eLife

Reviewed Preprint
v1 • April 1, 2026
Not revised

Reviewed Preprint
v2 • August 5, 2026
Revised by authors

Out-of-balance Growth Enables Cost-free Synthesis of the Flagellum and Other Proteins in a Single Bacterium

Mayra 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

For correspondence: cluzel@mcb.harvard.edu

Author Contributions: P.C. and M.G.-A. conceived the study and analyzed the data. M.G.-A. performed experiments and statistical analyses; J.C.K. performed the simulations. P.C. supervised the project. M.G.-A. J.C.K and P.C. wrote and prepared the manuscript.

Competing interests: No competing interests declared

Funding: See page 13

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

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 E. 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 - is consistent with the data.

https://doi.org/10.7554/eLife.110747.2.sa3

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 E. 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 [1-4]. Recent studies have built on this principle to derive powerful, predictive growth laws [2, 5-7]. 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 [6]. Consequently, the synthesis of unnecessary proteins is accompanied

by a growth burden that increases with protein expression levels [1, 4, 7]. 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 [8, 9]. Balanced growth also ensures that cell growth has reached a steady-state and does not run into resource bottlenecks that could impair its physiology [10, 11]. 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 [12-14], such as sudden and stochastic molecular events that inherently accompany growth and gene expression observed in single cells [15-22], 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 *E. coli* have reported that flagellar genes, responsible for the synthesis of one of the most energy-intensive bacterial organelles, exhibit large stochastic expression pulses [20]. 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 [23, 24] (Fig. S1A). During flagella synthesis, the cell directs a sizable portion of protein synthesis resources to this process [25-27]. 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 [28-30], 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 [30, 31]. The activity of flagellar genes follows a well-characterized transcriptional cascade [23, 32]. 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 (Fig. S1A). 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 [20]. 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' [33, 34], which allows bacteria to grow exponentially under controlled conditions for extended periods of time (Fig. S1B). 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 min resolution over ~30 generations (Fig. S1C). 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 (Fig. S2). All experiments were performed in an MG1655 background harboring a *motA* mutation that disables swimming motility but does not impair flagellar assembly [35, 36], thereby isolating the regulatory and growth-related dynamics of the flagellar system (see Materials and Methods, Fig. S3). 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 [20]. 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 min time window (Fig. 1 Left). These pairs were then sorted and grouped into bins based on their activity levels (Fig. S4). 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; [20, 37]) 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 (Fig. 1 Right, dashed line, Fig. S5), 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 (Fig. 1 Right, Fig. S5). 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 (Fig. S6, [28-30]; see Fig. S7 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 (Fig. 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 (Fig. S8). We interpret this 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 [20]. 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.

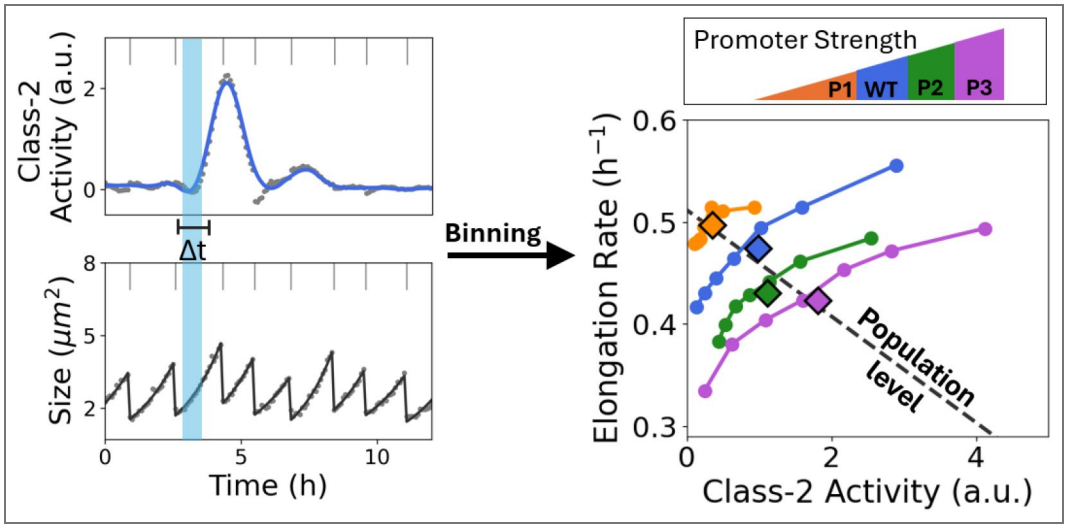


Fig. 1. Relationship between instantaneous growth rate and flagellar gene activity in single cells versus population.

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 $ER(t) = (1/S)(\Delta S/\Delta t)$. 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 min and then bin the pair activity-elongation rate (see Fig. S4). 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 [37]. 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 [30]. 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 Fig. S3D for all states. The number of time traces is n= 69, 107, 96, 67 for the strains with the promoter P1, WT, P2, and P4, respectively, all with a duration of ~47 hours.

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 (Fig. 2, top; Materials and Methods).

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 (Fig. 2, middle and bottom). This relationship remains consistent across pulse-to-pulse intervals of different lengths (see Fig. S9).

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 (Fig. S10, 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 (Fig. S11), 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 (Fig. S10). 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 (Fig. S12, S13). 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 (see Fig. S13) is consistent with the scenario described by Kiviet and Nghe et al. [38], 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 [39-42]. While Pavlou et al. found no immediate association between ribosome abundance and unequal division [41], several super-resolution studies suggest that ribosome heterogeneity arises from asymmetric partitioning during cell division [39, 40, 43]. 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 [15, 44]. With these results in mind and given that the 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) (Fig.

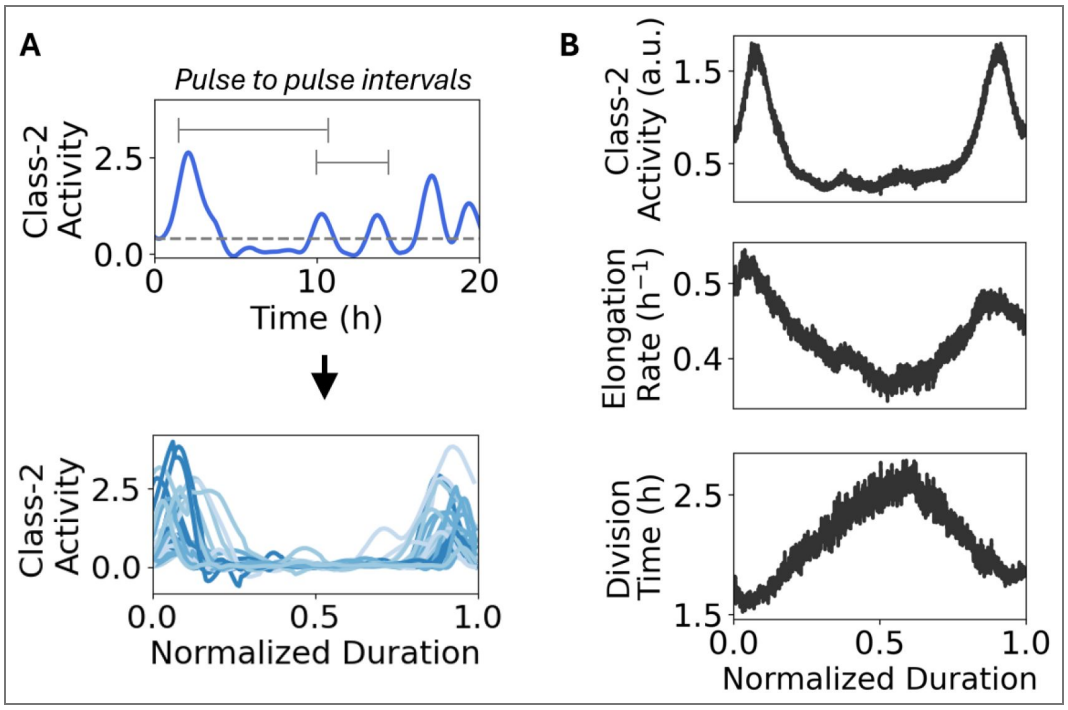


Fig. 2. Cells grow faster during pulses.

(**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.

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 (Fig. S14).

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 (Fig. 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 [45], and that aging at the old pole enhances phenotypic heterogeneity [46]. 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 [40]. Together, these observations support the hypothesis that differences in growth rate between sister cells are governed by difference of 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 (Fig. 3B, upper right) and compared their elongation rates (Fig. 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 (Fig. 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.

Next, we asked whether these observations could be strain dependent. For example, the MC4100 E. coli strain is known to exhibit growth rate variations over a much longer timescale than that of the MG1655 strain [44]. Consistent with previous work, MG1655 showed rapid fluctuation dynamics (1.5 generations), whereas the elongation rate of MC4100 fluctuated much more slowly (~10 generations). Our statistical analyses showed no correlation between MC4100's long timescale growth fluctuations and flagellar gene activity, indicating that these slow growth modes are largely decoupled from flagellar expression (Fig. S15).

Next, we explored whether the single-cell behavior observed in Fig. 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. Although our bulk measurements (Fig. 4A, diamonds, Fig. S16C) as in [5] 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. 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 (Fig. 4A; Fig. S15A, B; cross-correlation between elongation rate and activity in Fig. S17). The result in Fig. 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.

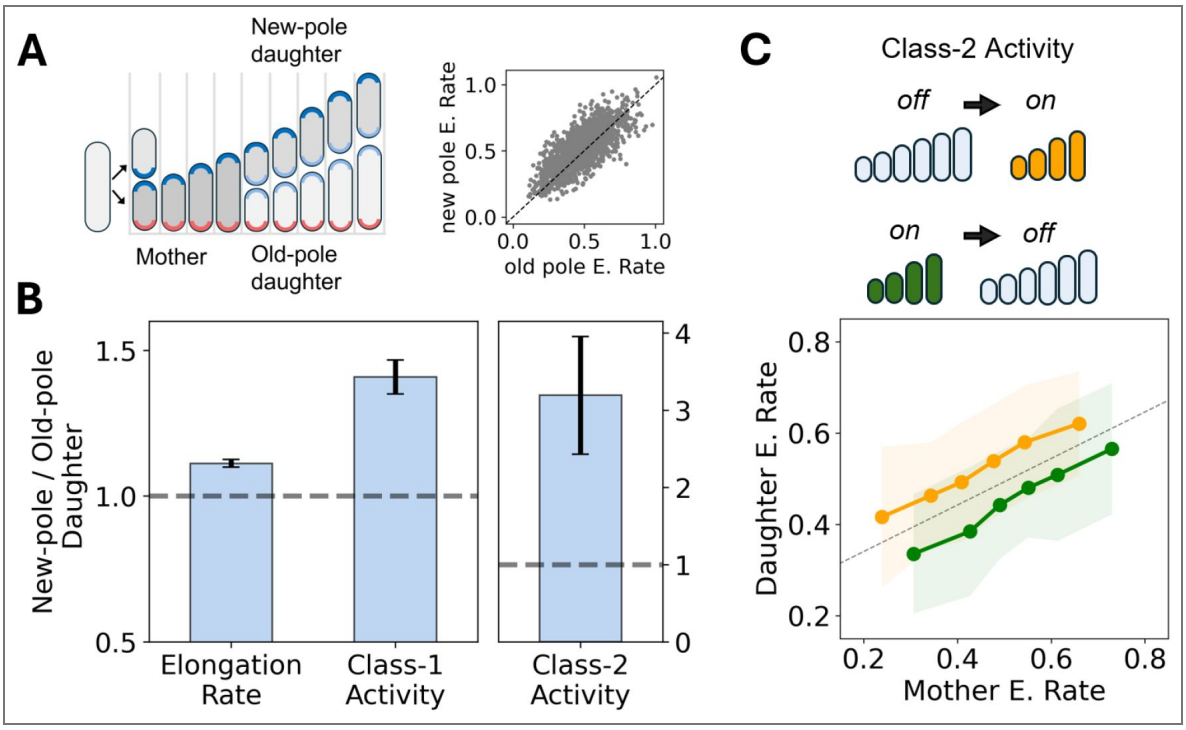


Fig. 3. Short-timescale shifts in Class-2 activity correlate with immediate changes in growth.

(**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.

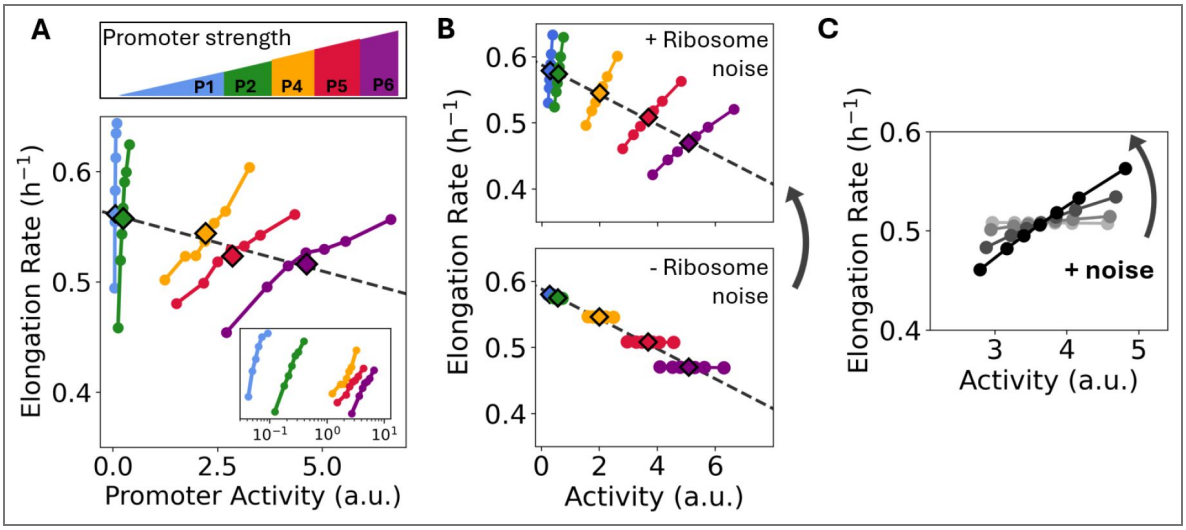


Fig. 4. Relationship between instantaneous growth and promoter activity in single cells versus population for different levels of constitutive expression of unnecessary proteins.

(**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 (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 (σ_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 σ_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 [14], and other simulation parameters are described in SI Section 1.

To illustrate this point, we performed single-cell simulations using a computational model (Fig. 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 (Fig. 4B, top). Moreover, the slope of this relationship becomes steeper as the noise in the distribution of growth factors becomes larger (Fig. 4C). In the absence of noise, the relationship shows no correlation between growth rate and protein synthesis (Fig. 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. [20] 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 [24]. While this molecular mechanism demonstrates that pulses can arise without feedback, it does not address how such pulses influence cellular growth.

Knowing that these pulses are not an essential feature of flagellar biosynthesis, especially since some strains can be genetically modified to produce flagella steadily [47], this raises the question of the biological advantage of such a 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 (Fig. 1, Fig. 4) is reminiscent of Simpson's paradox and arises from the separation of timescales inherent to these distinct analyses of the same sample [17, 18, 22]. 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 [21].

Materials and Methods

Strains

For a complete description of the strains see, Table 2 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 [37]). The strain construction is described in detail in [20].

Strains with constitutive expression of fluorescent proteins (MGPro-Venus; [48]): each strain carries a low-copy plasmid with an SC101 origin [49], in which a synthetic promoter [37] controls the expression of the sequence of the mVenus NB yellow fluorescent protein [50]. We renamed the promoters according to their reported strength and measured fluorescence level (P1: Pro2, P2: Pro4, P4: ProB, P5: ProC, P6: ProD).

MC*: The classic strain MC4100 [44, 51] 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 [20], and all transformations were performed via electroporation.

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 4 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/min for at least 1 hour to allow the inlets and outlets to be cleared.

Subsequently, the flow rate was decreased to 9 µl/min, 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 40x/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) [52] 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 $ER(t)$, we determined the first derivative of individual cell sizes $S(t)$ (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 $ER(t) = 1/S(t) \times dS(t)/dt$.

As a proxy for the promoter activity, we used $PA(t) = 1/S(t)\ dF(t)/dt$, where $S(t)$ is the cell size (number of pixels at time t), and $F(t)$ is the total fluorescence. As in previous literature [17, 20, 53-55], 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 $F(t)$ is the mean intensity of the cell, $C(t)$, multiplied by its size $S(t)$, one can show that PA can also be written as $PA(t) = C(t) \times ER(t) + dC(t)/dt$. 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 $C(t)$ with a window of 35 minutes, which we slid over the time series with an increment of 5 min. 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 $ER(t)$ and *Activity(t)* across all acquired time points, we first visualized the data using a scatter plot. We then binned the data based on $Activity(t)$ 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 $ER(t)$ 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 Fig. S6.

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

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

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 $Corr_{ER,A}(l)$, 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 autocorrelation function.

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*.

Acknowledgements

We thank Maximino Aldana, Mark Kim and Enrique Balleza 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. Funding: 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 United States 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 United States Government.

Additional files

Supplemental Material

Additional information

Funding

Funder	Grant reference number	Author
National Science Foundation (NSF)	1615487	Philippe Cluzel

Author ORCID iDs

Philippe Cluzel: https://orcid.org/0000-0001-8545-1879

References

1. **Dekel E,** Alon U (2005) Optimality and evolutionary tuning of the expression level of a protein. *Nature* **436**:588-592 https://doi.org/10.1038/nature03842 | PubMed
2. **Scott M,** Hwa T (2011) Bacterial growth laws and their applications. *Current Opinion in Biotechnology* **22**:559-565 https://doi.org/10.1016/j.copbio.2011.04.014 | PubMed
3. **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 https://doi.org/10.1534/genetics.107.085399 | PubMed
4. **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 https://doi.org/10.1128/jb.177.6.1497-1504.1995 | PubMed
5. **Basan M,** et al. (2015) Inflating bacterial cells by increased protein synthesis. *Molecular Systems Biology* **11**:836-836 https://doi.org/10.15252/msb.20156178 | PubMed
6. **Scott M,** et al. (2010) Interdependence of Cell Growth. *Science* **330**:1099-1102 https://doi.org/10.1126/science.1192588 | PubMed
7. **Shachrai I,** et al. (2010) Cost of Unneeded Proteins in E. coli Is Reduced after Several Generations in Exponential Growth. *Molecular Cell* **38**:758-767 https://doi.org/10.1016/j.molcel.2010.04.015 | PubMed
8. **Klumpp S,** Hwa T (2008) Growth-rate-dependent partitioning of RNA polymerases in bacteria. *Proceedings of the National Academy of Sciences of the United States of America* **105**:20245-20250 https://doi.org/10.1073/pnas.0804953105 | PubMed
9. **Klumpp S,** Zhang Z, Hwa T (2009) Growth Rate-Dependent Global Effects on Gene Expression in Bacteria. *Cell* **139**:1366-1375 https://doi.org/10.1016/j.cell.2009.12.001 | PubMed
10. **Dourado H,** Lercher MJ (2020) An analytical theory of balanced cellular growth. *Nature Communications* **11** https://doi.org/10.1038/s41467-020-14751-w | PubMed
11. **Jun S,** et al. (2017) Fundamental Principles in Bacterial Physiology - History, Recent progress, and the Future with Focus on Cell Size Control: A Review. *Rep Prog Phys* https://doi.org/10.1088/1361-6633/aaa628 | PubMed
12. **Korem Kohanim Y,** et al. (2018) A Bacterial Growth Law out of Steady State. *Cell Reports* **23**:2891-2900 https://doi.org/10.1016/j.celrep.2018.05.007 | PubMed
13. **Bollenbach T,** et al. (2009) Nonoptimal Microbial Response to Antibiotics Underlies Suppressive Drug Interactions. *Cell* **139**:707-718 https://doi.org/10.1016/j.cell.2009.10.025 | PubMed
14. **Kratz JC,** Banerjee S (2023) Dynamic proteome trade-offs regulate bacterial cell size and growth in fluctuating nutrient environments. *Communications Biology* **6** https://doi.org/10.1038/s42003-023-04865-4 | PubMed
15. **Biswas K,** et al. (2024) Single-cell Growth Rate Variability in Balanced Exponential Growth. *bioRxiv* https://doi.org/10.1101/2024.06.23.600237
16. **Park J,** et al. (2018) Molecular Time Sharing through Dynamic Pulsing in Single Cells. *Cell Systems* **6**:216-229.e15 https://doi.org/10.1016/j.cels.2018.01.011 | PubMed
17. **Patange O,** et al. (2018) Escherichia coli can survive stress by noisy growth modulation. *Nature Communications* **9** https://doi.org/10.1038/s41467-018-07702-z | PubMed
18. **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 https://doi.org/10.1038/s41564-021-00930-y | PubMed

19. **Goldman S,** Aldana M, Cluzel P (2023) Resonant learning in scale-free networks. *PLOS Computational Biology* **19**:e1010894 https://doi.org/10.1371/journal.pcbi.1010894 | PubMed
20. **Mark Kim J,** et al. (2020) Stochastic transcriptional pulses orchestrate flagellar biosynthesis in Escherichia coli. *Science Advances* **6** https://doi.org/10.1126/sciadv.aax0947 | PubMed
21. **Jain K,** et al. (2025) Pulsatile basal gene expression as a fitness determinant in bacteria. *Proceedings of the National Academy of Sciences* **122**:e2413709122-e2413709122 https://doi.org/10.1073/pnas.2413709122 | PubMed
22. **Sampaio NMV,** et al. (2022) Dynamic gene expression and growth underlie cell-to-cell heterogeneity in Escherichia coli stress response. *Proc. Natl. Acad. Sci. U.S.A* **119**:e2115032119 https://doi.org/10.1073/pnas.2115032119 | PubMed
23. **Chevance FF,** Hughes KT (2008) Coordinating assembly of a bacterial macromolecular machine. *Nat Rev Microbiol* **6**:455-465 https://doi.org/10.1038/nrmicro1887 | PubMed
24. **Sassi AS,** et al. (2020) Filtering input fluctuations in intensity and in time underlies stochastic transcriptional pulses without feedback. *Proceedings of the National Academy of Sciences of the United States of America* **117**:26608-26615 https://doi.org/10.1073/pnas.2010849117 | PubMed
25. **Erickson DW,** et al. (2017) A global resource allocation strategy governs growth transition kinetics of Escherichia coli. *Nature* https://doi.org/10.1038/nature24299 | PubMed
26. **Mori M,** et al. (2021) From coarse to fine: the absolute Escherichia coli proteome under diverse growth conditions. *Molecular Systems Biology* **17**:1-23 https://doi.org/10.15252/msb.20209536 | PubMed
27. **Schavemaker PE,** Lynch M (2022) Flagellar energy costs across the tree of life. *eLife* **11** https://doi.org/10.7554/elife.77266 | PubMed
28. **Honda T,** et al. (2022) Coordination of gene expression with cell size enables Escherichia coli to efficiently maintain motility across conditions. *PNAS* https://doi.org/10.1073/pnas.2110342119 | PubMed
29. **Ni B,** et al. (2017) Evolutionary Remodeling of Bacterial Motility Checkpoint Control. *Cell Reports* **18**:866-877 https://doi.org/10.1016/j.celrep.2016.12.088 | PubMed
30. **Ni B,** et al. (2020) Growth-rate dependent resource investment in bacterial motile behavior quantitatively follows potential benefit of chemotaxis. *Proceedings of the National Academy of Sciences of the United States of America* **117**:595-601 https://doi.org/10.1073/pnas.1910849117 | PubMed
31. **Yi X,** Dean AM (2016) Phenotypic plasticity as an adaptation to a functional trade-off. *eLife* **5**:e19307 https://doi.org/10.7554/eLife.19307 | PubMed
32. **Fitzgerald DM,** Bonocora RP, Wade JT (2014) Comprehensive Mapping of the Escherichia coli Flagellar Regulatory Network. *PLoS Genetics* **10** https://doi.org/10.1371/journal.pgen.1004649 | PubMed
33. **Wang P,** et al. (2010) Robust growth of escherichia coli. *Current Biology* **20**:1099-1103 https://doi.org/10.1016/j.cub.2010.04.045 | PubMed
34. **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**:1487-1487 https://doi.org/10.1039/c2lc00009a | PubMed
35. **Morimoto YV,** et al. (2010) Charged residues in the cytoplasmic loop of MotA are required for stator assembly into the bacterial flagellar motor. *Molecular Microbiology* **78**:1117-1129 https://doi.org/10.1111/j.1365-2958.2010.07391.x | PubMed
36. **Kojima S,** Blair DF (2001) Conformational Change in the Stator of the Bacterial Flagellar Motor. *Biochemistry* **40**:13041-13050 https://doi.org/10.1021/bi011263o | PubMed
37. **Davis JH,** Rubin AJ, Sauer RT (2011) Design, construction and characterization of a set of insulated bacterial promoters. *Nucleic Acids Research* **39**:1131-1141 https://doi.org/10.1093/nar/gkq810 | PubMed

38. **Kiviet DJ,** et al. (2014) Stochasticity of metabolism and growth at the single-cell level. *Nature* **514**:376-379 https://doi.org/10.1038/nature13582 | PubMed

39. **Chai Q,** et al. (2014) Organization of ribosomes and nucleoids in escherichia coli cells during growth and in quiescence. *Journal of Biological Chemistry* **289**:11342-11352 https://doi.org/10.1074/jbc.m114.557348 | PubMed

40. **Chao L,** et al. (2024) Spatial and temporal distribution of ribosomes in single cells reveals aging differences between old and new daughters of Escherichia coli. *eLife* **12**:RP89543 https://doi.org/10.7554/eLife.89543 | PubMed

41. **Pavlou A,** et al. (2025) Single-cell data reveal heterogeneity of investment in ribosomes across a bacterial population. *Nature Communications* **16**:285 https://doi.org/10.1038/s41467-024-55394-5 | PubMed

42. **Papagiannakis A,** et al. (2025) Nonequilibrium polysome dynamics promote chromosome segregation and its coupling to cell growth in Escherichia coli. *eLife* https://doi.org/10.7554/elife.104276 | PubMed

43. **van Heerden JH,** et al. (2024) Growth consequences of the inhomogeneous organization of the bacterial cytoplasm. *eLife* **13**:RP99996 https://doi.org/10.7554/eLife.99996

44. **Tanouchi Y,** et al. (2015) A noisy linear map underlies oscillations in cell size and gene expression in bacteria. *Nature* **523**:357-359 https://doi.org/10.1038/nature14562 | PubMed

45. **Shi C,** et al. (2020) Allocation of gene products to daughter cells is determined by the age of the mother in single Escherichia coli cells. *Proceedings of the Royal Society B: Biological Sciences* **287**:20200569 https://doi.org/10.1098/rspb.2020.0569 | PubMed

46. **Proenca AM,** et al. (2024) Progressive decline in old pole gene expression signal enhances phenotypic heterogeneity in bacteria. *Science Advances* **10**:eadp8784 https://doi.org/10.1126/sciadv.adp8784 | PubMed

47. **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 https://doi.org/10.1128/jb.186.22.7529-7537.2004 | PubMed

48. **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. Harvard Dataverse. https://doi.org/10.7910/DVN/T56JYW

49. **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 https://doi.org/10.1093/nar/25.6.1203 | PubMed

50. **Balleza E,** Kim JM, Cluzel P (2018) Systematic characterization of maturation time of fluorescent proteins in living cells. *Nature Methods* **15**:47-51 https://doi.org/10.1038/nmeth.4509 | PubMed

51. **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 https://doi.org/10.1128/jb.185.6.2017-2021.2003 | PubMed

52. **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 https://doi.org/10.1038/s41596-019-0216-9 | PubMed

53. **Locke JCW,** et al. (2011) Stochastic Pulse Regulation in Bacterial Stress Response. *Science* **334**:366-369 https://doi.org/10.1126/science.1208144 | PubMed

54. **Levine JH,** et al. (2012) Pulsed Feedback Defers Cellular Differentiation. *PLOS Biology* **10**:e1001252 https://doi.org/10.1371/journal.pbio.1001252 | PubMed

55. **Potvin-Trottier L,** et al. (2016) Synchronous long-term oscillations in a synthetic gene circuit. *Nature* **538**:514-517 https://doi.org/10.1038/nature19841 | PubMed

56. **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 https://doi.org/10.1038/ismej.2011.27 | PubMed

57. **Bakshi S,** et al. (2021) Tracking bacterial lineages in complex and dynamic environments with applications for growth control and persistence. *Nature Microbiology* **6**:783-791 https://doi.org/10.1038/s41564-021-00900-4 | PubMed

58. **Sassi AS,** et al. (2022) Protein Concentration Fluctuations in the High Expression Regime: Taylor's Law and Its Mechanistic Origin. *Physical Review X* **12** https://doi.org/10.1103/physrevx.12.011051 | PubMed

Peer reviews

Reviewer #1 (Public review):

Summary:

Garcia-Alcala, Kratz and Cluzel investigate to what extent our understanding of bacterial physiology in bulk experiments can be applied to single-cell observations. They find that intrinsic noise may be powerful enough to even inverse the trends found in the bulk. The authors hypothesize that asymmetric distribution of ribosomes to daughter cells during the cell division plays the dominant role in the intrinsic noise and is able to generate the observed phenomenon. They do not show it directly, but the data and its agreement with the model suffice to support this claim.

Strengths:

The experimental part is convincing: the positive correlation between the elongation rate and promoter activity of unnecessary protein is clear, as well as the negative correlation between the mean values while changing the promoter strength. This was demonstrated in both rich and poor media. The causality between the growth rate and the promoter activity was shown using the negative lag time of the cross-correlation function. A simple, reasonable model accounts well for the data. This paper demonstrates an interesting phenomenon and provides a plausible theory for it, advancing our understanding of bacterial physiology on the single-cell level.

Weaknesses:

(1) Mean-reversion timescales were assumed to be longer than the simulation time and much longer than the cell cycle time. It is not clear whether the results robust in case mean-reversion timescales become of the order of cell cycle or smaller. If not, is there an argument for such practically infinite reversion timescales?

(2) It is not easy to understand the simulation part unless one reads Ref. [14]. Is k(t) assumed to follow Eq. (1) from ref. [14]? Is this crucial that the ribosome noise appears only at the division? The ribosome noise strength \sigma_R=0.06 - is it lower or higher than the naively expected binomial division?

Also, more intuitive explanation of the Simpson paradox would help the reader.

(3) It would be useful for the reader to see the raw data and not only the filtered one to appreciate the measurement noise level.

(4) Negative lag time of the cross-correlation function is visible, but consider adding statistical test for it.

(5) Can you make similar cross-correlation plots using the model? Can you infer using it whether the data agrees better with the assumption that ribosomes noise appear only at division or continuous fluctuations during the cell cycle?

Comments on revised version:

The authors addressed the five comments listed above.

https://doi.org/10.7554/eLife.110747.2.sa2

Reviewer #2 (Public review):

Summary:

The manuscript by Garcia-Alcala et al. reports an interesting paradox: the cost of gene expression slows the population-average growth rate, whereas at the single-cell level, expression levels from these genes positively correlate with the growth rate. The effect is observed in the expression of flagellar genes and a gene under a synthetic promoter in E. coli. The findings are explained by the inheritance of growth factors, including ribosomes, during asymmetric division.

Strengths:

(1) The manuscript adds strength to an emerging body of literature showing that the population-level bacterial growth laws do not match correlations based on single-cell data. The evidence presented here is more striking than in previous works (such as Pavlou et al., Nat. Commun. 2025), as the trends in population-level data and single-cell data are reversed.

(2) A relatively simple model correctly explains the trends in the data.

Weaknesses:

(1) The differing behavior of the MG1655 and MC4100 strains remains a lingering question concerning the generality of the conclusions. It appears unlikely that ribosomes or other growth factors partition significantly differently in the MC4100 strain than in the MG1655 strain. Furthermore, based on Fig. S15, it is still unclear to what extent MC4100 exhibits growth-rate fluctuations, as stated in the text, rather than primarily size fluctuations, as shown in the figure. It is also unclear why such very slow fluctuations would lead to qualitatively different behavior, given that the proposed mechanism appears to be rather fundamental. It would be helpful for the authors to discuss these two points.

(2) It is unclear what fraction of the total proteome mVenus represents in different measurements. Adding this information would strengthen the conclusions.

https://doi.org/10.7554/eLife.110747.2.sa1

Author response:

The following is the authors' response to the original reviews.

> ***Public Reviews:***
>
> ***Reviewer #1 (Public review):***
>
> *Summary:*
>
> *Garcia-Alcala, Kratz and Cluzel investigate to what extent our understanding of bacterial physiology in bulk experiments can be applied to single-cell observations. They find that intrinsic noise may be powerful enough to even inverse the trends found in the bulk. The authors hypothesize that the asymmetric distribution of ribosomes to daughter cells during cell division plays the dominant role in the intrinsic noise and is able to generate the observed phenomenon. They do not show it directly, but the data and its agreement with the model are sufficient to support this claim.*

> *Strengths:*
>
> *The experimental part is convincing: the positive correlation between the elongation rate and promoter activity of unnecessary protein is clear, as well as the negative correlation between the mean values while changing the promoter strength. This was demonstrated in both rich and poor media. The causality between the growth rate and the promoter activity was shown using the negative lag time of the cross-correlation function. A simple, reasonable model accounts well for the data. This paper demonstrates an interesting phenomenon and provides a plausible theory for it, advancing our understanding of bacterial physiology on the single-cell level.*
>
> *Weaknesses:*
>
> *(1) Mean-reversion timescales were assumed to be longer than the simulation time and much longer than the cell cycle time. It is not clear whether the results are robust in case mean reversion timescales become of the order of the cell-cycle or smaller. If not, is there an argument for such practically infinite reversion timescales?*

Due to an error in the simulation code, the incorrect mean-reversion timescales were reported in the manuscript and have now been corrected. Instead of 1000 h and 100 h for τ_R and τ_U, respectively, they are 6.25 h and 0.0625 h, which are both much shorter than the total simulation time (75 h). The error had no effect on simulation results and was purely a timescale conversion mistake. We appreciate the referee's careful review of the manuscript which allowed us to catch this error.

Given the correct values, τ_U is well within the cell-cycle time, as is expected since we assume the source of noise in unnecessary protein expression is from stochastic gene expression. In contrast, τ_U is still significantly longer than the cell-cycle time and needs to be in order to generate the experimentally observed behavior (i.e., the increase in growth rate with an increase in unnecessary protein expression at the single-cell level). This is consistent with our biological hypothesis that some cells inherit ribosomal surpluses from their mothers which enable bursts in protein production. τ_U less than the cell-cycle time would correspond to a case where ribosomal composition quickly decays to the population average, meaning that daughter cells would never have time to capitalize on the benefit of receiving a ribosome surplus. Furthermore, τ_U being longer than the cell-cycle is biologically justifiable as proteins such as ribosomes are passed from mother to daughter at division, thus allowing for memory to persist over longer timescales than a single generation.

> *(2) It is not easy to understand the simulation part unless one reads Ref [14]. k(t) is assumed Equation (1) from Reference [14]? Is it crucial that the ribosome noise appears only at the division? The ribosome noise strength σ_R =0.06 - is it lower or higher than the naively expected binomial division? Also, a more intuitive explanation of the Simpson paradox would help the reader.*

$\kappa(t)$ is indeed Eq. (1) from Ref. [14]. To make the computational results clearer, the methods section has been updated to include the full set of equations used to perform the simulations, and the code used to produce the computational figures is now on GitHub. The relative standard deviation expected by modeling ribosome distribution at division by a binomial distribution with equal probability of being inherited by either daughter cell is in the range of 1-3% (0.01-0.03), assuming N~10^3-10^4 ribosomes. Thus, 0.06 is reasonable as it is in the same order of magnitude as what is predicted by binomial division. Furthermore, if clustering is present as already demonstrated in [39, 40, 43], we would expect the relative standard deviation to increase as clustering reduces the effective number of proteins which are distributed between daughters.

> *(3) It would be useful for the reader to see the raw data and not only the filtered one to appreciate the measurement noise level.*

We have included the direct calculations of cell size and promoter activity in the time-lapse plots in Fig. 1. In addition, we included Fig. S2, which shows typical time traces of Class-2 activity and elongation rate, displaying both the direct measurements and the corresponding smoothed traces for the flagellar reporter strains.

> *(4) Negative lag time of the cross-correlation function is visible, but consider adding a statistical test for it.*

We have included a statistical analysis of the lags of maximum correlation of elongation rate and activity for the flagellar reporter strains in Fig. S10. For each strain, we now show an overlay of the cross-correlation functions for all lineages and the distribution of the lag corresponding to the maximum correlation. In addition, we have added a section in Materials and Methods describing in detail how the cross-correlation and the strain-averaged lag were computed.

> *(5) Can you make similar cross-correlation plots using the model? Can you infer by using it, whether the data agrees better with the assumption that ribosomal noise appears only at division or continuous fluctuations during the cell cycle?*

The model in its current form is unable to capture the observed cross-correlation (the correlation is sharply peaked at zero). This is because the model coarse-grains transcription and translation into one process of protein production and thus lacks any delay or memory mechanisms which would make a significant positive or negative correlation.

Both continuous fluctuations and noise from division could in principle contribute to our observation of Simpson's paradox. We are unable to use the model in its present form to dissect the contributions from both mechanisms. However, our model simulations clearly show that noise at division by itself is sufficient to explain the observed effect with parameter values that are biologically plausible. As Chao et al., [40] has previously shown experimentally that a significant source of ribosomal noise comes from unequal distribution at division, which is the hypothesis we retained for the model.

> ***Reviewer #2 (Public review):***
>
> *Summary:*
>
> *The manuscript by Garcia-Alcala et al. reports an interesting paradox: the cost of gene expression slows the population-average growth rate, whereas at the single-cell level, expression levels from these genes positively correlate with the growth rate. The effect is observed in the expression of flagellar genes and a gene under a synthetic promoter in E. coli. The findings are explained by the inheritance of growth factors, including ribosomes, during asymmetric division.*
>
> *Strengths:*
>
> *(1) The manuscript adds strength to an emerging body of literature showing that the population-level bacterial growth laws do not match correlations based on single-cell data. The evidence presented here is more striking than in previous works (such as Pavlou et al., Nat. Commun. 2025), as the trends in population-level data and single-cell data are reversed.*
>
> *(2) A relatively simple model correctly explains the trends in the data.*
>
> *Weaknesses:*

> *(1) It is not clear whether flagellar proteins are expressed proportionally to the reporter signal. Furthermore, it is questionable if E. coli bacteria in the mother machine channels are flagellated. If they are, they could potentially swim out of the channels, which is not the case when they do not carry the MotA E98K mutation. The authors should provide some evidence that E. coli expresses the actual filament proteins in the channels.*

We agree that it is important to demonstrate that our reporter reflects the production of functional flagellar structures under our experimental conditions. To this end, we first tested the swimming capabilities of our strains before introducing the MotA E98K mutation, using a standard soft-agar motility assay. For strains in which only the Class-1 promoter was modified (Pro2, Pro4, Pro5), after 12 h of inoculation, the diameters of the swarming rings followed the expected order based on promoter strength: Pro2 showed the smallest ring, followed by Pro4, WT, and Pro5. In contrast, control strains carrying Pro4 together with either MotA E98K (non-rotating motors) or ΔfliC (no flagellin filament) did not form rings, consistent with their inability to swim. We also included an MG1655 strain carrying an IS5 insertion upstream of the Class-1 promoter, which is known to enhance flagellar expression [56]; this strain showed a larger ring, as expected. A qualitative summary of ring diameters for all strains is provided in Author response table 1. We have included the figure and section "Experimental validation of functional flagella expression in reporter strains" on Supplementary Material.

Label	Class-1 Promoter	Promoter label	E. coli Background	Observation Ring diameter
Pro2	Pro2	P1	MG1655	+
Pro4	Pro4	P2	MG1655	++
WT	WT	WT	MG1655	+++
Pro5	Pro5	P3	MG1655	++++
IS5	WT	WT	MG1655 IS5	+++++
E98K Pro4	Pro4	P2	MG1655 MotA E98K	-
ΔC Pro4	Pro4	P2	MG1655 ΔfliC	-
WT E98K	WT	WT	MG1655 MotA E98K	-

Author response table 1.

We also tested whether cells assemble functional flagella on the mother-machine. We compared MG1655 WT and MG1655 carrying the MotA E98K mutation under identical microfluidic growth conditions. Many WT cells left the channels during the experiment (~35% of lineages over ~20 h after the onset of exponential growth inside the device), consistent with active swimming, whereas the non-motile MotA E98K strain did not leave the channels. Because the only difference between these two strains is the MotA E98K mutation, which disables motor rotation but not flagellar assembly, this result indicates that (i) cells do express functional filaments and motors in the mother-machine environment, and (ii) the strain used for our main experiments is immobilized by MotA E98K.

Both experiments are included in the Supplementary Information section "Experimental validation of functional flagella expression in reporter strains" and Fig. S3.

> *(2) It is unclear what fraction of the total proteome mVenus represents in different measurements. Some quantification is needed (for example, using the Coomassie staining). Using f_U as high as 14.4% in simulations is questionable.*

We agree that we do not currently know the exact fraction of the proteome occupied by mVenus in our experiments, as we only quantified fluorescence and did not perform Coomassie staining or proteomics. The primary goal of our simulations is not to reproduce exact experimental conditions, but to illustrate how unequal ribosome partitioning affects

daughter cells across a range of protein synthesis burdens. Consequently, our use of values up to F_U = 14.4% in the simulations was intended as an exploratory upper range rather than as a direct estimate of the experimental condition.

To put our experimental burden in context, we compare our growth-rate reduction to a well–characterized high-burden case in the literature. In the study by T. Hwa's group [6], overexpression of β–galactosidase such that it constituted approximately 27% of the total proteome led to a 67% reduction in growth rate for *E. coli* growing in a medium similar to ours (differing only in the carbon source: glucose in their case, glycerol in ours). In our system, overexpression of mVenus from a plasmid result in a substantially smaller growth–rate reduction of 9% relative to the non–expressing control, and this is observed on the poorer carbon source (glycerol), under which burden effects are typically less pronounced [6].

Although we cannot convert our fluorescence measurements into an exact proteome fraction, the much smaller growth defect compared to the 27% β-galactosidase case strongly suggests that mVenus does not approach such an extreme fraction of the proteome under our conditions. Under the simplifying assumption that the qualitative relationship between unnecessary-protein fraction and growth-rate reduction is similar in the two systems, our data are therefore consistent with a modest fraction of unnecessary protein and make it unlikely that mVenus reaches very high fractions such as 27%. This gives us confidence that exploring F_U values up to 14.4% in the simulations represents a conservative upper range relative to our experimental burden, rather than an underestimate.

> *(3) The data from the MC4100 strain does not directly match the trends of MG1655. The justification for filtering out the low-frequency components of MC4100 is not particularly convincing. It appears unlikely that ribosomes or other growth factors partition significantly differently in the MC4100 strain than in the MG1655 strain. Further discussion and a plot similar to Figure 1 (Left) for this strain are warranted.*

We thank the reviewer for this helpful suggestion. We agree that the filtering analysis from the initial manuscript was not clear enough to be used as robust supplementary information, and we therefore removed it. Instead, we carried out with the 'unfiltered' MC4100 strain, the same single-cell analyses as with MG1655, including the binning analysis of instantaneous elongation rate versus Class-2 promoter activity, and a cross-correlation analysis between such measurements.

Both analyses did not reveal a significant correlation between growth and flagellar promoter activity in MC4100. We now present these results in the revised manuscript (Fig. S15), where we explicitly show the lack of association in a plot directly comparable to Fig. 1.

While ribosomes partition mechanism is likely to be the same between these two strains, MC4100 is known to exhibit very long oscillations of growth rate over 10 generations, which are absent in MG1655.

We now present MC4100 as an explicit counterexample to highlight that the positive correlation between short timescale growth fluctuations and flagellar expression observed in MG1655 is not universal across all *E. coli* strains, especially in strains like MC4100 whose growth rate fluctuations are dominated by long timescales, much longer than the division time. In MC4100 these slow modes are largely decoupled from flagellar gene regulation. We have also revised the text to clarify this point.

> *(4) The model needs to be described in more detail. A closed set of equations that have been simulated must be presented, along with all values of the model parameters and their sources. The authors should consider depositing their code on GitHub or another publicly accessible repository.*

The methods section has now been updated to include the full set of equations used to perform the simulations along with all parameter values and their sources. Additionally, the code used to produce the computational figures is now on GitHub. For a full derivation and biological justification of each model component, we still refer readers to ref [14] where this model was first published.

> ***Recommendations for the authors:***
>
> ***Reviewer #1 (Recommendations for the authors):***
>
> *(1) The first paragraph of Results still belongs to the Introduction Section, I feel.*

Thank you for the recommendation, we agreed with the change.

> *(2) Figure 1, left, is after some filter (Savitzky-Golay)? It might be useful to see the raw points.*

We have included the direct calculations of cell size and promoter activity in the time-lapse plots in Fig. 1. In addition, we included Fig. S2, which shows typical time traces of Class-2 activity and elongation rate, displaying both the raw (direct) measurements and the corresponding smoothed traces for the flagellar reporter strains.

> ***Reviewer #2 (Recommendations for the authors):***
>
> *(1) Is there a statistically significant positive slope in single-cell data of the elongation rate as a function of Class-2 activity? There should be an analysis of statistical significance for the slopes in Figure 1 (Right) and Figure 4A, including both population-average and single-cell data.*

Thank you for your recommendation. We have now added statistical analyses of the correlations between activity and elongation rate at both the single-cell and population levels. The results for the flagellar strains are presented in Fig. S5, S6, and S7. Also, the corresponding analysis for constitutive Venus expression is shown in Fig. S16.

> *(2) How does FlhC-YFP data compare with the CFP signal in the first set of measurements? What would Figure 1, Left look like for this signal?*

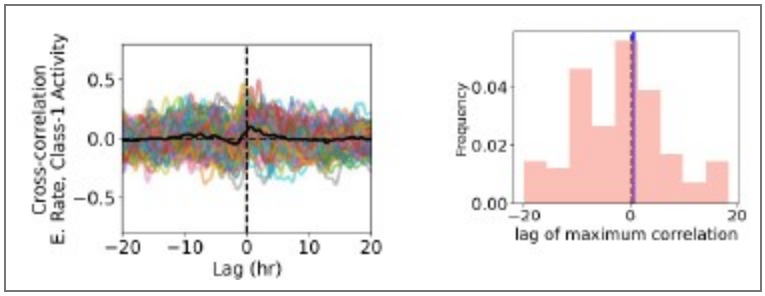
At the single-cell level, the relationship between Class-1 promoter activity and elongation rate within each strain is still positive, but clearly weaker than for Class-2, as shown in Fig. S8. When we bin the data, we can observe that the binned averages don't display a clear positive trend, even though the Pearson correlation for each flagellar reporter strain is still positive. We think this is because Class-1 controls a much smaller part of the proteome (it only encodes the two subunits of FlhDC) while Class-2 promoters drive many structural and export proteins. So, changes in Class-2 activity more directly reflect shifts in global translational capacity and are more tightly linked to growth, whereas Class-1 activity adds only a small translational load.

> *(3) The information from the ER-Activity cross-correlation functions is interesting but has not been interpreted or compared with the model. Which signal precedes the other? What can explain the observed lag time on the order of Tdiv? Why do some cross-correlations show negative values in Fig. S9 while others are positive (as they should)? Can the model explain the experimentally observed cross-correlation function?*

In our cross-correlation analysis, a negative lag at the maximum means that fluctuations in elongation rate (ER) precede fluctuations in promoter activity (A) by that lag. We now describe this explicitly in Materials and Methods and quantify lag distributions for all reporter strains in Fig. S10.

In Fig. 13 (previously Fig. 9), we mainly observe positive correlations between ER and PA with negative or near zero peak lags, i.e., ER tends to lead PA by a lag by about a division time. The reason governing the negative lags is not immediately clear, but it is consistent with a resource driven mechanism: cells that inherit more growth factors (e.g. ribosomes) at division, use them to prioritize housekeeping processes and thus increase growth rate first, and only subsequently use these extra resources to increase flagellar promoter activity.

As for Class–1 activity, the correlation with growth is much weaker as it was already observed in Kim et al. Averaging the cross–correlations across all lineages yields a modest peak (mean correlation 0.10, SD 0.12; see Author response image 1) at a small positive lag of 0.5 h (while average division time is T_{div} ~ 1.7h). This indicates a weak but real positive correlation between Class 1 activity and ER at short lags. However, the lags of the individual maxima are widely distributed (SD ≈ 9 h; mean −1.8 h, median −0.28 h, mode ≈ 0), with only ~53% of the cells showing a negative time lag. Thus, delays are roughly symmetrically spread around zero with only a slight negative bias.



Author response image 1. Left: cross–correlation between elongation rate and Class–1 activity for each lineage (N = 96, colored lines), and their average as a function of lag (black line). Right: distribution of the lags at which each lineage's cross–correlation attains its maximum. The dashed line indicates zero lag, and the full line marks the lag of the peak of the mean cross correlation.

An in-depth analysis about the sign and magnitude of the lag would require measurements from a broader set of promoters. The lag is likely to depend on what genes the promoter controls (e.g. stress response, housekeeping, or large structural modules), on its strength and regulation, and on growth conditions.

The model in its current form is unable to capture the observed cross-correlation (the correlation is sharply peaked at zero). This is because the model coarse-grains transcription and translation into one single process of protein production and thus lacks any delay or memory mechanism that could generate a phase shift.

Nonetheless, this memory-free formulation shows that stochastic redistribution of growth factors at division is by itself sufficient to generate the Simpson's paradox behavior. Capturing the experimentally observed lag time would require including explicit transcription/translation delays and additional regulatory dynamics between housekeeping and flagellar genes whose information we do not have.

> *(4) Figure 3, Left - it is not clear what this plot shows. Red and blue are scattered over the whole plot. How have daughter 1 and daughter 2 been assigned? Perhaps choosing one of the daughters with a higher growth rate and then plotting the data could reveal some trends.*

We agree that the original left panel of Fig. 3 was difficult to follow. In the original version, the daughter labels were assigned as "Daughter–1" for the cell at the closed end of the channel and "Daughter–2" for the cell closest to the open side. After discussing this with Dr

Camilla Ulla Rang (whom we now acknowledge in the manuscript), we relabeled the daughters as “new–pole daughter” and “old–pole daughter,” following the convention used in studies of aging and ribosome distribution in *E. coli*.

We now show the class–2 promoter activity comparison between new pole vs old pole, and in a separate panel, we plot the mean ratios of elongation rate, and class–1/class–2 promoter activities between the new–pole and old–pole daughters. These ratios show that the new–pole daughter tends to have both greater growth rate and flagellar gene activity than the old–pole daughter. Importantly, this new plot is in line with Rang and colleagues who demonstrated that new–pole daughters have greater ribosome density and faster growth rates than their old–pole sisters. Together, these results further support our hypothesis that excesses of ribosomes inherited at division underlies the observed growth boosts.

> *(5) Flagellar activity -> activity of flagellar gene synthesis (presumably no flagellar activity in these cells).*

We corrected the terms used to reference the flagellar gene activity.

> *(6) Page 7: "The strain MC4100, known to exhibit slow, long period oscillations in growth ..." - some reference is- needed here.*

We placed the reference some lines after, as such paper also includes the information of MG1655 short-term oscillations. Now the reference is [44]: Tanouchi, Y., et al., A noisy linear map underlies oscillations in cell size and gene expression in bacteria. Nature, 2015

> *(7) Page 7: Figure S11B - is Figure S11C perhaps meant?*

The correct panel indeed was Fig. S11C, and we have now corrected and updated the figure label and text accordingly.

> *(8) Page 11: Savitsky-Golay filter.*

We have corrected it, thank you.

https://doi.org/10.7554/eLife.110747.2.sa0