indicates a ratio of 1, corresponding to equal values between new- and old-pole daughters and is also the mean ratio obtained from a null model based on label-shuffling permutations between the daughters. Error bars indicate 95% confidence intervals. N=1669 pairs of the WT flagellar expression strain. (**C**) Elongation rates across two consecutive generations for cells experiencing an abrupt change in Class-2 activity: *off* to *on* (yellow) and *on* to *off* (green). For both cases, we plot the binned ER of the daughter cell (generation g+1) against the ER of the previous generation (g). A positive change in activity (*off* to *on*) is accompanied by an increase in ER across generations compared to the dotted line, reflecting faster cell growth in generation g+1. A decrease in activity (*on* to *off*) is associated with a reduced ER compared to the dotted line in generation g+1 (ER is lower at g+1 than at g). The number of divisions for each subset is ~300. The shaded regions correspond to the standard deviation for each bin, and the dotted line shows the linear fit for the entire dataset, irrespective of activity, representing a null model for ER across divisions.

The online version of this article includes the following figure supplement(s) for figure 3:

Figure supplement 1. Probability of bias advantage between the elongation rate of old and new pole daughters.

Figure supplement 2. MG1655 and MC4100 differ in size-oscillation dynamics but not in short-timescale growth–expression coupling, which is attenuated only at flagellar Class-2 activity.

Next, we asked whether these observations could be strain dependent. The MC4100 *E. coli* strain is known to exhibit long-timescale oscillatory dynamics in cell size, distinct from MG1655 (***Tanouchi et al., 2015***). We found that this slow, ~10-generation oscillation is present in MC4100's cell size but is absent from its elongation rate autocorrelation (***Figure 3—figure supplement 2A and B***). Instantaneous elongation rate remained positively correlated with Class-1 promoter activity in MC4100, comparable in magnitude to MG1655, indicating that the short-timescale growth-expression coupling itself is not strain-specific. However, this correlation was markedly attenuated at the level of Class-2 flagellar activity in MC4100 specifically (***Figure 3—figure supplement 2C***), suggesting that the discrepancy arises in the transmission of growth-coupled fluctuations through the flagellar regulatory cascade rather than in the growth-coupling mechanism itself.

Next, we explored whether the single-cell behavior observed in ***Figure 1***, namely the violation of the population-level growth law, is specific to the flagellum system or if it constitutes a more general phenomenon observable at the single-cell level. To this end, we repeated our experiments using unregulated synthetic expression systems that govern the expression of the yellow fluorescent protein, Venus, by means of synthetic constitutive promoters with various strengths (***Sassi et al.,***

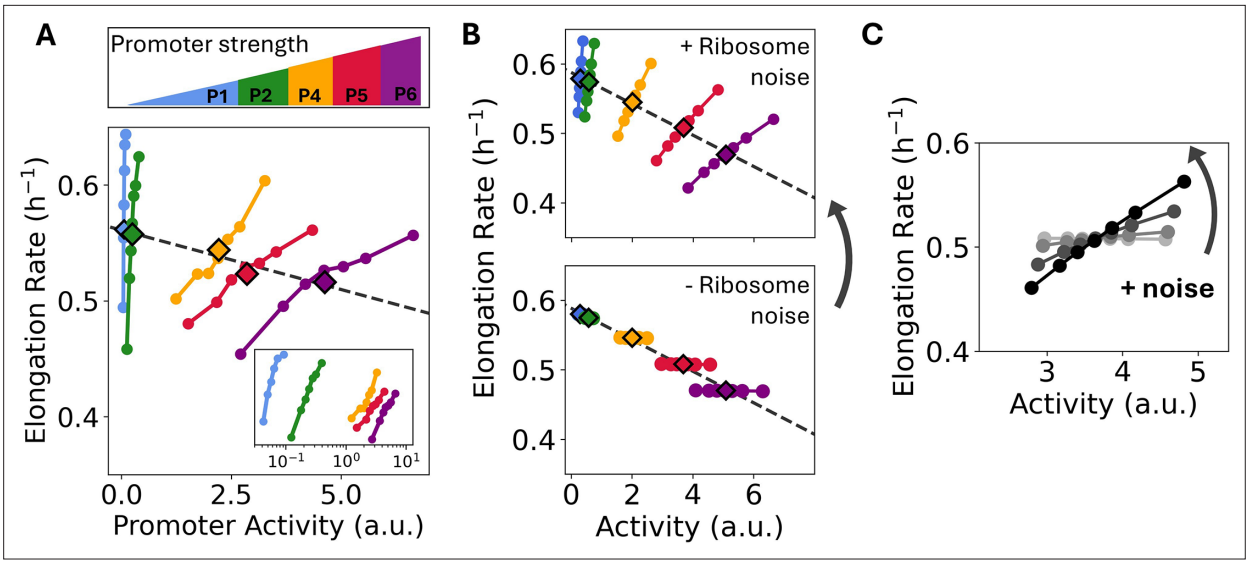


Figure 4. Relationship between instantaneous growth and promoter activity in single cells versus population for different levels of constitutive expression of unnecessary proteins. (**A**) For a set of strains expressing the fluorescent protein Venus, driven by promoters of varying strengths (P1–P6, as indicated at the top), we calculated the mean elongation rate as a function of promoter activity. The diamond markers represent the average for the entire population of each strain, and the black dotted line indicates the linear fit through these population averages. The colored dots show the binned single-cell data for each strain. In each strain, cells exhibit a positive correlation between elongation rate and constitutive promoter activity. The inset shows the same data with the x-axis plotted on a log scale. (**B**) (top) Single-cell simulations for increasing average allocation fraction to unnecessary protein expression ($\overline{f_U}$), capturing both the decreased elongation rate with increasing activity at the population level across conditions (diamonds), as well as the binned data, that shows the increase in elongation rate with activity at the single-cell level within conditions (colored dots). (bottom) Single-cell simulations of elongation rate as a function of activity with identical parameters to (top), except without noise due to unequal partitioning of ribosomes at division ($\sigma_R = 0$), showing that the positive slope between elongation rate and activity seen at the single-cell level (top) is driven by fluctuations in ribosome concentration at the single-cell level. (**C**) Simulations were performed for a fixed fraction of unnecessary proteins, $f_U = 0.073$, and varying noise amplitudes of ribosomes due to unequal partitioning $\sigma_R = (0, 0.02, 0.04, 0.06)$ (light gray to black). For each condition, the resulting scattered data of activity versus elongation rate was binned as in the experiments. All simulations were implemented using rules for proteome allocation and division timing from ***Kratz and Banerjee, 2023***, and other simulation parameters are described in SI Section 1.

The online version of this article includes the following figure supplement(s) for figure 4:

Figure supplement 1. Statistical analysis of activity–elongation rate relationships in constitutive fluorescent reporter strains.

Figure supplement 2. Elongation rate and promoter activity across constitutive expression levels and media.

Figure supplement 3. Growth-constitutive expression coupling is conserved across *E. coli* strains and conditions.

2022). Although our bulk measurements (***Figure 4A***, diamonds, ***Figure 4—figure supplement 1C***) as in ***Basan et al., 2015*** experimentally confirm the standard growth law, where growth rate decreases as unnecessary protein expression increases, binning analyses of instantaneous growth rate measurements revealed a strikingly different behavior at the single-cell level. Using a validated model relating growth-rate reduction to unnecessary-protein burden, based on the ribosome-allocation theory of ***Scott et al., 2010*** and calibrated for GFP-family proteins by ***Bienick et al., 2014***, we estimate that mVenus constitutes approximately 0.6–4.0% of dry cell weight across our promoter series (***Supplementary file 1C***). As with the flagellum system, our analysis shows that an instantaneous increase (decrease) in fluorescent protein synthesis is accompanied by a faster (slower) elongation rate (***Figure 4A***, ***Figure 4—figure supplement 1A and B***; cross-correlation between elongation rate and activity in ***Figure 4—figure supplement 2***). The result in ***Figure 4A*** appears to contradict the growth law, which posits that the increased synthesis of unnecessary proteins, such as the fluorescent proteins in our experiments, depletes ribosomal availability and should therefore reduce growth rate. However, our single-cell analyses reveal an opposite behavior, where transient boosts of growth accompany increases in protein synthesis. Consequently, we hypothesized that the observed behavior likely arises from an out-of-balance growth process over short timescales during which a surplus (or deficit) of beneficial growth factors is redistributed at division from the mother at no cost to the

daughter cell. The process is out-of-balance because the post-division abundance of growth factors (such as ribosomes) reflects unequal inheritance rather than balanced synthesis.

To test whether this coupling extends beyond our own strains and experimental conditions, we reanalyzed publicly available single-cell data from two independent studies of *E. coli* constitutively expressing a fluorescent protein: ***Wang et al., 2010***, in MG1655 and B/r, and ***Tanouchi et al., 2017***; ***Tanouchi et al., 2015***, in MC4100 at 25°C and 37°C. In all cases, we found the same positive, short-timescale correlation between instantaneous elongation rate and expression (***Figure 4—figure supplement 3***) indicating that the out-of-balance growth process we describe is not limited to our strains or growth conditions, but is a broader feature of single-cell *E. coli* physiology.

To evaluate whether the proposed out-of-balance growth mechanism can account for these observations, we performed single-cell simulations using a computational model (***Figure 4B***) that incorporates all the key elements that govern the growth laws as commonly defined from population measurements (simulation details in the Supplementary Materials). Importantly, the model also includes the stochastic redistribution of growth factors such as ribosomes and possibly other factors from the mother to daughter cells at division. We find that our binning procedure shows the same positively correlated relationship between growth rate and synthesis of unnecessary proteins observed experimentally (***Figure 4B***, top). Moreover, the slope of this relationship becomes steeper as the noise in the distribution of growth factors becomes larger (***Figure 4C***). In the absence of noise, the relationship shows no correlation between growth rate and protein synthesis (***Figure 4B***, bottom). In contrast, this correlation exhibits increased steepness with higher noise levels, indicating that stochastic partitioning of growth factors amplifies transient out-of-balance states.

This effect shows that instantaneous growth and expression changes are mainly dominated by inheritance noise rather than by the steady-state burden predicted by the growth law. In our model, single-cell growth rate is positively correlated with protein synthesis, so any mechanism generating molecular growth factors heterogeneity across generations will give rise to Simpson's paradox when comparing single-cell and population-level trends.

Altogether, these simulations illustrate that both behaviors can coexist: the growth burden caused by the synthesis of unnecessary proteins as seen at the population level, and the out-of-balance growth observed at single-cell level, which mediates instantaneous, burden-free synthesis of unnecessary proteins.

Rather than identifying the precise molecular source of heterogeneity, our goal is to examine the consequences of fluctuations of unspecified growth factors for balanced growth models. We therefore propose that these fluctuations underlie an intergenerational mechanism for cost-free ('pre-paid') protein production. We argue that mother cells can pass molecular growth factors to their daughters, either through excess of their production or asymmetric division, thereby giving these daughters an advantage in synthesizing costly proteins such as the flagellum.

Discussion

Kim et al., 2020 found that flagellar genes in *E. coli* are activated in stochastic pulses. Subsequently, some of us proposed that these pulses are generated by means of an inhibitor of the Class-1 transcription factor, called RflP (formerly YdiV), which creates an ultrasensitive switch that acts as a digital filter to remove small-amplitude input temporal fluctuations in Class-1 expression (***Sassi et al., 2020***). While this molecular mechanism demonstrates that pulses can arise without feedback, it does not address how such pulses influence cellular growth. This switch may also contribute to strain-specific differences: we found that in MC4100, the short-timescale coupling between elongation rate and expression, intact at the Class-1 level and in constitutively expressed proteins across strains (***Figure 3—figure supplement 2C***, ***Figure 4—figure supplement 3***), is markedly attenuated at Class-2. Rather than reflecting a general difference in growth-factor partitioning, which we show is comparable to MG1655 at the Class-1 level, this result suggests that the flagellar pulse-generating circuit itself may differ across strains in how, or whether, it exploits this generic growth-coupling noise.

Knowing that these pulses are not an essential feature of flagellar biosynthesis, especially since some strains can be genetically modified to produce flagella steadily (***Barker et al., 2004***), this raises the question of the biological advantage of such pulsing dynamics in wild-type strains. Here, we address this question through time series analyses to demonstrate that pulses of flagellar gene activity accompany sudden growth boosts that transiently alleviate the growth burden associated with

flagellar gene expression. By temporarily mitigating this burden, we speculate that some single cells may more effectively explore and colonize new environments as the cell does not have to immediately 'pay' for the cost of flagellum expression. In particular, it would be especially interesting to investigate whether this advantage is more pronounced in nutrient-depleted environments where seeder bacteria with a faster growth rate would also be the ones producing flagella. The fact that the relationship between growth rate and flagella expression, or more generally fluorescent protein expression, at the single-cell level contradicts the population-level trend across strains observed in bulk measurements (***Figures 1 and 4***) is reminiscent of Simpson's paradox and arises from the separation of timescales inherent to these distinct analyses of the same sample (***Jones and Uphoff, 2021***; ***Patange et al., 2018***; ***Sampaio et al., 2022***). Similar time series analyses based on the separation of timescales may also reveal that out-of-balance growth processes serve as a general mechanism for temporarily mitigating the cost of expressing other energetically demanding endogenous genes (***Jain et al., 2025***).

Materials and methods

Strains

For a complete description of the strains, see ***Supplementary file 1B*** in the Supplementary Materials. Briefly: MG1655 flagella reporter strains (E98KFD and E98KFD-Pro) carry the gene for yellow fluorescent protein (mVenus NB) inserted downstream of the Class-1 native promoter of the flhDC operon. At the attB site, a Class-2 promoter, fliFp, controls the activity of the cyan fluorescent protein SCFP3A. The E98KFD-Pro strains differ from E98KFD only in that the native Class-1 promoter is replaced with constitutive synthetic promoters with various strengths (P1: Pro2, P2: Pro4, P3: Pro5, from ***Davis et al., 2011***). The strain construction is described in detail in ***Kim et al., 2020***.

Strains with constitutive expression of fluorescent proteins (MGPro-Venus; ***Garcia Alcala et al., 2025***): each strain carries a low-copy plasmid with an SC101 origin (***Lutz and Bujard, 1997***), in which a synthetic promoter (***Davis et al., 2011***) controls the expression of the sequence of the mVenus NB yellow fluorescent protein (***Balleza et al., 2018***). We renamed the promoters according to their reported strength and measured fluorescence level (P1: Pro2, P2: Pro4, P4: ProB, P5: ProC, P6: ProD).

MC4100*: The classic strain MC4100 (***Peters et al., 2003***; ***Tanouchi et al., 2015***) was modified to correct for its flh5301 point mutation by replacing flhD with the same gene sequence as in MG1655 in order to recover its flagellar expression. We also removed the insertion element IS5 upstream of flhDp so as to keep Class-1 levels consistent with those in MG1655. We introduced the mutation (MotA E98K) to make the strain non-motile. Finally, the strain carries the Class-1 and Class-2 fluorescent reporters as in the MG1655 E98KFD strain. To implement mutations, we used a scarless chromosomal engineering strategy described in ***Kim et al., 2020***, and all transformations were performed via electroporation. Newly created strain (MC4100*) are described in ***Supplementary file 1B*** and are available from the corresponding author upon request, without restrictions.

Growth media

Microfluidic experiments were conducted using a modified version of Teknova's MOPS-EZ Medium: MOPS and Supplement EZ at 1x concentration; 0.4% glycerol serving as the carbon source, phosphate at 1.32 mM. Additionally, we added surfactant (Pluronic F-108, Sigma-Aldrich) at a final concentration of 0.85 g/L. Overnight cultures were prepared using LB medium.

Microfluidic device

We used a Mother Machine device customized for the lab setup. The master mold chip has four independent main feed channels where growth media flows, feeding smaller microchannels with a closed end (length × width × height = 25 μm × 1.2 μm × 1.2 μm).

We fabricated the microfluidic devices by casting a 10:1 mixture of degassed PDMS and curing agent (Sylgard 184, Dow Corning) onto the mold, curing it overnight at 65°C. The individual chips were then removed from the mold, and inlets and outlets were created using a 0.75 mm biopsy punch; then they were bonded to a glass slide (25 mm × 40 mm No. 1.5 coverglass [VWR]) using a plasma cleaner and kept at 65°C for another hour.

Experimental setup

Individual colonies were cultured overnight at 30°C, then concentrated by centrifugation for 1 minute at 8000 × *g* and loaded into the device by pipetting. After loading the device, the sample was left

for 1 hour at 30°C to increase the likelihood of the cells entering the channels by diffusion. Then the trapped cells in the channels began to grow exponentially. We flowed the media through the microfluidic chamber using Tygon tubing (VWR; inside diameter 0.02 inch × outside diameter 0.06 inch), and the flow-through was collected from an outlet via a second tubing into a waste beaker. Growth medium was first pumped at a rate of 20 μL/minute for at least 1 hour to allow the inlets and outlets to be cleared.

Subsequently, the flow rate was decreased to 9 μL/minute, and the temperature was maintained at 30°C. We allowed approximately 3 hours for the cells to adapt before initiating the measurements. The time-lapse images were acquired every 5 minutes on a Zeiss Axiovert 200M microscope equipped with a Plan-Apochromat 40×/1.3 Oil Ph3 Objective, a CCD camera (Hamamatsu C4742-98-24ERG), and a fluorescence excitation LED illumination source (SOLA SE II, Lumencor). We collected images from the phase-contrast and epifluorescent channels (yellow, blue, and/or red protein filters, depending on the strains). The typical duration of the experiment was 30–60 hours.

Microscopy image processing

We used the software BACMMAN (BACteria in Mother Machine Analyzer) (***Ollion et al., 2019***) for cell segmentation and tracking. Using this tool, we customized various parameters to process images according to each specific strain and condition, and then manually corrected any segmentation errors.

Elongation rate and promoter activity calculations

To obtain the elongation rate $ER\left(t\right)$, we determined the first derivative of individual cell sizes $S\left(t\right)$ (from birth to division) using a Savitzky–Golay filter (from Python's SciPy library) with a window of 7 data points (35 minutes). This derivative was then normalized by dividing by the cell size, resulting in the expression $ER\left(t\right)=1/S\left(t\right)\times dS\left(t\right)/dt$.

As a proxy for the promoter activity, we used $PA\left(t\right)=1S\left(t\right)dF\left(t\right)dt$, where $S\left(t\right)$ is the cell size (number of pixels at time t), and $F\left(t\right)$ is the total fluorescence. As in previous literature (***Levine et al., 2012***; ***Locke et al., 2011***; ***Kim et al., 2020***; ***Patange et al., 2018***; ***Potvin-Trottier et al., 2016***), we considered the rate normalized by cell size, thereby adjusting the production rate so that it is expressed per chromosomal equivalent. By normalizing in this way, we focus on the gene's specific activity, rather than the cell's total output.

Given that the total fluorescence $F\left(t\right)$ is the mean intensity of the cell, $C\left(t\right)$, multiplied by its size $S\left(t\right)$, one can show that PA can also be written as $PA\left(t\right)=C\left(t\right)\times ER\left(t\right)+dC\left(t\right)/dt$. We used this equation to calculate the PA, as the mean fluorescence effectively represents the intracellular concentration of the fluorescent signal, which changes more continuously and is less affected by the sudden halving associated with cell division, making the calculation of the derivative less prone to artifacts. We smoothed $C\left(t\right)$ with a window of 35 minutes, which we slid over the time series with an increment of 5 minutes. We calculated the time derivative using the Savitzky–Golay filter, also with a window of 35 minutes.

Binning

To analyze the single-cell relationship between $ER\left(t\right)$ and *Activity(t)* across all acquired time points, we first visualized the data using a scatter plot. We then binned the data based on $Activity\left(t\right)$ values, as the series displayed a clear pulsatile behavior and distinct mean values across strains. Each bin contained an equal number of data points. Finally, for each bin, we calculated the mean values of both $ER\left(t\right)$ and *Activity(t)*.

Pulse-to-pulse analysis

We defined Class-2 pulse-to-pulse intervals by empirically choosing a threshold to determine the on (above the threshold, the pulses) and *off* states (below the threshold). Then, for each sequence of on-off-on states, we identified the two maxima within the Class-2 activity time series and included the 15 timepoints before and after the maxima for each time interval.

We grouped those intervals so that they had similar durations and rescaled them from 0 to 1. Following the rescaling, we applied a sliding window to compute the average of the Class-2 activity within these normalized intervals, which allowed us to move systematically over the data points. This method provided us with a smoothed representation of the Class-2 activity, highlighting the significant

changes. To examine the 'instantaneous' relationship between the pulses and other quantities, we analyzed the same time intervals within the same cells initially defined by the Class-2 activity for the elongation rate, division time, etc. We then rescaled the intervals and used the same window size as for Class-2 activity to average them together. The averages of these quantities for different pulse-to-pulse duration intervals are presented in ***Figure 2—figure supplement 1***.

Autocorrelation and cross-correlation calculations

To quantify the cross-correlation between elongation rate $ER(t)$ and flagellar activity $A(t)$, we first subtracted the mean from each time series $ER'(t) = ER(t) - \langle ER \rangle$, $A'(t) = A(t) - \langle A \rangle$, and then used the function 'correlate' from the SciPy library in Python. The normalized cross-correlation is defined as

$$Corr_{ER,A}(\mathrm{lag}) = \frac{\sum_t ER'(t)\, A'(t+\mathrm{lag})}{\sqrt{\left(\sum_t ER'(t)^2\right)\left(\sum_t A'(t)^2\right)}},$$

where ℓ is the time lag. In practice, we evaluated the numerator using 'correlate' function from SciPy library in Python and normalized by the square root of the product of the zero-lag autocorrelations, by using the same function.

With this convention, the order of the arguments is important: in $Corr_{ER,A}(\ell)$, a maximum at negative ℓ indicates that changes in ER tend to precede changes in A, whereas a maximum at positive ℓ indicates that activity tends to precede elongation rate.

To average the cross-correlation for different lineages from the same strain, we adjusted the lag by dividing it by the average division time of the series. This rescaling ensures that fluctuations due to cell division are aligned across the different series. We applied the same methodology for the auto-correlation function.

Acknowledgements

We thank Maximino Aldana, Mark Kim, Enrique Balleza, and Yuhai Tu for early discussions, advice, and guidance, the Cluzel lab members, Fotios Avgidis for critical reading of the manuscript, and Camilla U Rang for drawing our attention to the importance of the old versus new pole bias. AI-assisted technologies were used to prepare the manuscript. This work was performed, in part, at the Center for Nanoscale Systems, a member of the National Nanotechnology Infrastructure Network, which is supported by the National Science Foundation under award no. ECS-0335765. The Center for Nanoscale Systems is part of Harvard University. This work was supported in part by grant 1615487 from the National Science Foundation and by the US government, ARPA-H. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the US government.

Additional information

Funding

Funder	Grant reference number	Author
National Science Foundation	1615487	Philippe Cluzel

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

Mayra Patricia Garcia Alcala, Conceptualization, Formal analysis, Investigation, Writing – original draft, Writing – review and editing; Josiah C Kratz, Formal analysis, Investigation, Writing – original draft, Writing – review and editing; Philippe Cluzel, Conceptualization, Formal analysis, Supervision, Writing – original draft, Writing – review and editing

Author ORCIDs
Mayra Patricia Garcia Alcala https://orcid.org/0000-0001-7425-1767
Josiah C Kratz https://orcid.org/0000-0001-5524-6209
Philippe Cluzel https://orcid.org/0000-0001-8545-1879

Peer review material
Version of Record:
Reviewer #1 (Public review): https://doi.org/10.7554/eLife.110747.4.sa1
Reviewer #2 (Public review): https://doi.org/10.7554/eLife.110747.4.sa2
Author response https://doi.org/10.7554/eLife.110747.4.sa3

Reviewed preprint v3
Reviewer #1 (Public review): https://doi.org/10.7554/eLife.110747.3.sa2
Reviewer #2 (Public review): https://doi.org/10.7554/eLife.110747.3.sa1
Author response: https://doi.org/10.7554/eLife.110747.3.sa0
eLife Assessment https://doi.org/10.7554/eLife.110747.3.sa3

Reviewed preprint v2
Reviewer #1 (Public review): https://doi.org/10.7554/eLife.110747.2.sa2
Reviewer #2 (Public review): https://doi.org/10.7554/eLife.110747.2.sa1
Author response: https://doi.org/10.7554/eLife.110747.2.sa0
eLife Assessment https://doi.org/10.7554/eLife.110747.2.sa3

Reviewed preprint v1
Reviewer #1 (Public review): https://doi.org/10.7554/eLife.110747.1.sa1
Reviewer #2 (Public review): https://doi.org/10.7554/eLife.110747.1.sa0
eLife Assessment https://doi.org/10.7554/eLife.110747.1.sa2

Additional files

Supplementary files

Supplementary file 1. Supplementary tables. (A) List of plasmids. (B) List of strains. (C) Estimated mVenus proteome fraction in *E. coli* strains expressing mVenus from constitutive promoters of varying strength.

MDAR checklist

Data availability

All the time series (Cell Size, Elongation Rate, Fluorescence, Promoter Activity, etc.) were deposited in Dataverse at https://doi.org/10.7910/DVN/T56JYW and are publicly available. The code to reproduce all the simulations is available on GitHub: https://github.com/josiah-k/Simpsons-paradox (copy archived at ***Kratz, 2026***).

The following dataset was generated:

Author(s)	Year	Dataset title	Dataset URL	Database and Identifier
Garcia Alcala M, Josiah CK, Philippe C	2025	Data for: Out-of-balance Growth Enables Cost-free Flagellum and Other Proteins in a Single Bacterium	https://doi.org/10.7910/DVN/T56JYW	Harvard Dataverse, 10.7910/DVN/T56JYW

References

Bakshi S, Leoncini E, Baker C, Cañas-Duarte SJ, Okumus B, Paulsson J. 2021. Tracking bacterial lineages in complex and dynamic environments with applications for growth control and persistence. *Nature Microbiology* **6**:783–791. DOI: https://doi.org/10.1038/s41564-021-00900-4, PMID: 34017106

Balleza E, Kim JM, Cluzel P. 2018. Systematic characterization of maturation time of fluorescent proteins in living cells. *Nature Methods* **15**:47–51. DOI: https://doi.org/10.1038/nmeth.4509, PMID: 29320486

Barker CS, Prüss BM, Matsumura P. 2004. Increased motility of *Escherichia coli* by insertion sequence element integration into the regulatory region of the flhD operon. *Journal of Bacteriology* **186**:7529–7537. DOI: https://doi.org/10.1128/JB.186.22.7529-7537.2004, PMID: 15516564

Basan M, Zhu M, Dai X, Warren M, Sévin D, Wang Y-P, Hwa T. 2015. Inflating bacterial cells by increased protein synthesis. *Molecular Systems Biology* **11**:836. DOI: https://doi.org/10.15252/msb.20156178, PMID: 26519362

Bienick MS, Young KW, Klesmith JR, Detwiler EE, Tomek KJ, Whitehead TA. 2014. The Interrelationship between promoter strength gene expression, and growth rate. *PLOS ONE* **9**:e109105. DOI: https://doi.org/10.1371/journal.pone.0109105

Biswas K, Sanderson AE, Salman H, Brenner N. 2024. Single-cell growth rate variability in balanced exponential growth. *bioRxiv*. DOI: https://doi.org/10.1101/2024.06.23.600237

Bollenbach T, Quan S, Chait R, Kishony R. 2009. Nonoptimal microbial response to antibiotics underlies suppressive drug interactions. *Cell* **139**:707–718. DOI: https://doi.org/10.1016/j.cell.2009.10.025, PMID: 19914165

Chai Q, Singh B, Peisker K, Metzendorf N, Ge X, Dasgupta S, Sanyal S. 2014. Organization of ribosomes and nucleoids in *Escherichia coli* cells during growth and in quiescence. *The Journal of Biological Chemistry* **289**:11342–11352. DOI: https://doi.org/10.1074/jbc.M114.557348, PMID: 24599955

Chao L, Chan CK, Shi C, Rang UC. 2024. Spatial and temporal distribution of ribosomes in single cells reveals aging differences between old and new daughters of *Escherichia coli*. *eLife* **12**:RP89543. DOI: https://doi.org/10.7554/eLife.89543.3

Chevance FFV, Hughes KT. 2008. Coordinating assembly of a bacterial macromolecular machine. *Nature Reviews. Microbiology* **6**:455–465. DOI: https://doi.org/10.1038/nrmicro1887, PMID: 18483484

Davis JH, Rubin AJ, Sauer RT. 2011. Design, construction and characterization of a set of insulated bacterial promoters. *Nucleic Acids Research* **39**:1131–1141. DOI: https://doi.org/10.1093/nar/gkq810, PMID: 20843779

Dekel E, Alon U. 2005. Optimality and evolutionary tuning of the expression level of a protein. *Nature* **436**:588–592. DOI: https://doi.org/10.1038/nature03842, PMID: 16049495

Dong H, Nilsson L, Kurland CG. 1995. Gratuitous overexpression of genes in *Escherichia coli* leads to growth inhibition and ribosome destruction. *Journal of Bacteriology* **177**:1497–1504. DOI: https://doi.org/10.1128/jb.177.6.1497-1504.1995, PMID: 7883706

Dourado H, Lercher MJ. 2020. An analytical theory of balanced cellular growth. *Nature Communications* **11**:1226. DOI: https://doi.org/10.1038/s41467-020-14751-w, PMID: 32144263

Erickson DW, Schink SJ, Patsalo V, Williamson JR, Gerland U, Hwa T. 2017. A global resource allocation strategy governs growth transition kinetics of *Escherichia coli*. *Nature* **551**:119–123. DOI: https://doi.org/10.1038/nature24299, PMID: 29072300

Fitzgerald DM, Bonocora RP, Wade JT. 2014. Comprehensive mapping of the *Escherichia coli* flagellar regulatory network. *PLOS Genetics* **10**:e1004649. DOI: https://doi.org/10.1371/journal.pgen.1004649, PMID: 25275371

Garcia Alcala M, Josiah CK, Philippe C. 2025. Out-of-balance growth enables cost-free synthesis of the flagellum and other proteins in a single bacterium. Harvard Dataverse. https://doi.org/10.7910/DVN/T56JYW

Goldman S, Aldana M, Cluzel P. 2023. Resonant learning in scale-free networks. *PLOS Computational Biology* **19**:e1010894. DOI: https://doi.org/10.1371/journal.pcbi.1010894, PMID: 36809235

Honda T, Cremer J, Mancini L, Zhang Z, Pilizota T, Hwa T. 2022. Coordination of gene expression with cell size enables *Escherichia coli* to efficiently maintain motility across conditions. *PNAS* **119**:2119. DOI: https://doi.org/10.1073/pnas.2110342119

Jain K, Hauschild R, Bochkareva OO, Roemhild R, Tkačik G, Guet CC. 2025. Pulsatile basal gene expression as a fitness determinant in bacteria. *PNAS* **122**:e2413709122. DOI: https://doi.org/10.1073/pnas.2413709122

Jones EC, Uphoff S. 2021. Single-molecule imaging of LexA degradation in *Escherichia coli* elucidates regulatory mechanisms and heterogeneity of the SOS response. *Nature Microbiology* **6**:981–990. DOI: https://doi.org/10.1038/s41564-021-00930-y, PMID: 34183814

Jun S, Si F, Pugatch R, Scott M. 2018. Fundamental principles in bacterial physiology-history, recent progress, and the future with focus on cell size control: a review. *Reports on Progress in Physics Physical Society* **81**:056601. DOI: https://doi.org/10.1088/1361-6633/aaa628, PMID: 29313526

Kim JM, Garcia-Alcala M, Balleza E, Cluzel P. 2020. Stochastic transcriptional pulses orchestrate flagellar biosynthesis in *Escherichia coli*. *Science Advances* **6**:eaax0947. DOI: https://doi.org/10.1126/sciadv.aax0947, PMID: 32076637

Kiviet DJ, Nghe P, Walker N, Boulineau S, Sunderlikova V, Tans SJ. 2014. Stochasticity of metabolism and growth at the single-cell level. *Nature* **514**:376–379. DOI: https://doi.org/10.1038/nature13582, PMID: 25186725

Klumpp S, Hwa T. 2008. Growth-rate-dependent partitioning of RNA polymerases in bacteria. *PNAS* **105**:20245–20250. DOI: https://doi.org/10.1073/pnas.0804953105, PMID: 19073937

Klumpp S, Zhang Z, Hwa T. 2009. Growth rate-dependent global effects on gene expression in bacteria. *Cell* **139**:1366–1375. DOI: https://doi.org/10.1016/j.cell.2009.12.001, PMID: 20064380

Kojima S, Blair DF. 2001. Conformational change in the stator of the bacterial flagellar motor. *Biochemistry* **40**:13041–13050. DOI: https://doi.org/10.1021/bi011263o, PMID: 11669642

Korem Kohanim Y, Levi D, Jona G, Towbin BD, Bren A, Alon U. 2018. A bacterial growth law out of steady state. *Cell Reports* **23**:2891–2900. DOI: https://doi.org/10.1016/j.celrep.2018.05.007, PMID: 29874577

Kratz JC, Banerjee S. 2023. Dynamic proteome trade-offs regulate bacterial cell size and growth in fluctuating nutrient environments. *Communications Biology* **6**:486. DOI: https://doi.org/10.1038/s42003-023-04865-4, PMID: 37147517

Kratz J. 2026. Simpsons-paradox. swh:1:rev:2cca41cb90e7fe563e758bea04199d61943a3109. Software Heritage. https://archive.softwareheritage.org/swh:1:dir:e5fed2ec621d066dc942dd180e459f237cf47f24;origin=https://github.com/josiah-k/Simpsons-paradox;visit=swh:1:snp:363ed42041d9216a9d10f959a006fb5fdd95ad4e;anchor=swh:1:rev:2cca41cb90e7fe563e758bea04199d61943a3109

Levine JH, Fontes ME, Dworkin J, Elowitz MB. 2012. Pulsed feedback defers cellular differentiation. *PLOS Biology* **10**:e1001252. DOI: https://doi.org/10.1371/journal.pbio.1001252, PMID: 22303282

Locke JCW, Young JW, Fontes M, Hernández Jiménez MJ, Elowitz MB. 2011. Stochastic pulse regulation in bacterial stress response. *Science* **334**:366–369. DOI: https://doi.org/10.1126/science.1208144, PMID: 21979936

Lutz R, Bujard H. 1997. Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic Acids Research* **25**:1203–1210. DOI: https://doi.org/10.1093/nar/25.6.1203, PMID: 9092630

Moffitt JR, Lee JB, Cluzel P. 2012. The single-cell chemostat: an agarose-based, microfluidic device for high-throughput, single-cell studies of bacteria and bacterial communities. *Lab on a Chip* **12**:1487–1494. DOI: https://doi.org/10.1039/c2lc00009a, PMID: 22395180

Mori M, Zhang Z, Banaei-Esfahani A, Lalanne JB, Okano H, Collins BC, Schmidt A, Schubert OT, Lee DS, Li GW, Aebersold R, Hwa T, Ludwig C. 2021. From coarse to fine: the absolute. *Molecular Systems Biology* **17**:1–23. DOI: https://doi.org/10.15252/msb.20209536

Morimoto YV, Nakamura S, Kami-ike N, Namba K, Minamino T. 2010. Charged residues in the cytoplasmic loop of MotA are required for stator assembly into the bacterial flagellar motor. *Molecular Microbiology* **78**:1117–1129. DOI: https://doi.org/10.1111/j.1365-2958.2010.07391.x, PMID: 21091499

Ni B, Ghosh B, Paldy FS, Colin R, Heimerl T, Sourjik V. 2017. Evolutionary remodeling of bacterial motility checkpoint control. *Cell Reports* **18**:866–877. DOI: https://doi.org/10.1016/j.celrep.2016.12.088, PMID: 28122238

Ni B, Colin R, Link H, Endres RG, Sourjik V. 2020. Growth-rate dependent resource investment in bacterial motile behavior quantitatively follows potential benefit of chemotaxis. *PNAS* **117**:595–601. DOI: https://doi.org/10.1073/pnas.1910849117, PMID: 31871173

Ollion J, Elez M, Robert L. 2019. High-throughput detection and tracking of cells and intracellular spots in mother machine experiments. *Nature Protocols* **14**:3144–3161. DOI: https://doi.org/10.1038/s41596-019-0216-9, PMID: 31554957

Papagiannakis A, Yu Q, Govers SK, Lin WH, Wingreen NS, Jacobs-Wagner C. 2025. Nonequilibrium polysome dynamics promote chromosome segregation and its coupling to cell growth in *Escherichia coli*. *eLife* **15**:4276. DOI: https://doi.org/10.7554/eLife.104276.2

Park J, Dies M, Lin Y, Hormoz S, Smith-Unna SE, Quinodoz S, Hernández-Jiménez MJ, Garcia-Ojalvo J, Locke JCW, Elowitz MB. 2018. Molecular time sharing through dynamic pulsing in single cells. *Cell Systems* **6**:216–229. DOI: https://doi.org/10.1016/j.cels.2018.01.011, PMID: 29454936

Patange O, Schwall C, Jones M, Villava C, Griffith DA, Phillips A, Locke JCW. 2018. *Escherichia coli* can survive stress by noisy growth modulation. *Nature Communications* **9**:5333. DOI: https://doi.org/10.1038/s41467-018-07702-z, PMID: 30559445

Pavlou A, Cinquemani E, Pinel C, Giordano N, Mathilde VMG, Mihalcescu I, Geiselmann J, de Jong H. 2025. Single-cell data reveal heterogeneity of investment in ribosomes across a bacterial population. *Nature Communications* **16**:285. DOI: https://doi.org/10.1038/s41467-024-55394-5, PMID: 39746998

Peters JE, Thate TE, Craig NL. 2003. Definition of the *Escherichia coli* MC4100 genome by use of a DNA array. *Journal of Bacteriology* **185**:2017–2021. DOI: https://doi.org/10.1128/JB.185.6.2017-2021.2003, PMID: 12618467

Potvin-Trottier L, Lord ND, Vinnicombe G, Paulsson J. 2016. Synchronous long-term oscillations in a synthetic gene circuit. *Nature* **538**:514–517. DOI: https://doi.org/10.1038/nature19841, PMID: 27732583

Proenca AM, Tuğrul M, Nath A, Steiner UK. 2024. Progressive decline in old pole gene expression signal enhances phenotypic heterogeneity in bacteria. *Science Advances* **10**:eadp8784. DOI: https://doi.org/10.1126/sciadv.adp8784

Sampaio NMV, Blassick CM, Andreani V, Lugagne JB, Dunlop MJ. 2022. Dynamic gene expression and growth underlie cell-to-cell heterogeneity in *Escherichia coli* stress response. *PNAS* **119**:e2115032119. DOI: https://doi.org/10.1073/pnas.2115032119, PMID: 35344432

Sassi AS, Garcia-Alcala M, Kim MJ, Cluzel P, Tu Y. 2020. Filtering input fluctuations in intensity and in time underlies stochastic transcriptional pulses without feedback. *PNAS* **117**:26608–26615. DOI: https://doi.org/10.1073/pnas.2010849117, PMID: 33046652

Sassi AS, Garcia-Alcala M, Aldana M, Tu Y. 2022. Protein concentration fluctuations in the high expression regime: Taylor's law and its mechanistic origin. *Physical Review* **12**:011051. DOI: https://doi.org/10.1103/physrevx.12.011051, PMID: 35756903

Schavemaker PE, Lynch M. 2022. Flagellar energy costs across the tree of life. *eLife* **11**:e77266. DOI: https://doi.org/10.7554/eLife.77266, PMID: 35881430

Scott M, Gunderson CW, Mateescu EM, Zhang Z, Hwa T. 2010. Interdependence of Cell Growth. *Science* **330**:1099–1102. DOI: https://doi.org/10.1126/science.1192588

Scott M, Hwa T. 2011. Bacterial growth laws and their applications. *Current Opinion in Biotechnology* **22**:559–565. DOI: https://doi.org/10.1016/j.copbio.2011.04.014, PMID: 21592775

Shachrai I, Zaslaver A, Alon U, Dekel E. 2010. Cost of unneeded proteins in *E. coli* is reduced after several generations in exponential growth. *Molecular Cell* **38**:758–767. DOI: https://doi.org/10.1016/j.molcel.2010.04.015, PMID: 20434381

Shi C, Chao L, Proenca AM, Qiu A, Chao J, Rang CU. 2020. Allocation of gene products to daughter cells is determined by the age of the mother in single *Escherichia coli* cells. *Proceedings Biological Sciences* **287**:20200569. DOI: https://doi.org/10.1098/rspb.2020.0569, PMID: 32370668

Stoebel DM, Dean AM, Dykhuizen DE. 2008. The cost of expression of *Escherichia coli* lac operon proteins is in the process, not in the products. *Genetics* **178**:1653–1660. DOI: https://doi.org/10.1534/genetics.107.085399, PMID: 18245823

Tanouchi Y, Pai A, Park H, Huang S, Stamatov R, Buchler NE, You L. 2015. A noisy linear map underlies oscillations in cell size and gene expression in bacteria. *Nature* **523**:357–360. DOI: https://doi.org/10.1038/nature14562, PMID: 26040722

Tanouchi Y, Pai A, Park H, Huang S, Buchler NE, You L. 2017. Long-term growth data of *Escherichia coli* at a single-cell level. *Scientific Data* **4**:170036. DOI: https://doi.org/10.1038/sdata.2017.36, PMID: 28350394

van Heerden JH, Berkvens A, de Groot DH, Bruggeman FJ. 2024. Growth consequences of the inhomogeneous organization of the bacterial cytoplasm. *eLife* **15**:9996. DOI: https://doi.org/10.7554/eLife.99996.1

Wang P, Robert L, Pelletier J, Dang WL, Taddei F, Wright A, Jun S. 2010. Robust growth of *Escherichia coli*. *Current Biology* **20**:1099–1103. DOI: https://doi.org/10.1016/j.cub.2010.04.045, PMID: 20537537

Wang X, Wood TK. 2011. IS5 inserts upstream of the master motility operon flhDC in a quasi-Lamarckian way. *The ISME Journal* **5**:1517–1525. DOI: https://doi.org/10.1038/ismej.2011.27, PMID: 21390082

Yi X, Dean AM. 2016. Phenotypic plasticity as an adaptation to a functional trade-off. *eLife* **5**:e19307. DOI: https://doi.org/10.7554/eLife.19307, PMID: 27692064

Appendix 1

Supplemental materials

Experimental validation of functional flagella expression in reporter strains

We performed a motility plate assay to qualitatively verify that the strains we claim have different levels of flagellar expression display phenotypes consistent with their promoter strengths as reported by ***Davis et al., 2011***. On low-agar plates, we inoculated strains with the native Class-1 promoter that was replaced by Pro2, Pro4, or Pro5 (P1–P3 in the main text), together with the following controls: the wild type MG1655, MotA E98K MG1655 (strain with a mutation that disables the rotation of the flagella; ***Kojima and Blair, 2001***; ***Morimoto et al., 2010***), MotA E98K Pro4, Pro4 ΔfliC, and MG1655 strain carrying the IS5 insertion (a motile strain known to have elevated flagellar expression; ***Wang and Wood, 2011***). 1 μL of liquid culture was spotted onto the plate and incubated for ~12 hours at 30°C. Motile strains formed characteristic swarming rings (***Figure 1—figure supplement 3A***). The strains with weaker promoters produced smaller swarm diameters than those with stronger promoters, in the same order expected from Davis et al. and consistent with our fluorescence measurements after introduction of the reporter. There was no growth ring for the strains carrying the MotA E98K nor ΔfliC mutations. This confirms that the reporter activity tracks the level of functional flagellar expression across strains.

Then, to address whether cells are flagellated and motile in the mother-machine channels, we compared the wild type MG1655 and MotA E98K MG1655 under identical microfluidic conditions. Cells were loaded in stationary phase and allowed to recover before time-lapse imaging. In the wild-type strain, we observed that ~35% of lineages left the growth channels during the experiment after 20 hours, whereas non-motile cells did not leave the channels. Because the only difference between these strains is the MotA E98K mutation, this strong differential escape behavior indicates that cells in the channels do assemble flagellar filaments and motors, and when motors are activated, the resulting motility is sufficient for cells to swim out of the channels. In ***Figure 1—figure supplement 3B***, we show one channel of the mother machine filled with cells outlined in blue, which suddenly leave the channels. Together with the motility plate assay, this supports that our reporter signal corresponds to expression of functional flagella under the conditions of our experiments.

Simulation details: Stochastic proteome allocation model with division-coupled noise

Single-cell simulations were implemented using rules for proteome allocation and division timing from ***Kratz and Banerjee, 2023***, allowing for analysis of instantaneous proteome allocation and growth rate at the level of the individual cell. For completeness, here we include a description of all relevant equations required to define the model, but for a detailed derivation and model justification, we refer readers to ***Kratz and Banerjee, 2023***. Each individual cell can be described by a set of state variables which include its coarse-grained proteome allocation to five sectors of proteins grouped together based on function and defined by their mass fractions ϕ_i for $i \in \{R, P, X, Q, U\}$, specifically ribosomal (R), metabolic (P), division (X), housekeeping (Q), and exogenously expressed 'unnecessary' (U) proteins. These sectors are coupled to the amino acid mass fraction of the cell, a. The ribosomal (ϕ_R) and amino acid (a) mass fractions together set the exponential growth rate, $\kappa = \mathrm{d} \ln V / \mathrm{d}t$, given by

$$\kappa = \kappa_t (a)\, \phi_R - \mu_{ns},$$

where μ_{ns} is the nonspecific degradation rate constant and $\kappa_t (a) = \kappa_t^0 g (a)$ is the translational efficiency, which is dependent on a such that translation becomes significantly attenuated at low amino acid levels (specifically $g (a) = (a/a_t)^2 / \left[1 + (a/a_t)^2\right]$). The dynamics of each sector are given by

$$\mathrm{d}\phi_i / \mathrm{d}t = \kappa_t (a)\, \phi_R \left(f_i (t) - \phi_i\right),$$

where $f_i (t)$ denotes the fraction of the total translational flux devoted to sector i. Both $f_R (t)$ and $f_X (t)$ obtain their time dependence through their dependence on other dynamic quantities, specifically $a (t)$ and $\phi_R^{\max} (t)$. These expressions are given by

$$f_x(a) = \alpha\left(\phi_R^{max} - f_R(a)\right) + \beta,$$

$$f_R\left(a, \phi_R^{\max}\right) = \frac{-f'(a)\, g(a)\, \phi_R^{\max}}{-f'(a)\, g(a) + f(a)\, g'(a)},$$

where $f(a) = 1/\left[1 + \left(a/a_n\right)^2\right]$, $\phi_R^{\max}(t)$ is a dynamic parameter constraining the maximum allocation to ribosomal proteins, and here $y'(a)$ denotes the derivative of y with respect to a. f_Q is taken to be invariant to nutrient changes, thus imposing the additional constraint $f_R + f_P + f_X = 1 - f_Q = \phi^{\max} = \phi_R^{\max} + \beta$. The dynamics of f_U are discussed below. The dynamics of ϕ_i are coupled to the dynamics of a, which are given by

$$\mathrm{d}a/\mathrm{d}t = \kappa_n(a, c)(\phi^{max} - \phi_R - \phi_X - \phi_U) + \mu_{ns} - \kappa_t(a)\phi_R,$$

where $\kappa_n(a, c) = \kappa_n^0(c) f(a)$ is the nutritional efficiency and c is the extracellular nutrient quality. Finally, a division event occurs when a sufficient number of X proteins have accumulated. The dynamics of X are given by

$$\frac{d\widetilde{X}}{dt} = \gamma f_X(a)\left(\kappa + \mu_{ns}\right) V - \mu_X \widetilde{X},$$

where here we track the normalized quantity $\widetilde{X} = X/X_0$ such that a cell divides when $\widetilde{X} = 1$, μ_X denotes the division protein degradation rate, and γ is a conversion factor relating protein mass fraction to protein number (see ***Kratz and Banerjee, 2023*** for details). To match experiments, at each division event a single daughter cell was simulated which inherited the intracellular composition of the mother cell, maintaining the overall population size.

Previous experimental work has shown that the major driver of heterogeneity in ribosome concentration across single cells is not stochastic production, but rather unequal partitioning of ribosomes at division due to the formation of larger clusters (***Chai et al., 2014***). Motivated by this, we modeled the redistribution of ribosomes at division through a stochastic differential equation in which perturbations to the maximum allocation to ribosomes only occurred at division times, t_d, then decayed slowly with a mean-reversion timescale $\tau_R \gg 0$. Specifically:

$$\mathrm{d}\phi_R^{\max} = \frac{1}{\tau_R}\left(\overline{\phi_R^{\max}} - \phi_R^{\max}\right)\mathrm{d}t + \phi_R^{\max}\delta\left(\mathrm{t} - \mathrm{t_d}\right)\xi_\mathrm{d}$$

where $\overline{\phi_R^{\max}}$ is the mean maximum allocation to ribosomes, $\delta(\cdot)$ is the Dirac delta function, and $\xi_d N\left(0, \sigma_R^2\right)$ denotes random jump magnitudes drawn from a Gaussian distribution of mean 0 and variance σ_R^2. In the absence of clustering, $\sigma_R = 0$.

In contrast, we assumed the main driver of heterogeneity in unnecessary protein allocation is due to stochastic gene expression. As such, we modeled noise in unnecessary protein expression by assuming it continuously fluctuated around its mean allocation value $\bar{f_U}$ with timescale τ_U and with volatility coefficient σ_U, yielding

$$\mathrm{d}f_U = \frac{1}{\tau}\left(\overline{f_U} - f_U\right)\mathrm{d}t + \sigma_U f_U \mathrm{d}W_t,$$

where W_t denotes a Wiener process. Assuming constant cellular density, protein mass fraction is proportional to protein concentration, allowing us to rewrite our previous definition of promoter activity as $A_U(t) = \phi_U(t)\,\kappa(t) + \frac{d\phi_U(t)}{dt}$ where $\phi_U(t)$ is the mass fraction of unnecessary proteins and $\kappa(t)$ is the single-cell growth rate at time t. Using this expression, along with our definition of sector dynamics from above $\left(\frac{d\phi_U(t)}{dt} = \kappa(t)\left(f_U(t) - \phi_U(t)\right)\right)$, the activity of unnecessary protein expression could then be calculated as

$$A_U(t) = f_U(t)\,\kappa(t).$$

To simulate, we numerically integrated the set of coupled stochastic differential equations detailed above for each cell. The code used to generate all simulation results is available at https://github.com/josiah-k/Simpsons-paradox (copy archived at ***Kratz, 2026***). The single-cell growth rate and corresponding activity for 3000 individual cells was computed at intervals of Δt=3 min for 75 hours total.

Parameters

Translational efficiency, κ_t^0 (h^{-1}):	5
Nutritional efficiency, κ_n^0 (h^{-1}):	1.5
Mean-reversion timescale (ribosomes), τ_R (h):	6.25
Mean-reversion timescale (unnecessary proteins), τ_U (h):	0.0625
Unnecessary protein expression volatility coefficient σ_U:	0.3
Ribosome protein expression volatility coefficient σ_R:	0.06
Average unnecessary protein allocation fraction, $\overline{f_U}$:	0.005, 0.010, 0.037, 0.073, 0.108, 0.144
Division protein degradation rate, μ_X (h^{-1}):	0.6
Average maximum flux allocation to ribosomes, $\overline{\phi_R^{max}}$ (***Kratz and Banerjee, 2023***):	0.55
Translation attenuation threshold, a_t (***Kratz and Banerjee, 2023***):	10^{-4}
Feedback inhibition threshold, a_n (***Kratz and Banerjee, 2023***):	10^{-3}
Nonspecific degradation rate, μ_{ns} (h^{-1}) (***Kratz and Banerjee, 2023***):	0
Basal allocation to division protein production, $\gamma\beta$ (μm^{-3}) (***Kratz and Banerjee, 2023***):	1.1
Dynamic allocation to X protein production, $\gamma\alpha$ (μm^{-3}) (***Kratz and Banerjee, 2023***):	4.5