Please cite this article in press as: Taheri-Araghi et al., Cell-Size Control and Homeostasis in Bacteria, Current Biology (2015), http://dx.doi.org/10.1016/j.cub.2014.12.009

Current Biology 25, 1–7, February 2, 2015 ©2015 Elsevier Ltd All rights reserved http://dx.doi.org/10.1016/j.cub.2014.12.009

Report

Cell-Size Control and Homeostasis in Bacteria

Sattar Taheri-Araghi,[1,7] Serena Bradde,[2,7] John T. Sauls,[1] Norbert S. Hill,[3] Petra Anne Levin,[4] Johan Paulsson,[5] Massimo Vergassola,[1,*] and Suckjoon Jun[1,6,*]
[1]Department of Physics, University of California San Diego, La Jolla, CA 92093, USA
[2]Initiative for the Theoretical Sciences, The Graduate Center, City University of New York, 365 Fifth Avenue, New York, NY 10016, USA
[3]Department of Molecular & Cell Biology, University of California, Berkeley, CA 94720, USA
[4]Department of Biology, Washington University, Saint Louis, MO 63130, USA
[5]Department of Systems Biology, Harvard Medical School, Longwood, MA 02115, USA
[6]Section of Molecular Biology, Division of Biological Science, University of California San Diego, La Jolla, CA 92093, USA

Summary

How cells control their size and maintain size homeostasis is a fundamental open question. Cell-size homeostasis has been discussed in the context of two major paradigms: "sizer," in which the cell actively monitors its size and triggers the cell cycle once it reaches a critical size, and "timer," in which the cell attempts to grow for a specific amount of time before division. These paradigms, in conjunction with the "growth law" [1] and the quantitative bacterial cell-cycle model [2], inspired numerous theoretical models [3–9] and experimental investigations, from growth [10, 11] to cell cycle and size control [12–15]. However, experimental evidence involved difficult-to-verify assumptions or population-averaged data, which allowed different interpretations [1–5, 16–20] or limited conclusions [4–9]. In particular, population-averaged data and correlations are inconclusive as the averaging process masks causal effects at the cellular level. In this work, we extended a microfluidic "mother machine" [21] and monitored hundreds of thousands of Gram-negative *Escherichia coli* and Gram-positive *Bacillus subtilis* cells under a wide range of steady-state growth conditions. Our combined experimental results and quantitative analysis demonstrate that cells add a constant volume each generation, irrespective of their newborn sizes, conclusively supporting the so-called constant Δ model. This model was introduced for *E. coli* [6, 7] and recently revisited [9], but experimental evidence was limited to correlations. This "adder" principle quantitatively explains experimental data at both the population and single-cell levels, including the origin and the hierarchy of variability in the size-control mechanisms and how cells maintain size homeostasis.

[7]Co-first author
*Correspondence: massimo@physics.ucsd.edu (M.V.), suckjoon.jun@gmail.com (S.J.)

Results

At the Population Level, New Experimental Data Confirm the Growth Law

Population-level parameters derived from our single-cell data followed established patterns for microbial growth known as the growth law [1]: the average newborn cell volume $\langle v_b \rangle$ increased and the average generation time $\langle \tau_d \rangle$ decreased, respectively, as the nutrient-imposed growth rate $\langle \lambda \rangle = \langle 1/\tau_d \rangle \ln 2$ increased (newborn refers to the cells right after birth; Figure 1A). The newborn cell volume depended exponentially on the nutrient-imposed growth rate (hereafter referred to as growth rate, unless otherwise noted), $\langle v_b \rangle = A \exp(B\langle \lambda \rangle)$, in quantitative agreement with the growth law [1] (Figure 1C, red symbols and line; A is the y intercept, and B is the slope of the red line). Moreover, newborn length $\langle s_b \rangle$ and width $\langle w_b \rangle$, averaged over the entire set of individual cells in each growth condition, also showed an exponential dependence on the average growth rate $\langle \lambda \rangle$ (Figure S1A available online).

The size of individual cells also increased exponentially as $s(t) = s_b 2^{\alpha t}$ (where α is the instantaneous elongation rate), and… their width did not change significantly between birth and division (Figure S1B; [21]; hereafter, we use size and volume synonymously). The average instantaneous elongation rate was identical to the average growth rate of the population since $\langle 1/s\ ds/dt \rangle = \langle \alpha \rangle \ln 2 = \langle 1/\tau_d \rangle \ln 2 = \langle \lambda \rangle$.

At the Single-Cell Level, Individual Cells Show Systematic Deviations from the Growth Law

Individual cells, however, exhibited intrinsic variability even under constant growth conditions, and we asked whether the quantitative relationship between the average size and the average growth rate also applied at the single-cell level. For example, the SDs of the growth rate and the newborn cell size were ~15% and ~14% of their respective means (Figure 1B). Therefore, when the growth-rate distributions for two different growth conditions partially overlapped as shown in Figure 1B, individual cells in the overlap region could have had the same growth rate $\lambda = (\ln 2)/\tau_d$. Thus, if the growth rate solely defined the cell's growth physiology, individual cells with the same λ should have had on average the same size as described by the growth law $\langle v_b \rangle = A \exp(B\langle \lambda \rangle)$. We found this was not the case. For all seven growth conditions, the size versus growth rate measured from individual cells, v_b versus λ, systematically deviated from the population-level growth law (Figure 1C, blue symbols and lines versus red symbols and line). This deviation indicates that, at the single-cell level, the size of individual cells is controlled by a mechanism that is different from the growth law $\langle v_b \rangle = A \exp(B\langle \lambda \rangle)$ (see below).

Correlations of Growth and Size Parameters Contradict Both Sizer and Timer Models

The newborn cell size (s_b) and the generation time (τ_d) of individual cells were negatively correlated (Figure 1D, left), which excluded the timer model of cell-size control. Otherwise, we would have seen constant τ_d with respect to s_b. Furthermore, timer models showed instability when accounting for the

Please cite this article in press as: Taheri-Araghi et al., Cell-Size Control and Homeostasis in Bacteria, Current Biology (2015), http://dx.doi.org/10.1016/j.cub.2014.12.009

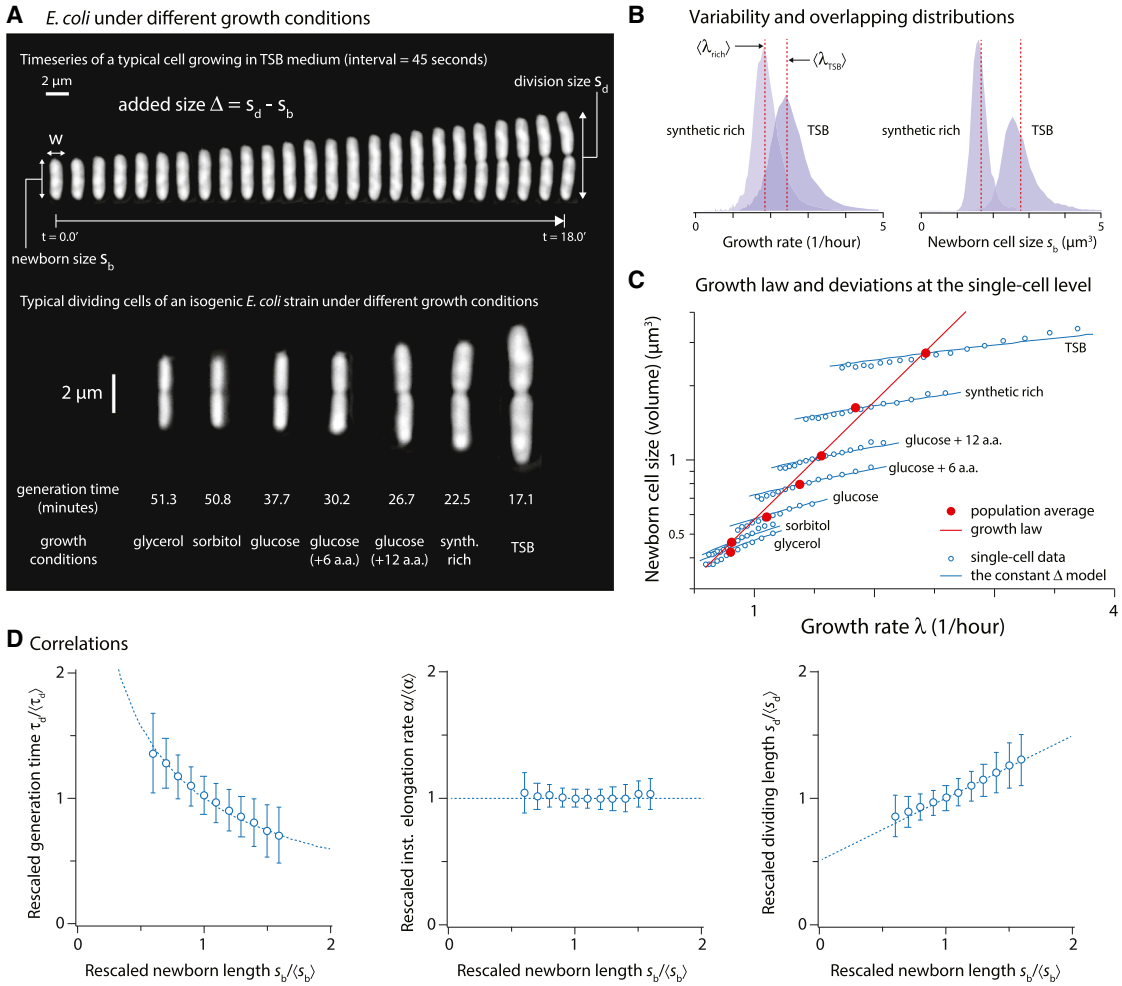


Figure 1. Growth Law at the Population Level and Systematic Deviations at the Single-Cell Level

(A) Top: time series of a typical cell growing in a nutrient-rich medium. Bottom: sample images of dividing *E. coli* cells in steady-state exponential growth at 37°C in seven different growth media.

(B) Partially overlapping distributions of the growth rate and the newborn size measured from individual cells in two different growth conditions. The vertical lines show the population average values. Cells in the overlap region can have the same growth rate or newborn cell size.

(C) Population average of single-cell measurements demonstrates exponential dependence of newborn cell volume on the average growth rate (red). However, s_b versus λ of individual cells (binned data in empty blue circles; measured by following them from birth to division) shows systematic deviations from the average growth law. Thus, although the cells in the overlap region in (B) can have the same growth rate or newborn cell size, the size of individual cells are controlled by a mechanism that is different from the growth law. Otherwise, all blue symbols would have fallen on top of the red line.

(D) Correlations between rescaled growth parameters at the single-cell level with SDs from the entire set of *E. coli* data. Left: generation time versus size at birth. Middle: elongation rate versus size at birth. Right: size at division versus size at birth. Dashed lines indicate predictions from the adder principle from this work. The first correlation falsifies the timer model, whereas the last correlation falsifies the sizer model.

See also Figure S1.

observed exponential growth of individual cells (Supplemental Information). The fact that cells born small take on average more time before they divide is in principle consistent with a sizer model. However, the strong positive correlations between the dividing size s_d and s_b (Figure 1D, right) ruled out the model because the sizer predicted that s_d should be constant.

Cells Instead Employ "Adder" Principle

Our data instead support a model in which the size added between birth and division ($\Delta = s_d - s_b$) is constant for given growth conditions. We found that, although Δ varied significantly between growth conditions and also between individual cells, Δ was on average constant irrespective of the newborn size s_b in each growth condition (Supplemental Information). In fact, the entire conditional distribution $\rho(\Delta|s_b)$ had the same shape as the nonconditional distribution $\rho(\Delta)$, and distributions of Δ from different experimental conditions collapsed onto a single curve when rescaled by their mean (Figure 2, right; Figure S2). The distribution of the size added in each generation, Δ, was thus independent of the newborn cell size.

We also confirmed the constancy of Δ in two additional *E. coli* strains from our previous work (K12 MG1655 and B/r) [21] (Figure S3) and *E. coli* size mutants (Δpgm and ftsA*) [16]. Furthermore, we also confirmed the validity of the model in the Gram-positive *B. subtilis* (Figures 2B and 2C).

The collapse of the conditional distributions in Figure 2 established the constant Δ model, or adder (as opposed to

Please cite this article in press as: Taheri-Araghi et al., Cell-Size Control and Homeostasis in Bacteria, Current Biology (2015), http://dx.doi.org/10.1016/j.cub.2014.12.009

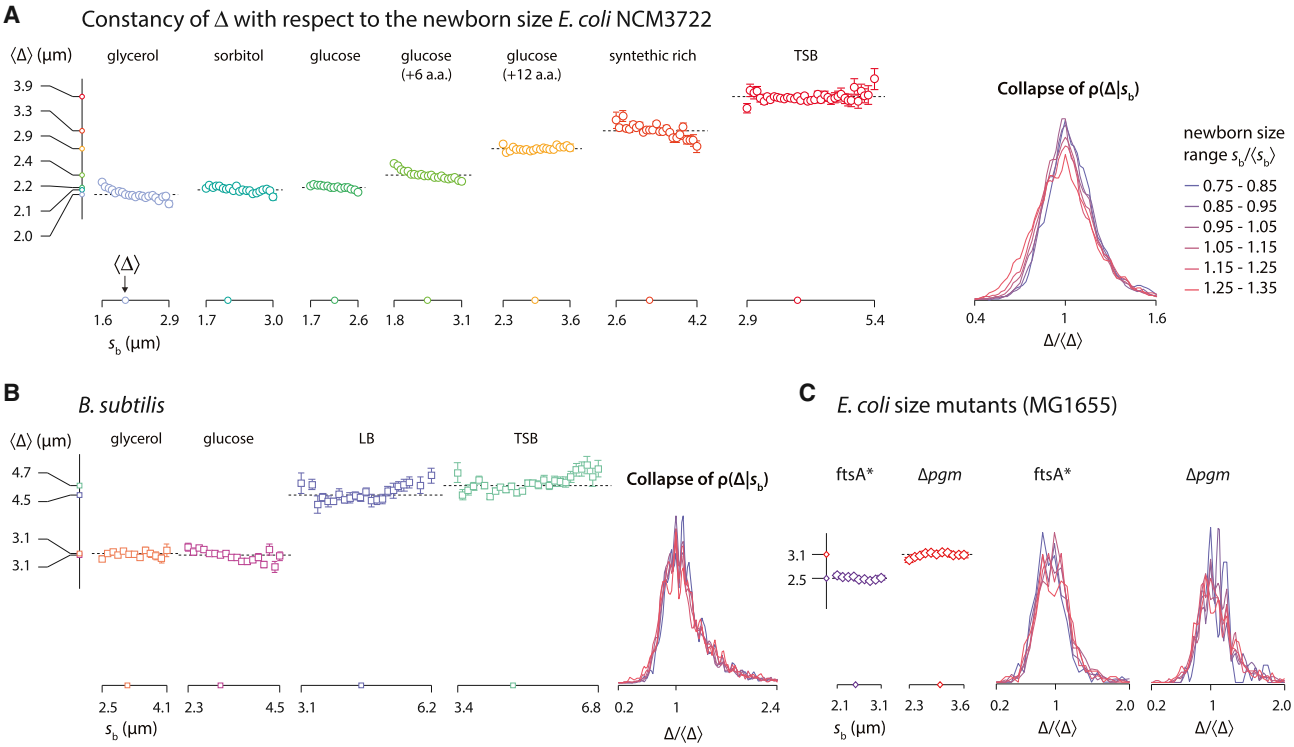


Figure 2. Experimental Evidence of Constancy of Δ in Bacteria

(A) *E. coli*: average Δ with respect to the newborn size s_b, with each bin containing >10^3 cells.

(B and C) *B. subtilis* (B) and *E. coli* size mutants (C). All rescaled distributions conditional to different newborn size ranges collapse onto one another, demonstrating that *E. coli* and *B. subtilis* cells grow by a constant size for division, independent of the newborn cell size.

See also Figures S2 and S3.

"timer" or "sizer"). Next, we explain quantitatively consequences of this adder principle on cell-size homeostasis.

Adder Ensures Size Homeostasis

An immediate consequence of addition of constant Δ is that it automatically ensured size homeostasis because at every cell division, the cell approached (albeit passively) the population average as illustrated in Figure 3A (data depict the average behavior in all growth conditions). If a cell born at size $s_b = \langle s_b \rangle + \delta s_b$ stochastically added an uncorrelated size Δ and divided in the middle with some precision, then the daughter sizes on average were $\langle s_b \rangle + \delta s_b/2$. After *n* consecutive divisions, the original size deviation of the newborn cell on average decreased as $\delta s_b/2^n$ (Figure 3A). The size homeostasis principle was confirmed by our data for both *E. coli* and *B. subtilis* (Figures 3B and 3C).

Addition of Constant Size and Exponential Elongation Explain Correlations

The constant Δ model predicted that autocorrelations of s_b, s_d, and τ_d decayed by a factor of two in each generation and that the correlation coefficient between the generation time of the mother and its daughters was −1/4, which was also confirmed by the data (Figure S4). Intuitively, the negative correlation reflects the increased generation time of the daughter cells that were born smaller than s_b due to stochastic, premature division of the mother cell [4]. Since all cells elongated exponentially with the elongation rate proportional to the cell length, cells born at $s_b < \langle s_b \rangle$ would require more time to elongate by Δ for division than cells born at $s_b > \langle s_b \rangle$ (Figure 1D, left, dashed line).

Distributions of the Growth and Division Parameters Collapse when Rescaled by Their Respective Means

The constant Δ model in fact provides a quantitative explanation for the distributions of quantities involved in growth and size control. The six distributions of the relative septum position $s_{1/2}$, elongation rate α, division size s_d, newborn size s_b, generation time τ_d, and size increment Δ are shown in Figure 4A. The coefficients of variation (CVs) of four distributions are related in the Δ model as

$$\left(\frac{\sigma_\Delta}{\langle\Delta\rangle}\right)^2 \approx \frac{1}{\ln 2}\left(\frac{\sigma_\tau}{\langle\tau\rangle}\right)^2 \approx 3\left(\frac{\sigma_b}{\langle s_b\rangle}\right)^2 \gtrsim 3\left(\frac{\sigma_d}{\langle s_d\rangle}\right)^2, \quad \text{(Equation 1)}$$

where σ denotes the SD of the distribution (see theory section in Supplemental Information for details). This predicted hierarchy of variability was confirmed by our data for both *E. coli* and *B. subtilis* (Figure 4B). Note that the size at birth s_b was slightly more variable than the size at division s_d because of the small variability of the septum position $s_{1/2}$. The elongation rate α was subject to its own physiological control and variability and showed negligible correlations with the distributions determined by Δ (Supplemental Information).

The constancy of Δ was finally supported by the scale invariance of the distributions shown in Figure 4A. In the constant Δ model, the average of the three size variables are related as $\langle\Delta\rangle = \langle s_b\rangle = \langle s_d\rangle/2$ and, if ρ(Δ) shows scale invariance, the three distributions ρ(s_b), ρ(s_d) and ρ(τ_d) also inherit the property of scale invariance of ρ(Δ) (theory section in Supplemental Information). In support of our theoretical prediction, all experimental ρ(Δ) and other size distributions collapsed onto each

Please cite this article in press as: Taheri-Araghi et al., Cell-Size Control and Homeostasis in Bacteria, Current Biology (2015), http://dx.doi.org/10.1016/j.cub.2014.12.009

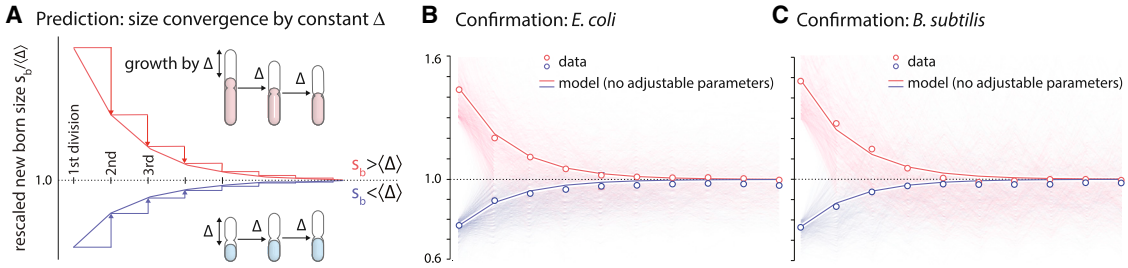


Figure 3. Mechanism of Size Homeostasis Following the Adder Principle

(A–C) For all newborn cells regardless of their size, if the cells always add a constant Δ and divide in the middle, their respective newborn size automatically converges to Δ (A). If Δ is subject to fluctuations without correlations from one generation to the next, and the cell divides in the middle with some precision, the newborn size on average still converges to $\langle\Delta\rangle$. Our data confirm this size homeostasis mechanism for both *E. coli* (B) and *B. subtilis* (C). Data in (B) and (C) show the average from all growth conditions used for each organism.
See also Movie S1.

other (Figure 4A; [23, 24]). Hence, the variation of all the statistics with growth conditions is determined by the unique parameter $\langle\Delta\rangle$.

Discussion

Proteome and Biological Origin of Constancy of Added Size

Since the proteome is a good proxy for cell size, the constant Δ is consistent with the "structural models" discussed by Fantes et al. [22]. Key features of the structural models include the following: (1) individual cells elongate exponentially, (2) initiators of cell cycle are produced at the same rate as the cell elongation rate, and (3) accumulation of the initiators to a threshold triggers the cell cycle [22]. Since the cellular volume and total number of proteins increase with the growth rate, the cellular fraction of protein initiators should reduce to maintain the constancy of the threshold. In a recent work by Scott et al. [11], the bacterial proteome is partitioned into three "sectors": R, containing ribosomal proteins; Q, containing housekeeping proteins; and P, containing the rest of the proteins. Using proteome data for the relative fraction φ_P of the P-sector proteins in *E. coli* (Figure 4E, left; [11]) and the respective average cell volume $\langle V\rangle$ (Figure 1C, red line), we found that the total number of P-sector proteins per cell $N_P = \varphi_P \times \langle V\rangle$ is relatively constant in all growth conditions for different *E. coli* strains (see Figure 4 and Supplemental Information). Thus, proteins in the P sectors behave as the initiators postulated in [22]. This leads to the prediction that the majority of proteins involved in metabolism (e.g., nutrient transporters and metabolic sensors [15]) and the cell cycle should belong to the P sector of the bacterial proteome (with their constant basal level to the Q sector). Note that the total proteome per cell increases exponentially with respect to the average growth rate; the growth law ([5]; Figure 1C) can thus be interpreted as a response of the average cell size (total proteome per cell) to nutrient conditions such that the average P-sector proteins per cell is approximately constant with respect to the nutrient-imposed growth rate. There is a clear experimental avenue for the future that will investigate how Δ will change when the proteome composition is perturbed by, e.g., transcription or translational inhibitors.

Extension to Other Organisms

The growing number of modern single-cell data sets provides a unique opportunity to determine the applicability of our findings to other bacteria as well as to eukaryotes. Analysis of bacteria, such as *Caulobacter* [25, 26], and single-celled eukaryotes should illuminate the role played by programmed degradation of regulatory proteins in cell-size homeostasis. Fantes [27] considered structural models for fission yeast *S. pombe* and dismissed them based on existing data sets. While differences might indeed be expected between eukaryotes and bacteria, extensive modern single-cell data sets are now available in, e.g., budding yeast [28], and could be used to address the question [26]. It will also be of great interest to determine whether other non-rod-shaped organisms, particularly those that exhibit tip growth and/or nonuniform morphologies, including mycobacteria, hyphal fungi, and protists like *Stentor*, also add constant volume or maintain their size through other independent mechanisms. We finally remark that the size and the shape of cells play a major role in their physiology in multicellular organisms as well, namely during *Xenopus* embryogenesis [29].

Hierarchy of Growth Parameters and the Meaning of Biological Noise

We showed that only two parameters, the elongation rate α and the added size Δ, are sufficient to reproduce the distributions of all growth and division parameters of both *E. coli* and *B. subtilis* in all growth conditions without any adjustable parameters (Equation 1 and Figure 4C; Supplemental Information). We thus propose that α and Δ represent two basic controls of physiology and size homeostasis and that the size at birth and division, as well as generation time, are slaved to them.

Ordering the variances of the rescaled distributions, the distribution of the septum position $s_{1/2}$ is the smallest, and the added size Δ is the largest (Figure 4A). Previously, sizer was supported because the coefficient of variance for division size (10%) was smaller than that for generation time (40%–60%) [19]. Therefore, interpreting coefficient of variance as a biological "noise" should be taken with caution since Δ is a basic control parameter for size homeostasis, yet Δ shows the largest variability.

Conclusions

We demonstrated that both *E. coli* and *B. subtilis* maintain cell-size homeostasis by adding a constant size Δ. The constant Δ model quantitatively explains the distributions of growth-related parameters and their variability. How bacteria can overlap their cell cycles without making fatal mistakes in the absence of eukaryotic-like cell-cycle checkpoints is a long-standing open question [30, 31]. Our results provide a new

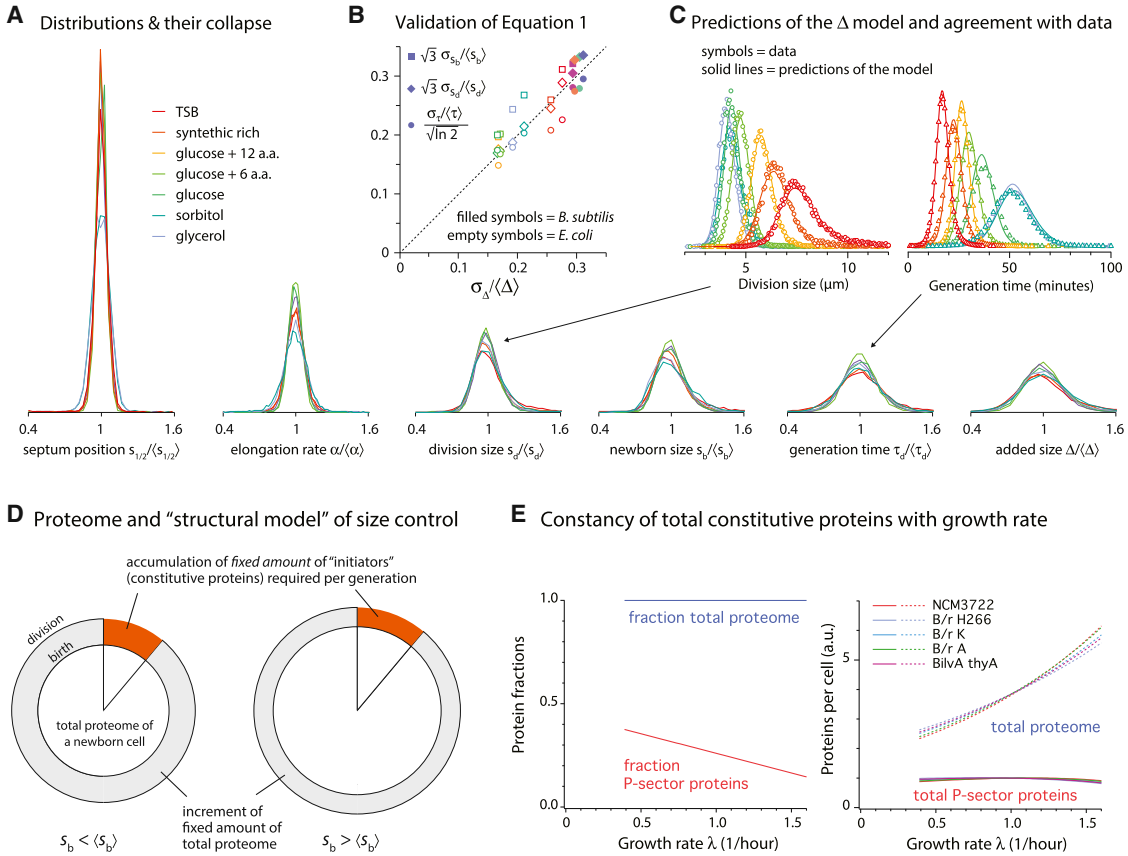


Figure 4. Origin and Quantitative Consequences of Constancy of Added Size Δ

(A) Six distributions are shown in the ascending order of their relative widths. All growth parameters from different growth conditions show scale invariance, i.e., collapse when rescaled by their respective means.

(B and C) Among the six distributions in (A), four distributions are determined by Δ (division size s_d, newborn size s_b, generation time τ_d, and Δ) (B). See Equation 1. Thus, $\rho(\Delta)$ and $\rho(\alpha)$ are sufficient to reproduce all distributions for all growth conditions for both *E. coli* (C) and *B. subtilis* (Supplemental Information) without any adjustable parameters.

(D and E) Constant Δ is consistent with the "structural models" discussed in [22], which assume that the cell grows to accumulate fixed amounts of cell-cycle regulators in each generation. Since metabolism and cell-cycle proteins are neither housekeeping nor ribosomal proteins, this prediction can be quantitatively tested using the proteome data [11] and the growth law in Figure 1C. Indeed, the total P-sector proteins per cell is constant in all growth conditions (E). See also Figure S4 and Table S3.

perspective on this issue and in the search for the underlying molecular mechanisms. A direction to be pursued in the future is the constancy of the added size Δ and its relationship with the proteome [11]. That hints at an ensemble of molecular players and entails both exponential dependency of the average cell size on growth rate (the growth law) and constancy of Δ at steady state. It will thus be important to interfere with protein synthesis and assess the resulting effects on the cell-size distributions.

Experimental Procedures

Strains

For physiological study, it is important to use a prototrophic strain. For *E. coli*, we chose the strain K12 NCM3722, constructed, sequenced, and extensively tested by Sydney Kustu's laboratory [32]. We used SJ202, a nonmotile derivative of NCM3722 (ΔmotA). For *B. subtilis* experiments, we chose a strain in the 3610 background with ComI (Q12L) mutation to allow competence. We used a derivative with reduced motility and biofilm formation by deleting *epsH* and a flagellin protein hag, respectively.

Growth Media

E. coli growth experiments were performed in seven different nutrient conditions. The average generation time in these conditions evenly spanned from 17.1 to 51.4 min at 37°C. The growth medium is based on MOPS, developed by Fred Neidhardt [33], and is commercially available from Teknova (http://www.teknova.com). *B. subtilis* growth experiments were performed in four different growth conditions with average doubling times between 16.9 and 38.9 min. The details of the growth media are listed in Tables S1 and S2. Prior to growth of the cells in the microfluidics device, all cultures were grown in a 37°C water bath shaker, shaking at 240 rpm.

Sample Preparations

All experimental steps—from inoculation to imaging—were performed at 37°C ± 0.1°C. To this end, all equipment was stationed in a 5′ × 7′ environmental chamber to eliminate any side effect of temperature fluctuations in the cell growth and physiology. Within the chamber, the temperature distribution was homogeneous, with forced air circulation within ±0.1°C, and constantly monitored at multiple locations. See Supplemental Information for more information.

Microscopy

Image acquisition and analysis were performed with an inverted microscope (Nikon Ti-E) equipped with Perfect Focus (PFS 3), a 100× oil immersion objective lens (NA 1.45), and white LED transmission light (TLED, Sutter Instruments, 400–700 nm), and an Andor NEO sCMOS camera was used for phase-contrast imaging. The illumination condition was 50 ms exposure with illumination intensity set at 10% of the maximum TLED intensity. The frequency of the time-lapse imaging was chosen such that about 20 or

more images were taken per generation time. Imaging in phase contrast eliminated potential artifacts common in fluorescence imaging. Analysis of the large number of phase-contrast images required development of custom high-throughput image analysis software as described in Supplemental Information.

Model for the Δ Control

We denote by s the cell size along the elongating axis of the rod and by s_b and s_d the size of cells at birth and division. We assume the width of the cell is roughly constant. If s(t) is the size of a cell at the current time t, its added size is denoted $\Delta(t) = s(t) - s_b$. The Δ model posits that the mechanism of control involves the single variable, $\Delta = s_d - s_b$, the size added between birth and division. The density of cells n(s, Δ′, t) having size s and added size Δ′, with g(s) = ds/dt, obeys the continuity equations

$$\partial_t n(s,\Delta',t) + \partial_s\left[g(s)n(s,\Delta',t)\right] + \partial_{\Delta'}\left[g(s)n(s,\Delta',t)\right] = -\gamma(\Delta')g(s)n(s,\Delta',t); \quad \text{(Equation 2)}$$

$$g(s)n(s,0,t) = 4g(2s)\int_0^\infty \gamma(x)n(2s,x,t)dx. \quad \text{(Equation 3)}$$

The left-hand side in Equation 2 is the total time derivative, and the two drift terms are due to the elongation of the cells, i.e., ds/dt = g(s) and dΔ(t)/dt = g(s). The right-hand side accounts for the division of cells. The Poissonian splitting rate function γ(Δ) is related to the distribution $\rho_{\Delta d}(\Delta)$ for the size added at division of individual cells as $\rho(\Delta) = \gamma(\Delta)\exp(-\int\gamma(x)dx)$. Indeed, the exponential term is the probability that the cell will not divide up to Δ and γ(Δ)dΔ is the probability of division in the range (Δ, Δ + dΔ). Simple algebra leads then to

$$\gamma(\Delta) = \frac{\rho_{\Delta d}(\Delta)}{1 - \int_0^\Delta dx\rho_{\Delta d}(x)}. \quad \text{(Equation 4)}$$

The conversion of the rate of division to unit time involves the Jacobian |dΔ(t)/dt| = g(s) that appears in the right-hand side of Equation 2. Finally, Equation 3 is the boundary condition that accounts for cells having all Δ = 0 at birth, irrespective of their size 2s at division.

Equation 2 goes back at least to [34, 35], and the formalism was then expanded and utilized for the sizer, the timer, and their combinations in a series of papers and books (see, e.g., [8, 24, 36–39]). We took the pragmatic approach of extracting the functions g and γ from the distribution of the sizes at division and of the elongation rates and using them to simulate the cell-size control process at the level of individual cells. We then compared statistical observables alternative to those used for the calibration of the model. As detailed in the Supplemental Information, this procedure allowed us to rule out timer and sizer models and to establish the consistency of the Δ model.

Supplemental Information

Supplemental Information includes Supplemental Experimental Procedures, four figures, three tables, and one movie and can be found with this article online at http://dx.doi.org/10.1016/j.cub.2014.12.009.

Author Contributions

J.P., M.V., and S.J. designed the research; S.T.-A., S.B., J.T.S., N.S.H., P.A.L., M.V., and S.J. performed the research; and S.T.-A., S.B., J.P., M.V., and S.J. wrote the paper. S.T.-A. and S.B. contributed equally to this work.

Acknowledgments

This work was supported by the Paul G. Allen Foundation, the Pew Charitable Trusts, and the National Science Foundation CAREER Award (to S.J.) and by the NIH grant GM 64671 (to P.A.L.).

Received: October 24, 2014
Revised: November 23, 2014
Accepted: December 2, 2014
Published: December 24, 2014

References

1. Schaechter, M., Maaløe, O., and Kjeldgaard, N.O. (1958). Dependency on medium and temperature of cell size and chemical composition during balanced grown of Salmonella typhimurium. J. Gen. Microbiol. *19*, 592–606.
2. Cooper, S., and Helmstetter, C.E. (1968). Chromosome replication and the division cycle of *Escherichia coli* B/r. J. Mol. Biol. *31*, 519–540.
3. Donachie, W.D. (1968). Relationship between cell size and time of initiation of DNA replication. Nature *219*, 1077–1079.
4. Koch, A.L., and Schaechter, M. (1962). A model for statistics of the cell division process. J. Gen. Microbiol. *29*, 435–454.
5. Powell, E.O. (1964). A note on Koch & Schaechter's hypothesis about growth and fission of bacteria. J. Gen. Microbiol. *37*, 231–249.
6. Sompayrac, L., and Maaloe, O. (1973). Autorepressor model for control of DNA replication. Nat. New Biol. *241*, 133–135.
7. Voorn, W.J., Koppes, L.J., and Grover, N.B. (1993). Mathematics of cell division in *Escherichia coli*: comparison between sloppy-size and incremental-size kinetics. Current Topics in Mol. Genet. *1*, 187–194.
8. Osella, M., Nugent, E., and Cosentino Lagomarsino, M. (2014). Concerted control of *Escherichia coli* cell division. Proc. Natl. Acad. Sci. USA *111*, 3431–3435.
9. Amir, A. (2014). Cell size regulation in bacteria. Phys. Rev. Lett. *112*, 208102.
10. Son, S., Tzur, A., Weng, Y., Jorgensen, P., Kim, J., Kirschner, M.W., and Manalis, S.R. (2012). Direct observation of mammalian cell growth and size regulation. Nat. Methods *9*, 910–912.
11. Scott, M., Gunderson, C.W., Mateescu, E.M., Zhang, Z., and Hwa, T. (2010). Interdependence of cell growth and gene expression: origins and consequences. Science *330*, 1099–1102.
12. Mitchison, J., and Werner, D. (1977). Mitosis Facts and Questions: Proceedings in Life Sciences, M. Little, N. Paweletz, C. Petzelt, H. Ponstingl, D. Schroeter, and H.-P. Zimmermann, eds. (Berlin: Springer), pp. 1–19.
13. Fantes, P.A., and Nurse, P. (1981). Division timing: controls, models and mechanisms. In The Cell Cycle, P.C.L. John, ed. (Cambridge: Cambridge University Press), pp. 11–34.
14. Kafri, R., Levy, J., Ginzberg, M.B., Oh, S., Lahav, G., and Kirschner, M.W. (2013). Dynamics extracted from fixed cells reveal feedback linking cell growth to cell cycle. Nature *494*, 480–483.
15. Weart, R.B., Lee, A.H., Chien, A.C., Haeusser, D.P., Hill, N.S., and Levin, P.A. (2007). A metabolic sensor governing cell size in bacteria. Cell *130*, 335–347.
16. Hill, N.S., Kadoya, R., Chattoraj, D.K., and Levin, P.A. (2012). Cell size and the initiation of DNA replication in bacteria. PLoS Genet. *8*, e1002549.
17. Wold, S., Skarstad, K., Steen, H.B., Stokke, T., and Boye, E. (1994). The initiation mass for DNA replication in *Escherichia coli* K-12 is dependent on growth rate. EMBO J. *13*, 2097–2102.
18. Bates, D., and Kleckner, N. (2005). Chromosome and replisome dynamics in *E. coli*: loss of sister cohesion triggers global chromosome movement and mediates chromosome segregation. Cell *121*, 899–911.
19. Schaechter, M., Williamson, J.P., Hood, J.R., Jr., and Koch, A.L. (1962). Growth, cell and nuclear divisions in some bacteria. J. Gen. Microbiol. *29*, 421–434.
20. Helmstetter, C.E., and Cummings, D.J. (1963). Bacterial synchronization by selection of cells at division. Proc. Natl. Acad. Sci. USA *50*, 767–774.
21. Wang, P., Robert, L., Pelletier, J., Dang, W.L., Taddei, F., Wright, A., and Jun, S. (2010). Robust growth of *Escherichia coli*. Curr. Biol. *20*, 1099–1103.
22. Fantes, P.A., Grant, W.D., Pritchard, R.H., Sudbery, P.E., and Wheals, A.E. (1975). The regulation of cell size and the control of mitosis. J. Theor. Biol. *50*, 213–244.
23. Trueba, F.J., Neijssel, O.M., and Woldringh, C.L. (1982). Generality of the growth kinetics of the average individual cell in different bacterial populations. J. Bacteriol. *150*, 1048–1055.
24. Giometto, A., Altermatt, F., Carrara, F., Maritan, A., and Rinaldo, A. (2013). Scaling body size fluctuations. Proc. Natl. Acad. Sci. USA *110*, 4646–4650.
25. Iyer-Biswas, S., Wright, C.S., Henry, J.T., Lo, K., Burov, S., Lin, Y., Crooks, G.E., Crosson, S., Dinner, A.R., and Scherer, N.F. (2014). Scaling laws governing stochastic growth and division of single bacterial cells. Proc. Natl. Acad. Sci. USA *111*, 15912–15917.
26. Jun, S., and Taheri-Araghi, S. (2014). Cell-size maintenance, universal strategy revealed. Trends Microbiol. Published online December 11, 2014. http://dx.doi.org/10.1016/j.tim.2014.12.001.
27. Fantes, P.A. (1977). Control of cell size and cycle time in *Schizosaccharomyces pombe*. J. Cell Sci. *24*, 51–67.

Please cite this article in press as: Taheri-Araghi et al., Cell-Size Control and Homeostasis in Bacteria, Current Biology (2015), http://dx.doi.org/10.1016/j.cub.2014.12.009

28. Di Talia, S., Skotheim, J.M., Bean, J.M., Siggia, E.D., and Cross, F.R. (2007). The effects of molecular noise and size control on variability in the budding yeast cell cycle. Nature *448*, 947–951.
29. Good, M.C., Vahey, M.D., Skandarajah, A., Fletcher, D.A., and Heald, R. (2013). Cytoplasmic volume modulates spindle size during embryogenesis. Science *342*, 856–860.
30. Mitchison, J.M. (1972). The Biology of the Cell Cycle (Cambridge: Cambridge University Press).
31. Marshall, W.F., Young, K.D., Swaffer, M., Wood, E., Nurse, P., Kimura, A., Frankel, J., Wallingford, J., Walbot, V., Qu, X., and Roeder, A.H.K. (2012). What determines cell size? BMC Biol. *10*, 101.
32. Soupene, E., van Heeswijk, W.C., Plumbridge, J., Stewart, V., Bertenthal, D., Lee, H., Prasad, G., Paliy, O., Charernnoppakul, P., and Kustu, S. (2003). Physiological studies of *Escherichia coli* strain MG1655: growth defects and apparent cross-regulation of gene expression. J. Bacteriol. *185*, 5611–5626.
33. Neidhardt, F.C., Bloch, P.L., and Smith, D.F. (1974). Culture medium for enterobacteria. J. Bacteriol. *119*, 736–747.
34. Collins, J.F., and Richmond, M.H. (1962). Rate of growth of *Bacillus cereus* between divisions. J. Gen. Microbiol. *28*, 15–33.
35. Tyson, J.J., and Diekmann, O. (1986). Sloppy size control of the cell division cycle. J. Theor. Biol. *118*, 405–426.
36. Diekmann, O., Lauwerier, H., Aldenberg, T., and Metz, J. (1983). Growth, fission and the stable size distribution. J. Math. Biol. *18*, 135–148.
37. Perthame, B. (2007). Transport Equations in Biology (Basel: Birkhäuser Verlag).
38. Wheals, A.E. (1982). Size control models of *Saccharomyces cerevisiae* cell proliferation. Mol. Cell. Biol. *2*, 361–368.
39. Robert, L., Hoffmann, M., Krell, N., Aymerich, S., Robert, J., and Doumic, M. (2014). Division in *Escherichia coli* is triggered by a size-sensing rather than a timing mechanism. BMC Biol. *12*, 17.

Extended Supplementary Information for

Cell-Size Control and Homeostasis in Bacteria

Sattar Taheri-Araghi, Serena Bradde, John T. Sauls, Norbert S. Hill, Petra Anne Levin, Johan Paulsson, Massimo Vergassola, and Suckjoon Jun

Current Biology, Volume 25
Doi number: http://dx.doi.org/10.1016/j.cub.2014.12.009

CONTENTS

I. EXPERIMENTAL MATERIALS AND METHODS

A. Strain and growth conditions

Strains: For physiological study, it is important to use a prototrophic strain. For *E. coli* experiments, we chose *E. coli* K12 NCM3722 constructed, sequenced, and extensively tested by Sydney Kustu's lab [1]. We used SJ202, a nonmotile derivative of NCM3722 (ΔmotA). For *B. subtilis* experiments, we chose a strain in the 3610 background with Coml (Q12L) mutation to allow competence. We used a derivative with reduced motility and biofilm formation by deleting *epsH* and a flagellin protein *hag*, respectively.

Growth media: E. coli growth experiments were performed in seven different nutrient conditions. The average generation time in these conditions evenly spans from 17.1 to 51.4 minutes at 37°C. The growth medium is based on MOPS, developed by Fred Neidhardt [2] and is commercially available from Teknova (Teknova.com). *B. subtilis* growth experiments performed in four different growth conditions with average doubling times between 16.9 and 38.9 minutes. The details of the media used for *E. coli* and *B. subtilis* are listed in the Tables SII & SI. Prior to growth of the cells in the microfluidics device, all cultures were grown in a 37°C water bath shaker, shaking at 240 rpm.

Sample preparations: Cell culture for each experiment was carried out in three steps: seed culture, pre-culture and experimental culture, all in the growth medium of the corresponding experiment. The culture volume was 3 ml in 950 mm × 150 mm culture tubes to ensure aeration. For seed culture, cells were inoculated into growth media from -80 glycerol stock, cultured at 37°C with shaking overnight. The culture was then backdiluted 1000-fold in the identical growth medium and cultured in 37°C water bath shaker (pre-culture) until it reaches exponential phase.

The pre-culture was then diluted 1000-fold in the identical growth medium and cultured in 37°C water bath shaker (experimental culture). After a few cell doublings, the experimental culture was centrifuged, and concentrated 20 folds in the identical media containing 50mg/ml of bovine serum albumin (BSA), to reduce nonspecific adsorption of cells to glass surface in the microfluidic device. The concentrated culture was injected in the microfluidics device for single cell imaging. Fresh pre-warmed media was pumped in the device immediately after loading of the cells in the growth channels.

Media name (as used in the text)	Buffer	Carbon source (v/w) concentration	Supplement	Mass doubling time at 37°C
glycerol	MOPS modified buffer	Glycerol 0.4%	–	51.4
sorbitol	MOPS modified buffer	Sorbitol 0.2%	–	50.9
glucose	MOPS modified buffer	Glucose 0.2%	–	37.7
glucose + 6 a.a.	MOPS modified buffer	Glucose 0.2%	table I.b	30.1
glucose + 12 a.a.	MOPS modified buffer	Glucose 0.2%	tables I.b and I.c	26.6
Synthetic rich	MOPS modified buffer	Glucose 0.2%	table I.d	22.5
Tryptic soy broth (TSB)	Tryptone 1.7%, Soytone 0.3%, sodium chloride 0.5%, glucose 0.25%, and dipotassium phosphate 0.25%			17.1

Table I.a: MOPS modified buffer

Components	Concentration
MOPS (MW 209.3)	40mM
Tricine (MW 179.2)	4.0mM
Iron Sulfate Stock	0.01mM
Ammonium Chloride	9.5mM
Potassium Sulfate	0.276mM
Calcium Chloride	0.0005mM
Magnesium Chloride	0.525mM
Sodium Chloride	50mM
Ammonium Molybdate	3×10^{-9} M
Boric Acid	4×10^{-7} M
Cobalt Chloride	3×10^{-8} M
Cupric Sulfate	10^{-8} M
Manganese Chloride	8×10^{-8} M
Zinc Sulfate	10^{-8} M
Potassium Phosphate Dibasic	1.32mM

Table I.b: supplements for glucose+6 a.a. and +12 a.a.

Components	Concentration (μg/ml)
L-Methionine	5
L-Histidine	5
L-Arginine	5
L-Proline	5
L-Threonine	5
L-Tryptophan	5

Table I.c: supplements for glucose+12 a.a.

Components	Concentration (μg/ml)
L-Serine	5
L-Leucine	5
L-Tyrosine	5
L-Alanine	5
L-Asparagine	5
L-Aspartic Acid	0.25

Table I.d: supplements for synthetic rich media

Components	Concentration
L-Alanine	0.8mM
L-Arginine	5.2mM
L-Asparigine	0.4mM
L-Aspartic Acid, Potassium Salt	0.4mM
L-Glutamic Acid, Potassium Salt	0.66mM
L-Glutamine	0.6mM
L-Glycine	0.8mM
L-Histidine HCl H_2O	0.2mM
L-Isoleucine	0.4mM
L-Proline	0.4mM
L-Serine	10mM
L-Threonine	0.4mM
L-Tryptophan	0.1mM
L-Valine	0.6mM
L-Leucine	0.8mM
L-Lysine	0.4mM
L-Methionine	0.2mM
L-Phenylalinine	0.4mM
L-Cysteine HCl	0.1mM
L-Tyrosine	0.2mM
Thiamine	0.01mM
Calcium Pantothenate	0.01mM
para-Amino Benzoic Acid	0.01mM
para-Hydroxy benzoic Acid	0.01mM
di Hydroxy Benzoic Acid	0.01mM
Potassium Hydroxide	1.5mM
Adenine	0.2mM
Cytosine	0.2mM
Uracil	0.2mM
Guanine	0.2mM

Table SI List of growth media, carbon sources and the supplements that we used for *E. coli* in this study.

Media name (as used in the text)	Buffer	Carbon source (v/w) concentration	Supplement	Mass doubling time at 37°C
$S7_{50}$ glycerol	tables II.a and II.c	Glycerol 1%	table II.b	35.6
$S7_{50}$ glucose	tables II.a and II.c	Glucose 1%	table II.b	38.9
LB	Tryptone 1%, Yeast extract 0.5% and NaCl 1%			17.8
Tryptic soy broth (TSB)	Tryptone 1.7%, Soytone 0.3%, sodium chloride 0.5%, glucose 0.25%, and dipotassium phosphate 0.25%			16.9

Table II.a: $S7_{50}$ salts

Components	Concentration
MOPS buffer	50mM
Ammonium sulfate $(NH_4)_2SO_4$	1mM
Potassium phosphate monobasic KH_2PO_4	5mM

Table II.b: supplemets

Components	Concentration
glutamate	6.8mM
Na-Citrate	250uM
$FeCl_3$	250uM

Table II.c: $S7_{50}$ metals

Components	Concentration
$MgCl_2$	2mM
$CaCl_2$	700μM
$MnCl_2$	50μM
$ZnCl_2$	1μM
$FeCl_3$	5μM
Thiamine-HCl	1mM
HCl	20μM

Table SII List of growth media, carbon sources and the supplements that we used for *B. subtilis* in this study.

Figure S1. A view of the environmental chamber. Equipment is fully housed in the 37°C environmental chamber in order to avoid temperature fluctuations.

B. Temperature control

All experimental steps - from inoculation to imaging - were performed at 37 ± 0.1°C. To this end, all equipment was stationed in a 5' x 7' environmental chamber (Figures S1 and S2) to eliminate any side effect of temperature fluctuations on the cell growth and physiology. Within the chamber, the temperature distribution was homogeneous with forced air circulation within ± 0.1°C, and constantly monitored at multiple locations.

C. Microfluidics

To monitor single cell growth, we employed the microfluidic mother machine that we developed to study cell growth and the cell cycle [3]. Master molds, from each of which many PDMS microfluidic devices were cast, were fabricated using standard nanofabrication techniques (detailed protocols are available in "The Mother Machine Handbook" from the web site of Jun lab at http://jun.ucsd.edu and a video at http://www.youtube.com/watch?v=RGfb9XU5Oow). PDMS was prepared from a Sylgard 184 Silicone Elastomer kit: polymer base and curing agent were mixed in a 10 to 1 ratio, air bubbles were purged from the mixture in a vacuum chamber, the degassed mixture was poured over the master, and the devices were cured about 24 hours at 65°C. Cured PDMS has a consistency like rubber; devices were peeled from the master mold. Devices were treated with pentane and then acetone to remove residual uncured

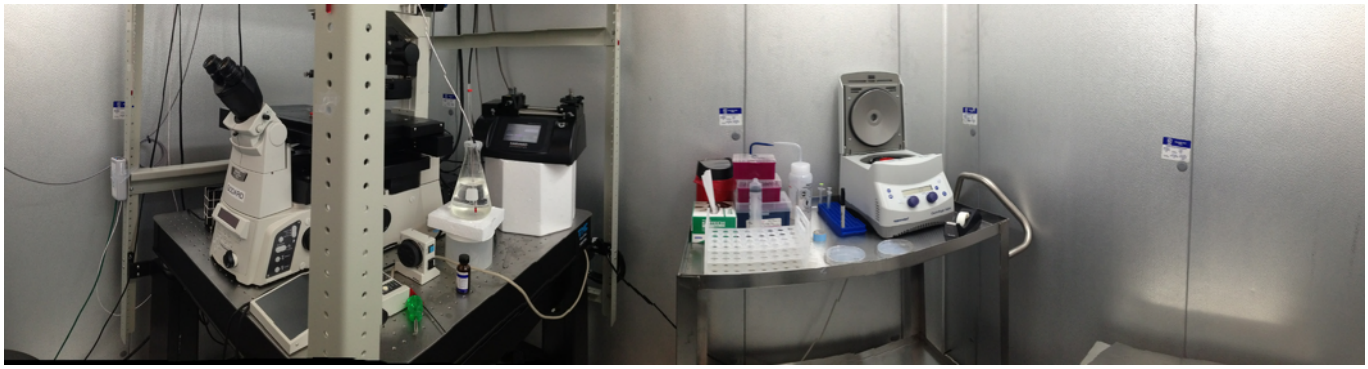


Figure S2. A panoramic view of the interior of the environmental chamber showing the microscope and the other facilities that we employed for the growth and the imaging of live cells.

polymer from the PDMS matrix.

The inlet and outlet of the microfluidic device were bolstered with a thicker layer of PDMS. To bond the PDMS layers, the surfaces were exposed to oxygen plasma for 15 seconds at 30 watts in a Harrick Plasma system. Oxygen plasma makes exposed PDMS and glass reactive, so that covalent bonds form between surfaces brought into contact with one another. The seal between PDMS surfaces was established for 10 minutes at 65 °C.

The PDMS mother machine device was sealed to a Willco-dish glass bottom dish (Willco Wells) with no.1.5 glass thickness (170 μm). Before use, the glass surfaces were wiped with ethanol and then immediately with water, to remove residual ethanol. As before, the PDMS and glass surfaces were exposed to oxygen plasma for 15 seconds. Then the PDMS microfluidic device was sealed to the cover slip and the bonds were allowed to set for 10 minutes at 65 °C. After constructing the device, the channel walls were incubated with 50 mg/ml BSA (Sigma) for surface passivation. The growth medium also contained 0.5 mg/ml BSA in the background.

Because the *E. coli* cell width changes with respect to the growth condition, each growth condition required a unique mother machine device optimized for the experiment.

D. Microscopy and image acquisition

An inverted microscope (Nikon Ti-E) equipped with Perfect Focus (PFS 3), a 100x oil immersion objective lens (NA 1.45), white LED transmission light (TLED, Sutter Instruments, 400-700nm) and Andor NEO sCMOS camera was used for phase contrast imaging. The illumination condition was 50ms exposure with illumination intensity set at 10% of the max TLED intensity. The frequency of the time lapse imaging was chosen such that about 20 or more images were taken per generation time.

E. Image analysis

We developed custom high-throughput image analysis software optimized for our growth experiments in the mother machine. The procedure includes processing of phase contrast images to eliminate background patterns from the PDMS device. The overall steps are as follows:

- Step 1: Subtract the image of an empty channel from the channels containing cells. The purpose of this step is to eliminate patterns from the PDMS device.
- Step 2: Find the mid-cell axis that runs through the long axis of the cells.
- Step 3: Project the brightness profile of the cells along their mid-cell axis.
- Step 4: From the brightness profile, find the cell poles and the length of the cells.

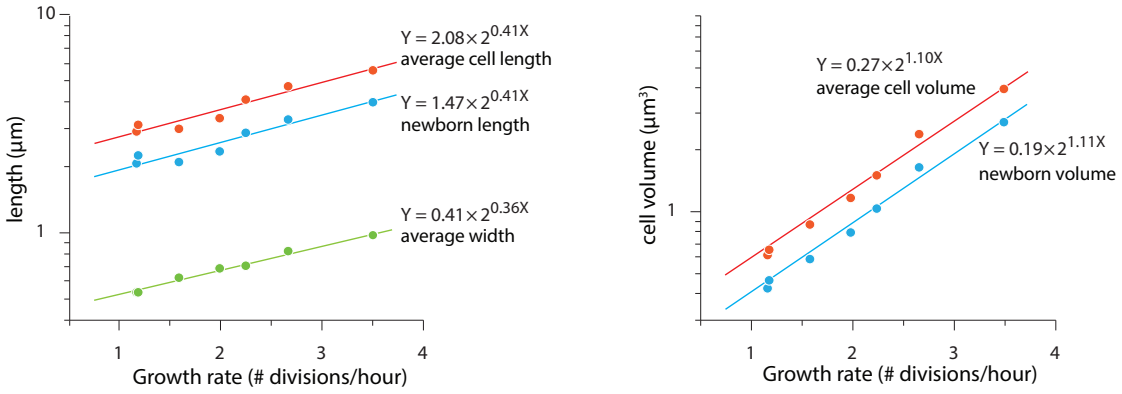


Figure S3. The behavior of *E. coli* width, length and volume as a function of growth rate (divisions/hour). (left) Width, length at birth and average length of the cells as a function of the growth rate. (right) Newborn cell volume, and average cell volume as a function of the growth rate. Both figures are on semi-log panels.

II. FULL EXPERIMENTAL DATA

A. Growth law: variation of cell volume, length and width as a function of growth rate

The growth law, based on the two classic works [4; 5], established an exponential relationship between the average cell size and the growth rate (number of divisions per hour). In our study, using single cell data, we went a step beyond and tested the growth law at the single cell level. Figure 1C of the main text confirms the exponential relationship between average volume (or, equivalently, the volume of newborn cells) and the growth rate of *E. coli*, consistent with the classic literature.

At the single cell level, we measured both the width and the length of the cell and the resulting data is shown in Figure S3. Importantly, *E. coli* cells maintained a constant width during elongation. This constancy can be seen in Figure S4A, where the cell width varies less than 1% between birth and division (by linear fit; see also Table SII). Furthermore, the width of the cell was independent of its length (Figure S4B).

Because the cell width was constant for each growth condition, we were able to calculate the volume of the cell assuming the geometry of the cell is cylindrical with two hemispherical caps. The resulting values are reported as a function of the growth rate.

The growth law between the average newborn cell volume and the growth rate raises a question whether the newborn cell length and width also show an exponential dependence on the growth rate. Figure S3 suggests that may indeed be the case.

B. Correlation of parameters with respect to the newborn size

In Figure 1D of the main text we presented the correlation between a set of parameters of interest, viz., the rescaled elongation rate, the generation time and the size of the cell at division, as a function of the rescaled newborn size. Here, we present data separately by showing rescaled scatter plots for the various growth conditions in Figure S5. The trend is manifestly similar for all the growth conditions.

C. Constancy of the added size with respect to the size of the cell at birth

As a quantitative test of the Δ model, we analyzed the distribution of the added size at division, $\Delta_d \equiv s_d - s_b$, where s_d and the s_b are the size of the cell at division and birth, respectively. In the SI, we shall use for size increments at division the notation Δ_d, with the index meant to distinguish the value at division from the value during elongation that we shall need later. Figure 2 in the main text shows the distributions of the added size, $P(\Delta_d|s_b)$, conditional to s_b, which were observed to be independent of s_b. Here, we make a direct comparison with the conditional distribution of the size at division, $P(s_d/\langle s_d\rangle|s_b)$, for all the experiments. Figure S6 depicts $P(\Delta_d|s_b)$

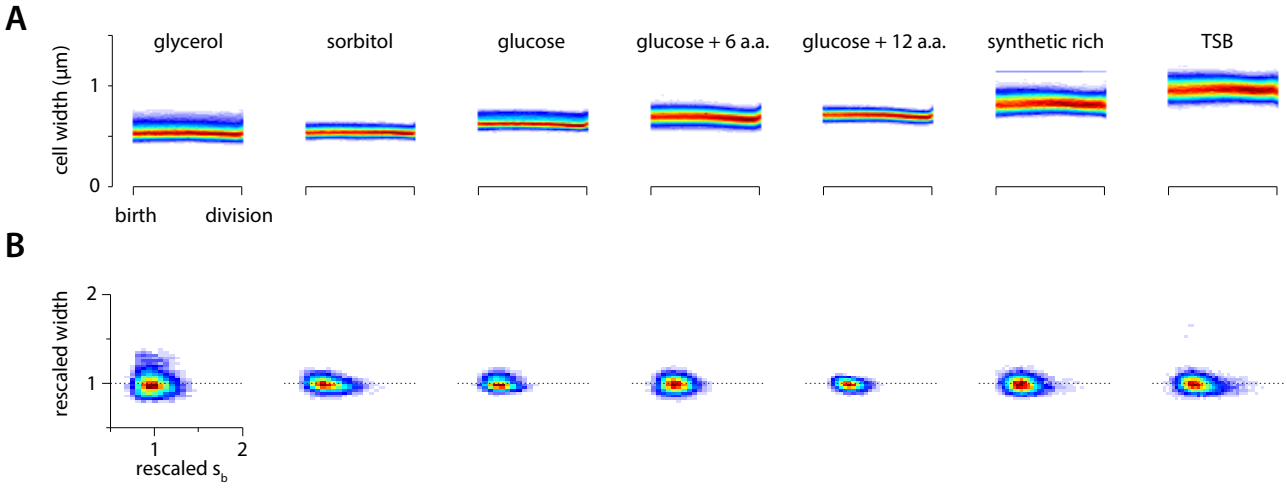


Table of statistics

	panel A	panel B		
growth condition	width change from birth to division in μm	linear fit slope	% of width change over 20% of length change	width CV (%)
glycerol	-0.024	-0.019	0.4	9.03
sorbitol	-0.013	-0.079	1.6	5.18
glucose	-0.026	-0.037	0.7	5.03
glucose + 6. a.a.	-0.046	-0.031	0.6	6.05
glucose + 12 a.a.	-0.036	-0.06	1.2	4.22
synthetic rich	-0.015	-0.034	0.7	5.84
TSB	-0.006	-0.054	1.1	6.11

Figure S4. The behavior of the width of *E. coli* cells in the steady states at given growth conditions. (A) The width remains constant throughout the cell cycle, from birth to division. (B) The absence of correlation between the length at birth (rescaled by its mean value) of the cells and their width (rescaled by its mean value). The resulting correlation between width and length of the cells is negligible and smaller than the coefficient of variation of the width itself.

Table SII A numerical comparison of the variation of the cell width with respect to the newborn size and to the age of the cells. The second column (data from panel A) shows the change of the average width from birth to division, based on a linear fit to the data. The columns three and four (data from panel B) show the dependence of the cell width on the newborn length and the change of the average width normalized by the coefficient of variation (CV) of the newborn length, which is about 20%. The last column shows the coefficient of variation (CV) of the cell width in different conditions.

(top panels) and $P(s_d|s_b)$ (bottom panels) in different growth conditions. We observe a consistent trend in all of the experiments: the distributions of Δ_d are independent of the newborn cell size while the distributions of s_d clearly depend on s_b. The conditional distribution of the size at division, $P(s_d/\langle s_d\rangle|s_b)$, manifestly shifts as the newborn size s_b increases.

D. Constant Δ evidence for E. coli K-12MG1655 and B/r strains

Using data from published work by Wang et al. [3], we have looked into the constancy of Δ with respect to newborn size of the cells. The experiments have been performed in similar microfluidics device in LB media at 37°C. Figure 7 shows that the added size is indeed constant with respect to newborn size and that the distribution of $\Delta/\langle\Delta\rangle$ is independent of newborn size.

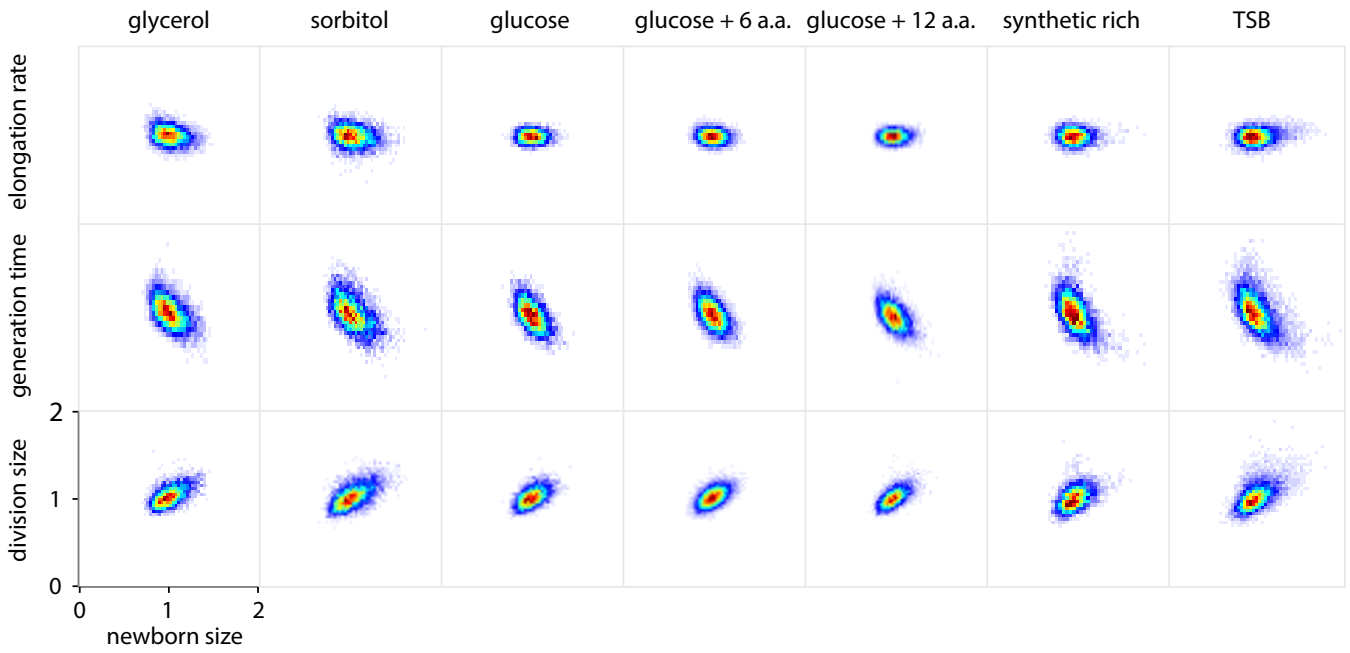


Figure S5. Correlations between *E. coli* rescaled parameters in different growth conditions. The three rows show the rescaled (by their mean) elongation rate, generation time and size of the cell at division *vs* the size of the cell at birth. Note that the trends are similar for different conditions, strongly suggesting that the size control mechanism is independent of the growth conditions. The plot in Figure 1D of the main text is a combination of the data presented here.

E. Correlation between growth rate λ and added size Δ

To justify use of the growth rate distribution $\rho(\lambda)$ and the added size distribution $\rho(\Delta)$ as two independent basic controls of physiology and size homeostasis, we measured their correlations. We confirm the two show negligible correlations (Fig. S8).

F. Relaxation of atypical cell sizes to typical values

The Δ model implies that if a group of cells is prepared with a size initially deviating from the population average value, their average size will approach the population average as generations elapse (see Section IV for details). To test this, we collected a subpopulation of the cells in each experiment, namely we sorted one group of large cells and one group of small cells. We then followed these subpopulations for 10 generations and monitored their average newborn size after each division. After 10 generations, the cell size of each group essentially converged to the average size ($\langle s_b \rangle = \langle \Delta_d \rangle$) for the corresponding growth condition. In Figure 3b of the main text, we presented the convergence of the rescaled cell sizes for a combination of all experiments and we observed consistency with the Δ model. Figure S9 shows the convergence of the size for each growth condition separately. Blue and red markers/lines indicate the large and small subpopulations, respectively. Markers with error bars are the results of the experiments; solid lines show the theoretical prediction of the Δ model, which will be presented in Section IV.

G. Distributions of the various growth parameters

Growth parameters, such as elongation rate, generation time, and size at birth and division, have some variability over the populations in a given growth condition. Figure 4 in the main text showed the remarkable collapse of the various distributions once the variables are rescaled by their mean. Here we present in Figure S10 the probability distributions of the same parameters yet with no rescaling by their mean. The table in Figure S10 reports the numerical values of the mean, the standard deviation, and the coefficient of variation (CV). The standard deviation for all the parameters increases as the mean increases. Conversely, the CV does not show any particular dependence

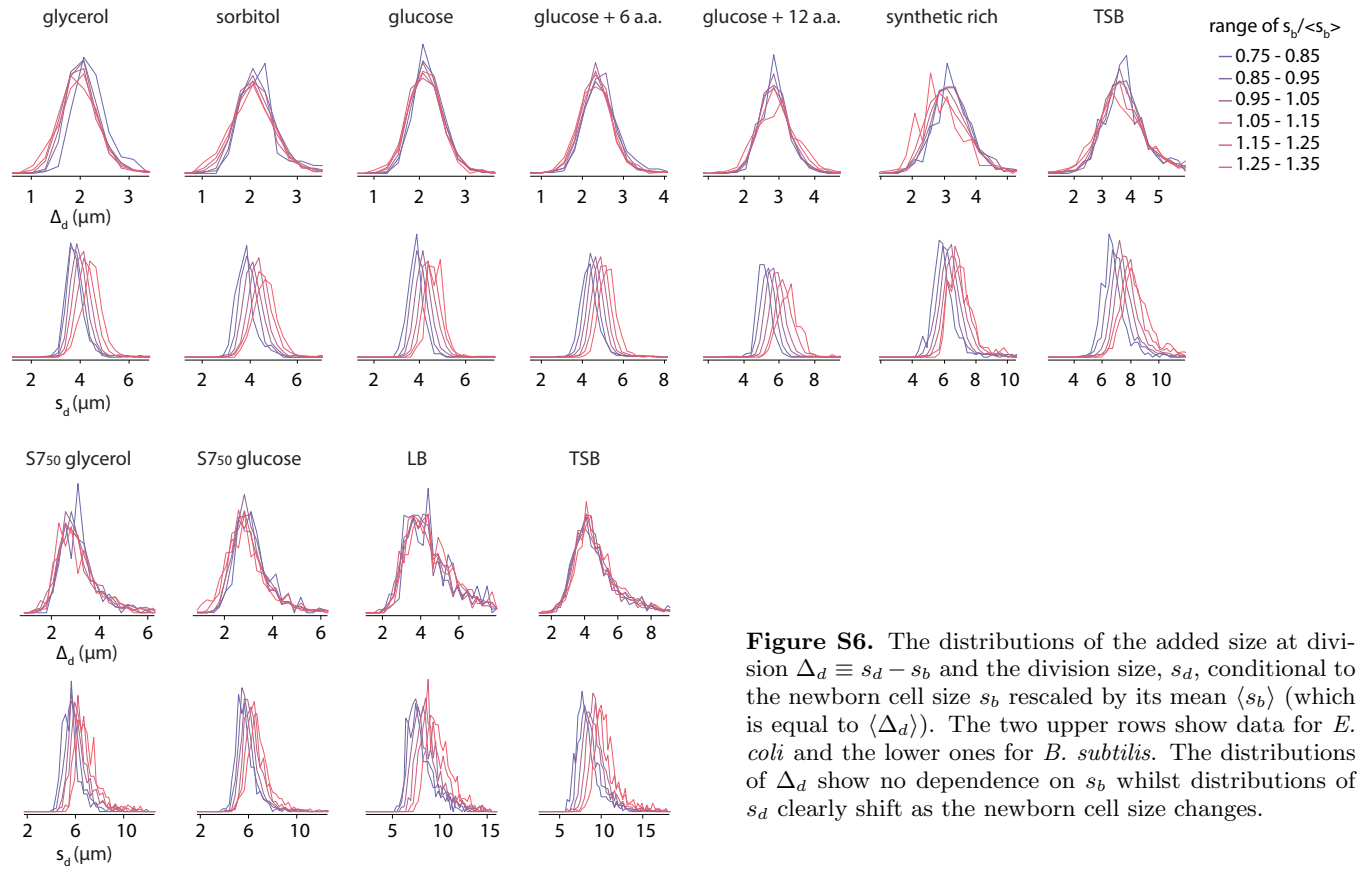


Figure S6. The distributions of the added size at division $\Delta_d \equiv s_d - s_b$ and the division size, s_d, conditional to the newborn cell size s_b rescaled by its mean $\langle s_b \rangle$ (which is equal to $\langle \Delta_d \rangle$). The two upper rows show data for *E. coli* and the lower ones for *B. subtilis*. The distributions of Δ_d show no dependence on s_b whilst distributions of s_d clearly shift as the newborn cell size changes.

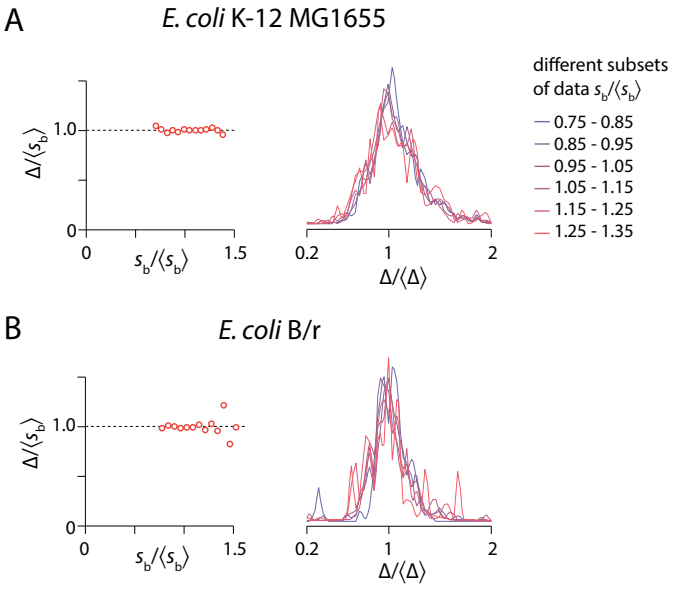


Figure S7. Data from *E. coli* strains K-12 MG1655 and B/r growth in steady-state in LB media at 37°C shows that the added size is indeed constant with respect to newborn size and that the distribution of $\Delta/\langle \Delta \rangle$ is independent of newborn size

on the mean. As discussed in more detail in Section V, this behavior is consistent with a scale-invariant form for the probability distributions. An additional prediction is that higher-order moments of order p normalized by the power p of the mean should be a constant (dependent on p) with respect to the growth conditions. Such behavior is consistent with the data, as shown in Figure S11.

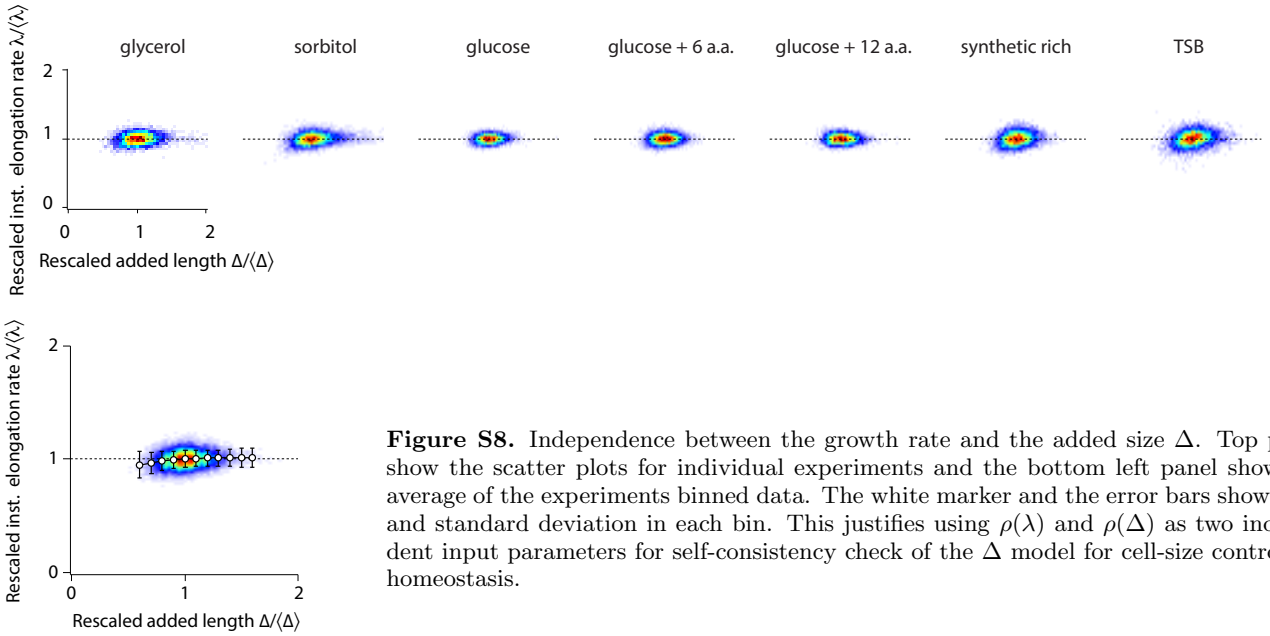


Figure S8. Independence between the growth rate and the added size Δ. Top panels show the scatter plots for individual experiments and the bottom left panel shows the average of the experiments binned data. The white marker and the error bars show mean and standard deviation in each bin. This justifies using $\rho(\lambda)$ and $\rho(\Delta)$ as two independent input parameters for self-consistency check of the Δ model for cell-size control and homeostasis.

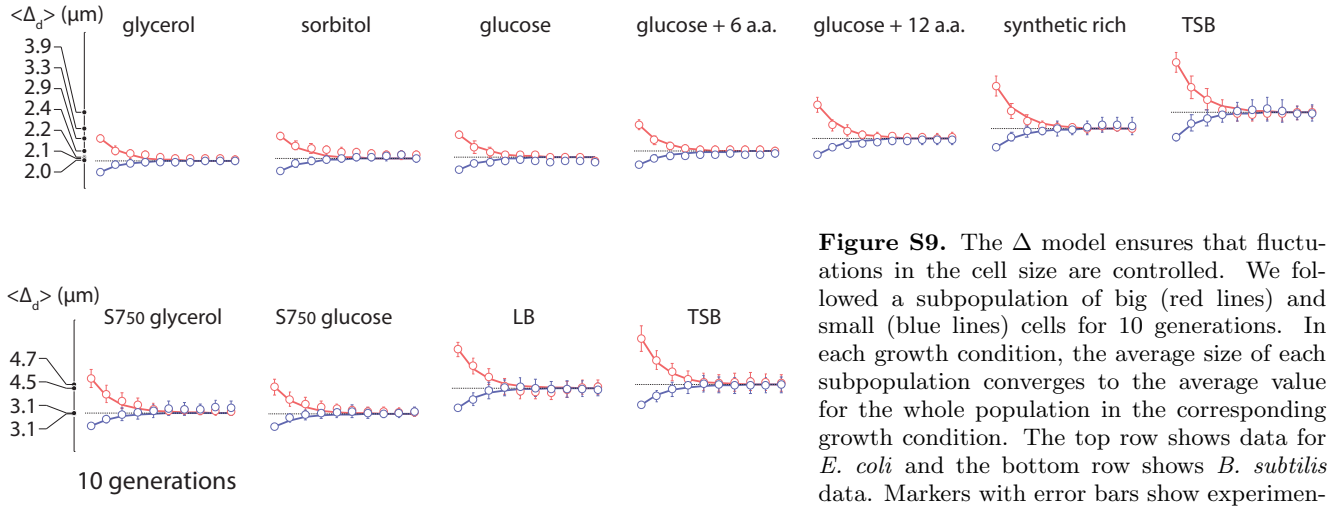
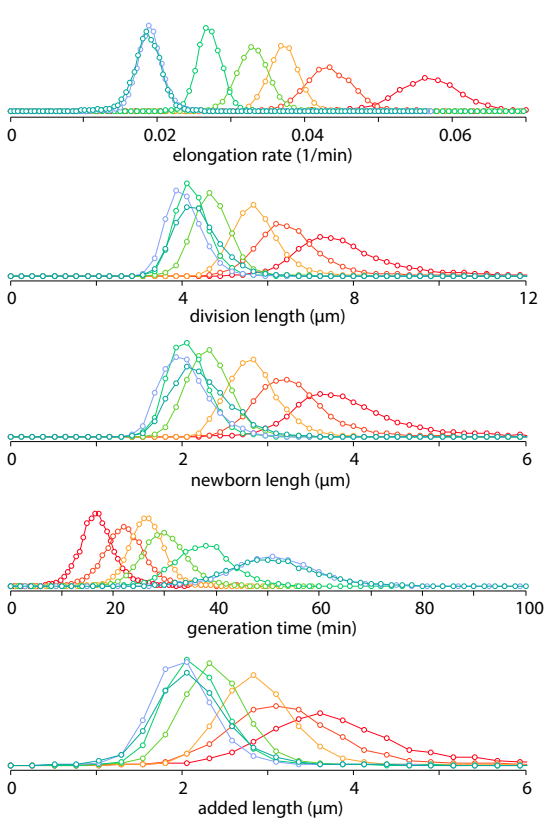


Figure S9. The Δ model ensures that fluctuations in the cell size are controlled. We followed a subpopulation of big (red lines) and small (blue lines) cells for 10 generations. In each growth condition, the average size of each subpopulation converges to the average value for the whole population in the corresponding growth condition. The top row shows data for *E. coli* and the bottom row shows *B. subtilis* data. Markers with error bars show experimental data while solid lines show the theoretical predictions for the Δ model discussed in Section IV.

H. Autocorrelations

Autocorrelation functions can reveal repeating patterns or memory in a series of data and the decay of memory across generations. The autocorrelation function of various parameters of interest, including elongation rate, generation time, newborn size, division size, and added size, are presented in Figure S12. The corresponding theoretical predictions for the Δ model are discussed in the Section IV.A.



elongation rate (1/min)	TSB	Synth. rich	glc+12a.a.	glc+6a.a.	glucose	sorbitol	glycerol
mean	0.0569	0.0433	0.0372	0.0330	0.0270	0.0187	0.0189
sdev σ	0.0045	0.0032	0.0021	0.0022	0.0016	0.0020	0.0018
CV	0.0795	0.0741	0.0561	0.0666	0.0605	0.1062	0.0950
division length (μm)							
mean	7.8912	6.6099	5.7748	4.7535	4.2820	4.3958	4.1129
sdev σ	1.3201	0.9380	0.5880	0.4718	0.4168	0.5448	0.4456
CV	0.1673	0.1419	0.1018	0.0993	0.0973	0.1239	0.1083
newborn length (μm)							
mean	3.9839	3.3361	2.8754	2.3564	2.1091	2.2672	2.0789
sdev σ	0.7170	0.5005	0.3268	0.2747	0.2429	0.3510	0.2936
CV	0.1800	0.1500	0.1137	0.1166	0.1152	0.1548	0.1412
generation time (min)							
mean	17.0861	22.5043	26.6143	30.1391	37.6550	50.9376	51.4759
sdev σ	3.7549	4.6321	3.7947	4.6266	5.8287	9.6385	9.6270
CV	0.2198	0.2058	0.1426	0.1535	0.1548	0.1892	0.1870
added length Δs (μm)							
mean	3.9073	3.2738	2.8994	2.3971	2.1729	2.1287	2.0341
sdev σ	1.0801	0.8400	0.4873	0.4116	0.3595	0.4504	0.3900
CV	0.2764	0.2566	0.1681	0.1717	0.1654	0.2116	0.1917

Figure S10. Normalized distributions of various *E. coli* parameters in different growth conditions. The table on the right shows the mean, the standard deviation and the coefficient of variation (CV) of the quantities plotted on the left.

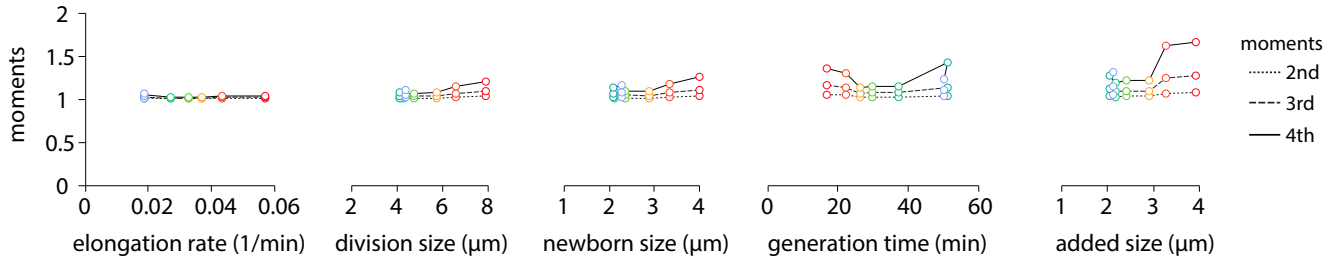


Figure S11. Higher-order moments of the distributions of various *E. coli* parameters. The moments of order p ($p = 2, 3, 4$), divided by the power p of the mean, are plotted as a function of the mean. For a scale-invariant distribution, those normalized moments should be constant with respect to the value of the mean, which is consistent with the data. Deviations observed for the fourth-order moment are due to partial statistical convergence, as witnessed by their increase moving from the left to the right graphs (the direction of increasing fluctuations, see Figure S10.)

III. THEORY: A SUMMARY OF PREVIOUS MODELS FOR THE CELL SIZE CONTROL

Mechanisms previously proposed in the literature for the control of the cell size can be grouped into the following categories: (A) the control depends directly on the size of the cell ("sizer"); (B) the control is based on the age of the cell ("timer"); (c) combinations of the two previous controls. We briefly review the various mechanisms below. We also compare their predictions with the experimental data presented in the main text, illustrating the limitations of the existing models and the necessity of the Δ model discussed later, based on the control of the added size.

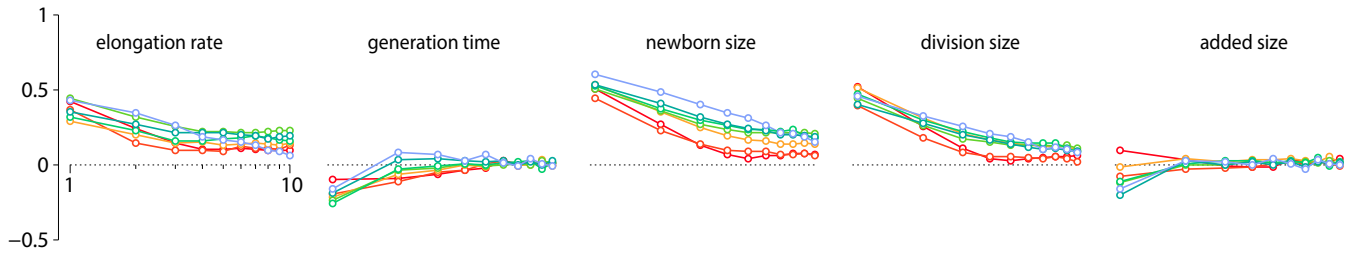


Figure S12. The autocorrelations of various *E. coli* parameters across the generations. The first point (with the unit value on the abscissae) refers to the correlation among mother and daughters; correlations are measured up to 10 generations, when they have essentially vanished.

A. The "sizer"

The main assumption of the model is that cells divide after reaching a threshold size, which is drawn from a probability distribution that depends on growth conditions. This mechanism explains by its very definition the observation that cells have a well-defined distribution of sizes. However, as shown below, its specific predictions are not supported quantitatively by our experimental data.

As shown by Figure S4, *E. coli* cells keep their width roughly constant during their growth and elongate along the axis of the rod. We shall denote the size of the cells along the axis by s and assume $\dot{s} = g(s)$, where the dot indicates the time derivative. Sizes of the cells at birth and division are denoted by s_b and s_d, respectively. The number of individuals in a population of bacteria having size s at time t is denoted $n(s,t)$. The sizer mechanism posits that the rate of division of the cells depends on their size s only. The equation that governs the evolution in time of $n(s,t)$ reads:

$$\partial_t n(s,t) + \partial_s(g(s)\, n(s,t)) = -g(s)\,\gamma(s) n(s,t) + 4g(2s)\gamma(2s) n(2s,t)\,. \tag{1}$$

The left-hand side in (1) is the total time-derivative, with the s-derivative drift term accounting for the elongation of the cells. The right-hand side of (1) stems from the division of cells and $\gamma(s)$ is the so-called splitting rate function [see, e.g., [6]]. The function is defined as the local Poisson rate of cell division, i.e., $\gamma(s)\,ds$ is the probability that a cell of size s divides while growing from s to $s+ds$. Therefore, the probability that a cell with initial size s_b has not divided while reaching the size s_d is $\exp\left[-\int_{s_b}^{s_d}\gamma(s')\,ds'\right]$. Finally, the conversion of the division rate from unit size to unit time involves the factor $|ds/dt| = g(s)$, which justifies its presence in the first term (of loss) on the right hand side of (1).

The second term in the right-hand side of (1) describes the gain in the number of cells in the size range $(s, s+ds)$. The gain originates from those bacteria in the size range $(2s, 2s+2ds)$ which divide and thus halve their size (fluctuations in the size of the offsprings are neglected here for simplicity but will be discussed later). The structure of the second term is thus analogous to the previous one, with s replaced by $2s$ in the arguments of the functions. The factor 4 arises from the product of two factors 2: the first is due to the $2ds$ range of dividing cells and the second is due to division producing two offsprings.

Equation (1) has a long history, which goes back at least to Ref. [7], where the basic formalism was introduced. The formalism was further developed and expanded in [8] and then utilized in a series of papers and books that include [6; 9; 10; 11; 12; 13]. Additional features, such as e.g. fluctuations in the rate of elongation or in the sizes of the offsprings due to the positioning of the septum can be included in the formalism [see, e.g., the Supplementary Material in [13]], but will not be needed here.

The solution to the equation (1) depends on the specific form of the splitting rate $\gamma(s)$ and the elongation law $g(s)$. It is possible though to extract a few general relations that we proceed to derive.

First, integrating (1) over s, we obtain the equality

$$\partial_t \int ds\, n(s,t) = \int ds\, g(s)\,\gamma(s) n(s,t)\,, \tag{2}$$

where the right-hand side is the total number of individuals which divide in the time interval $[t, t+dt]$. For the rate

of growth of the number of individuals in the population, we obtain then

$$\partial_t \ln \int ds\, n(s,t) = \int ds\, g(s)\, \gamma(s) \frac{n(s,t)}{\int ds\, n(s,t)} \equiv \langle g(s)\gamma(s) \rangle \,. \tag{3}$$

At long times, we expect from the very definition of the model that $n(s,t)/\int ds\, n(s,t)$ will reach a stationary distribution that we denote by $\rho_n(s)$. The corresponding rate of growth of the population is denoted $k = \int ds g(s)\gamma(s)\rho_n(s)$.

Second, multiplying (1) by s and integrating, we can obtain for the average size

$$\partial_t \langle s \rangle = \langle g(s) \rangle - \langle g(s)\gamma(s) \rangle \langle s \rangle \,. \tag{4}$$

In the stationary state, the two terms on the right-hand side of (4) will balance. Namely, for an exponential elongation $g(s) = \alpha s \ln 2$, it follows from (4) that the consistency condition $\langle s\, \gamma(s) \rangle = 1$ is satisfied. The factor $\ln 2$ in $g(s)$ is introduced to conform to the microbiology convention of cells elongating as $s(t) = s_b 2^{\alpha t}$.

Finally, multiplying (1) by s^q and integrating, we obtain a series of relations for higher-order moments. In the stationary state and for an exponential elongation $g(s) = \alpha s \ln 2$, the relations read $(q-1)\langle s^q \rangle = (1-2^{1-q})\langle s^{q+1}\gamma(s) \rangle$, where the integer $q > 1$ and we have made use of (4).

1. Comparison and consistency with experiments

The analytical solution of (1) is not available for generic forms of the elongation rate $g(s)$ and of the splitting rate function $\gamma(s)$. Imposing a sharp maximum value S for the size distribution, i.e. $n(s > S, t) \equiv 0$, permits to patch the solution and have it in closed form (see the SI in [11] for details) yet the resulting expressions are not particularly transparent. Therefore, we decided to take a more pragmatic approach: $g(s)$ and $\gamma(s)$ are extracted from experimental data, e.g. from the distribution of the sizes at division and of elongation rates, and we use them to simulate the cell size control process at the level of individual cells. We can then compare additional statistical observables, other than those used to calibrate the model, to assess the consistency and validity of the sizer mechanism for the size control.

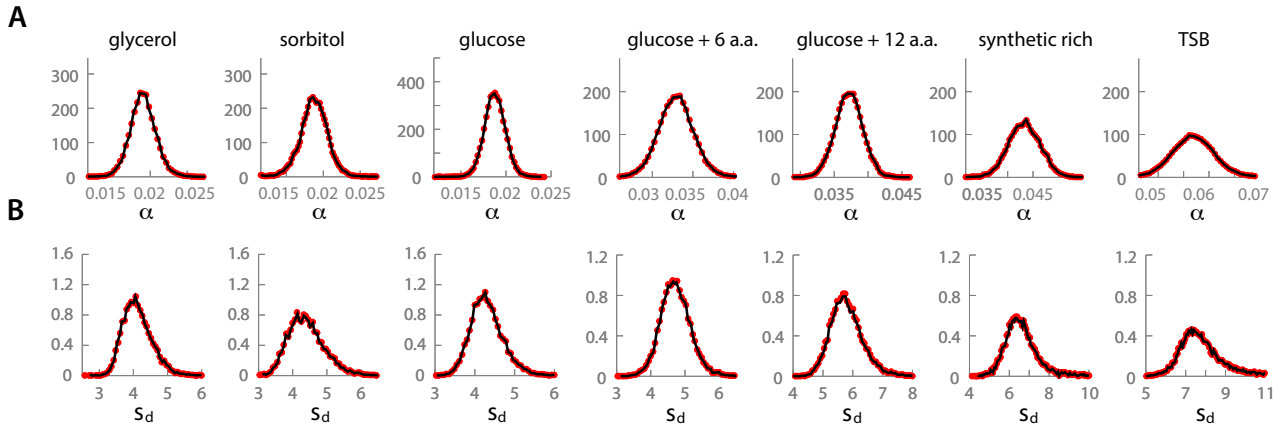


Figure S13. Calibration of the "sizer" model for the control of the cell size. Each cell's elongation rate is a random variable independently drawn from the experimental *E. coli* distribution, $\rho_\alpha^{ex}(\alpha)$, shown in red in panel (A) for the seven different growth conditions discussed in the Section I.A (in the order of increasing growth rate: glycerol, sorbitol, glucose, glucose + 6 a.a., glucose + 12 a.a., synthetic rich, TSB). Correlations among the elongation rates of mothers and siblings are weak and not taken into account; we verified that results are not modified by including them. The splitting rate $\gamma(s)$ is computed from the distribution of the size of the cells at division $\rho_d^{ex}(s_d)$ as detailed in the text (see eqs. (5) and (6)). The experimental distributions are shown in red in panel (B) for the same growth conditions as for the elongation rate. The curves in black are the results of numerical simulations of the model detailed in the text. Their agreement with the experimental curves confirms that the parameters (distribution of the elongation rate α and the splitting rate function γ) of the model are appropriately calibrated.

Specifically, the calibration of the model goes as follows. As for the elongation rates, experimental data are consistent with an exponential law, $g(s) = \alpha s \ln 2$, where the elongation rate α can vary from cell to cell. The values

of α that we use in the simulations are drawn directly from the experimental distribution, $\rho_\alpha^{ex}(\alpha)$ shown in Figure S13 for *E. coli*. Draws are independent for different bacteria. Furthermore, experimental data indicate that the correlation between the elongation rates of mother and daughter cells is weak. We therefore assume that the elongation rate α for each cycle of elongation of a bacterium is drawn independently of its past and future cycles of elongation.

As for the splitting rate function $\gamma(s)$, we compute it from the the experimental size distributions $\rho_d^{ex}(s_d)$ and $\rho_b^{ex}(s_b)$ of the sizes at division and birth, respectively. From the very definition of the splitting rate $\gamma(s)$ we have at the steady state the following relation

$$\rho_d(s_d) = \gamma(s_d) \int_0^{s_d} \rho_b(s_b) e^{-\int_{s_b}^{s_d} \gamma(s')\, ds'}\, ds_b\,. \tag{5}$$

Indeed, $\rho_b(s_b)$ is the probability that a newborn cell has size s_b, the exponential term is the probability that the cell will not divide up to s_d and, finally, $\gamma(s_d)\, ds_d$ is the probability of division in the range $(s_d, s_d + ds_d)$. The integral runs over all possible values of s_b respecting the constraint $s_d \geq s_b$.

Using the experimental distributions for ρ_d and ρ_b, we can extract the splitting rate function γ from (5). In practice, the calculation is simplified by remarking that the distributions at division and birth are well separated and very weakly overlap in their tails, i.e. γ is small where ρ_b is appreciable. A good approximation to (5) is then to assume $\rho_b(s_b) = \delta(s_b)$, which gives

$$\rho_d(s_d) = -\frac{d}{ds_d} e^{-\int_0^{s_d} \gamma(s')\, ds'} \;\Rightarrow\; 1 - \int_0^{s_d} \rho_d(s')\, ds' = e^{-\int_0^{s_d} \gamma(s')\, ds'} \;\Rightarrow\; \gamma(s_d) = \frac{\rho_d(s_d)}{1 - \int_0^{s_d} \rho(s')\, ds'}\,, \tag{6}$$

where we first used that the cumulative distribution and its density are related by $\rho_d(s_d) = -d\left[1 - \int_0^{s_d} \rho_d(s')\, ds'\right]/ds_d$ and then took a logarithm of the second equality and derived it with respect to s_d. The final approximation coincides with the expression for γ given in [6]. Experimental data for the size at division are well reproduced by the approximation (6), as shown in Figure S13.

We now proceed to simulate the dynamics of a growing population using the distribution of the elongation rates α and the splitting function $\gamma(s)$ extracted from the experimental data as we just described. Each cycle of elongation of a cell proceeds at the constant (random) rate α and division occurs with the Poissonian rate $\gamma(s)$. After an initial transient time period, distributions for the various observables reach a stationary form, as expected. The resulting distributions for the final size s_d and α, which we used to calibrate the model, are shown in Figure S13 for the case of *E. coli*. Similar curves are obtained for *B. subtilis* (not shown). Non-trivial tests for the sizer mechanism of control are provided by comparing the results of numerical simulations and experiments for additional observables.

The distribution for newborn sizes s_b is shown in Figure S14. The agreement is good, yet not independent of the calibration process itself since the two distributions of the size at birth and at division are simply related to each other as shown in panel (B) of the figure. A germane test of the model is provided by the distributions of the added size $\Delta_d = s_d - s_b$ and of the generation time $\tau_d = \log_2(s_d/s_b)/\alpha$ for each cell during the growth of the colony. Results obtained by simulations are compared with *E. coli* experimental data in Figure S15 : the former are found to have fatter tails than the latter. The disagreement can be intuitively traced back to the fact that the sizer mechanism posits that division is essentially controlled by the cell size only. Indeed, given an initial size s_b the probability of division in the interval $(s_d, s_d + ds_d)$ is $\propto -d \exp\left[-\int_{s_b}^{s_d} \gamma(s')\, ds'\right]/ds_d$ [see (5)] and the dependency on s_b is weak, as we discussed for the derivation of (6). Conversely, experimental data show that the dependency between the initial and the final size of a cell is strong, as clearly illustrated by the shifts with s_b of the conditional probabilities $P(s_d|s_b)$ shown by experimental data presented in the Figure S6. Discrepancies in Figure S15 are then intuited because the shift of $P(s_d|s_b)$ with s_b reduces the contribution to the left (right) tails of $\Delta_d = s_d - s_b$ and τ_d from those cells whose initial size is larger (smaller) than the average.

A final check that the dependency of the division rate on s_b is indeed the crucial element missing in the sizer mechanism is provided in Figure S16. Picking randomly s_b and s_d from the two distributions $\rho_b^{ex}(s_b)$ and $\rho_d^{ex}(s_d)$ (with the only constraint $s_d \geq s_b$ and no additional correlations) we show that the sizer model predictions are recovered. In conclusion, the sizer model ensures a stationary distribution of the cell size, yet its quantitative predictions disagree with the experiments. Discrepancies are due to the fact that the sizer mechanism neglects the role of the size at birth s_b on the control of the cell size at division s_d.

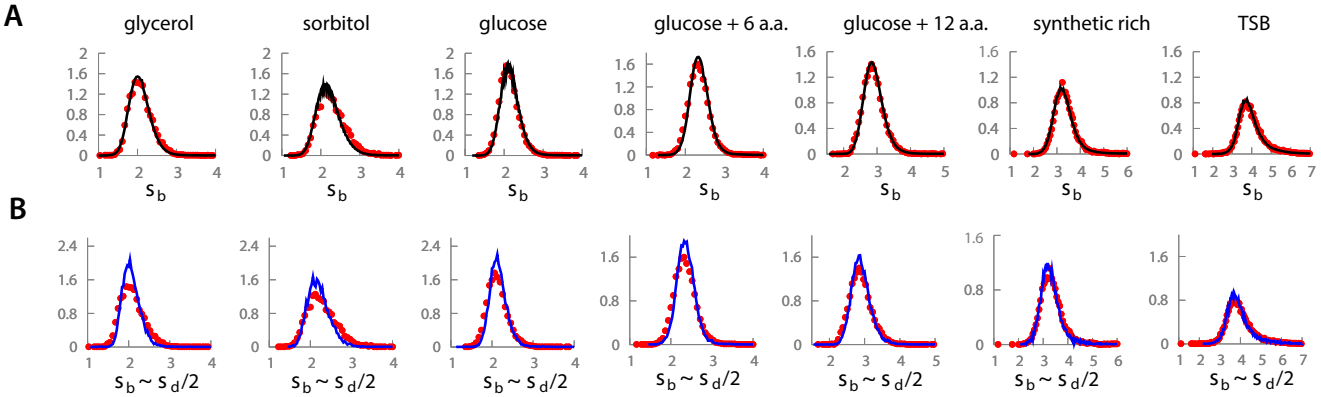


Figure S14. The "sizer" model for the control of the cell size reproduces some features of the experimental data but not all (see Figures S15 and S16). In panel (A), the model calibrated as in Figure S13 is simulated numerically and the distributions of the newborn cell size s_b are reported and compared to *E. coli* experimental data for the same growth conditions as in Figure S13. Black curves refer to numerics while red dots are the experimental data; the agreement is excellent. However, the agreement is strongly related to the calibration process since $s_b \simeq s_d/2$. Additional noise due to the fluctuations in the halving of the size of the daughter cells and the positioning of the septum is relatively weak, as shown in panel (B). The panel shows the experimental distributions of s_b (red dots) on top of the experimental distribution of the quantity $s_d/2$ (blue lines) showing their strong overlap.

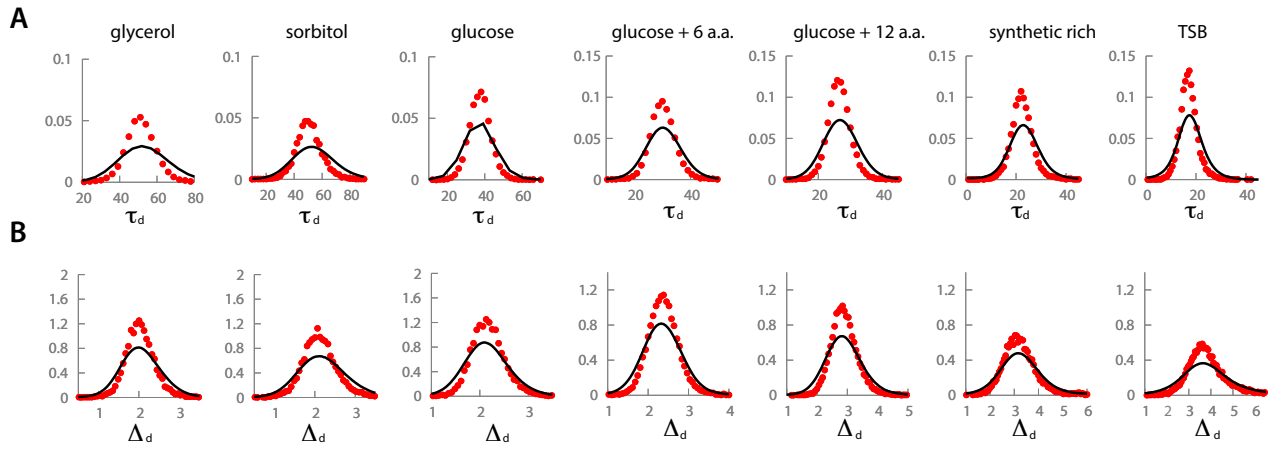


Figure S15. Test of the "sizer" model for the control of the cell size. The model calibrated as in Figure S13 is simulated numerically and we report the distribution of the generation time $\tau_d = \log_2(s_d/s_b)/\alpha$ and of the added size $\Delta_d = s_d - s_b$ for the case of *E. coli*. Black curves refer to numerics and red dots to experiments, respectively. The seven panels in (A) show the curves for the generation time τ_d for all the growing conditions while panel (B) presents the distribution of the Δ_d. Tails in the experimental data are significantly less pronounced for the reasons discussed in the body of the text. A similar disagreement is observed for *B. subtilis*.

B. The "timer"

The "timer" mechanism posits that cell division is controlled by the age of the cell, i.e. the time elapsed since its birth. The state of cells is therefore described by their size s and their age ξ and the corresponding number of cells at time t is denoted by $n(s, \xi, t)$. The equation for the evolution of n reads:

$$\partial_t n(s,\xi,t) + \partial_s \left[g(s) n(s,\xi,t)\right] + \partial_\xi\, n(s,\xi,t) = -\gamma(\xi) n(s,\xi,t)\,; \quad n(s,0,t) = 4\int \gamma(\xi') n(2s,\xi',t)\, d\xi'\,. \tag{7}$$

The left hand side is the total time derivative: the drift in s is due to the elongation of cells $\dot{s} = g(s)$ while the drift in ξ is due to the aging of cells $\dot{\xi} = 1$ (the discontinuity in ξ occurring at division will be addressed momentarily).

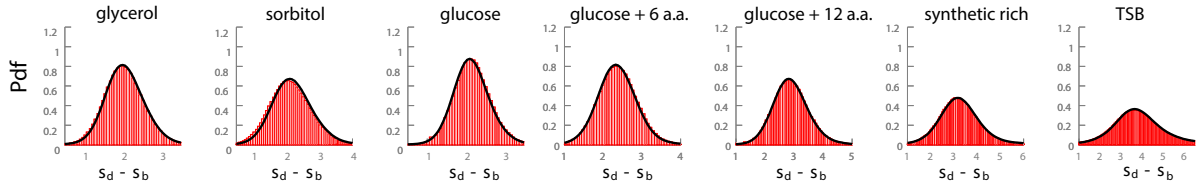


Figure S16. The "sizer" model is inconsistent with experimental data because it neglects the effects of the initial size s_b onto the control of the size s_d at division of the cells. The point is directly demonstrated by choosing randomly s_b and s_d from their respective experimental distributions $\rho_b^{ex}(s_b)$ and $\rho_d^{ex}(s_d)$ for *E. coli*. The only constraint we impose in the random draw is $s_d \geq s_b$ and no additional correlations are retained. The resulting distributions for the "scrambled" added size $s_d - s_b$ (shown in red) are indeed in agreement with the sizer model results (black curves).

The right hand side of the first equation in (7) is the loss term due to the division of cells. The Poisson splitting rate function γ depends now on the age of the cell, ξ. Furthermore, since $\dot{\xi} = 1$ there is no additional factor coming from the conversion of the rate of division from unit age to unit time (see (1)). From the definition of the splitting rate it follows that

$$\rho_{\tau_d}(\tau) = \gamma(\tau)\, e^{-\int_0^\tau \gamma(\xi')\, d\xi'} \;\Rightarrow\; \gamma(\tau) = \frac{\rho_{\tau_d}(\tau)}{1 - \int_0^\tau d\xi' \rho_{\tau_d}(\xi')}\,, \tag{8}$$

where $\rho_{\tau_d}(\tau)$ is the probability density for the generation time τ_d of a given cell and the derivation proceeds as for (6). Finally, the last equation in (7) is the boundary condition that accounts for newborn cells having all the same age $\xi = 0$, irrespective of their size $2s$ that gets halved. The integral $\int d\xi\, \gamma(\xi) n(2s, \xi, t)$ represents the total number of cells that divide in the unit time; as in (1), the factor 4 is the product of the factor 2 resulting from the $2\, ds$ width of the range of dividing cells and the factor 2 due to division producing two newborn daughters cells.

The dynamics of the age of the cells is independent of their size, as it can be easily seen by integrating (7) over s. The resulting equations for the marginal distribution $n(\xi, t) = \int n(s, \xi, t)\, ds$ reads

$$\partial_t n(\xi, t) + \partial_\xi n(\xi, t) = -\gamma(\xi) n(\xi, t)\,; \quad n(0, t) = 2 \int d\xi\, \gamma(\xi)\, n(\xi, t)\,. \tag{9}$$

It is also immediate to integrate (7) over s and ξ, obtaining for the rate k of growth of the population

$$\partial_t \int n(s, \xi, t)\, ds\, d\xi = \int \gamma(\xi) n(s, \xi, t)\, ds\, d\xi = \int \gamma(\xi) n(\xi, t) d\xi\,, \tag{10}$$

which depends on the marginal distribution $n(\xi, t)$ only. The well-known solution [14] to (9) is $n(\xi, t) = e^{kt} \tilde{n}(\xi)$, where

$$\tilde{n}(\xi) = A e^{-\int_0^\xi \gamma(\xi')\, d\xi' - k\xi}\,; \quad 2k \int_0^\infty e^{-k\xi} e^{-\int_0^\xi \gamma(\xi')\, d\xi'}\, d\xi = 1\,. \tag{11}$$

The first equation is simply obtained by integrating the differential equation in (9) and the constant A depends on the initial size of the population. The second relation is obtained from the boundary condition in (9) by an integration by parts. A similar integration by parts also allows to check that the second equation in (11) is equivalent to the growth rate relation $k \int \tilde{n}(\xi)\, d\xi = \int \tilde{n}(\xi) \gamma(\xi)\, d\xi$ derived in (10).

As for the dynamics of the size of cells, we can obtain the equation for the mean size multiplying (7) by s and integrating over s and ξ. The resulting expression reads

$$\partial_t \langle s \rangle = \langle g(s) \rangle - k \langle s \rangle\,, \tag{12}$$

where $\langle s \rangle \equiv \int s\, n(s, \xi, t)\, ds\, d\xi / \int n(s, \xi, t)\, ds\, d\xi$ and similar definitions apply for other averages. For a stationary state, the left-hand side is required to vanish, i.e. elongation should balance the growth. In the case of linear elongation, $g(s) = \text{const.}$ a stationary distribution is reached and $\langle s \rangle = \text{const.}/\langle \gamma(\xi) \rangle$, where we used $k = \langle \gamma(\xi) \rangle$. Conversely, for exponential elongation $g(s) = \alpha s \ln 2$, a stationary distribution obtains only for the special choice $\alpha \ln 2 = k$. If the equality is not satisfied, then the average size grows to infinity or decay to zero exponentially fast (see below).

The lack of control of the size of the cells by the timer mechanism was remarked in [15; 16] and it is intuitively understood by considering the sizes at birth $s_b^{(n)}$ and $s_b^{(n+1)}$ over two consecutive generations n and $n + 1$. For an

exponential elongation, $\log_2 s_b^{(n+1)} - \log_2 s_b^{(n)} \simeq \alpha\tau_d - 1$, where τ_d is again the generation time and α is the elongation rate for the bacterium in the n-th generation. Exponential elongation of the size of the cells makes that the average value of $\alpha\tau_d$ must be precisely tuned in order to avoid a systematic drift of $\log_2 s_b$. Notice that even in the absence of drift, the long term behavior of $\log_2 s_b$ will be analogous to a random walk (assuming that the values of $\alpha\tau_d$ fluctuate and decorrelate over the generations). Therefore, the variance of the size of the cells will grow across the generations and no effective control of the size of the cells is achieved by the timer mechanism. Diluting bacteria by washing them out, e.g. a term $-Dn(s,\xi,t)$ is added to the equation (7), will not modify the previous conclusion unless dilution is coupled to the size, i.e. the dilution rate depends on s.

In practice, the elongation rate is deviating from a linear behavior at very small and very large sizes so that the logarithm of the size will not go to zero or diverge to infinity. However, its behavior will depend very sensitively on the details of the elongation law at very small and large sizes [13; 17] and the resulting size distributions are generally too wide (data not shown). Most importantly, the timer mechanism disagrees with our experimental data in that the conditional distribution $P(\tau_d|s_b)$ of the generation time τ_d *vs* the initial size of cells s_b should be independent of s_b. Conversely, data shown in the main text indicate a clear dependency on s_b.

C. Mixed models

Diverse combinations of sizer and timer mechanisms are conceivable. A well-known instance is the bilinear model proposed in [18]. The proposed dynamics is summarized as follows: cells grow linearly in time $g(s) = u$ until they reach a threshold size s_Λ. The threshold is argued to be universal with respect to growth conditions and its value $s_\Lambda \simeq 2.78\,\mu m$ to be twice the minimal possible length for *E. coli* cells. After reaching s_Λ, cells keep growing for a fixed amount of time (~ 20 mins) at a velocity $2u$ (see [19; 20; 21] for a discussion of bilinear vs exponential elongation). The velocity u depends on the growth conditions and the last 20 mins phase accounts for the dependency of the distribution of the final size s_d on u, and thus on the growth conditions. Ref. [18] sharply poses relevant questions and deserves all the influence it had in the field. However, the specific mechanism which was proposed is not supported by modern experimental data that allow for better statistics. In Figure S17, the size s_d and the generation time τ_d at division *vs* the initial size s_b are presented and the disagreement between the predictions and the data is manifest.

More recently, other combinations of age and size control have been proposed. The intuitive motivation comes from plots like Figure 1D in the main text, where a behavior intermediate between a timer and a sizer is observed. The most recent proposition is [12], where a control mechanism operating concertedly (and not sequentially as in [18]) on the size and the age of the cell is discussed. Namely, no major inconsistency with experimental data (for a single growth condition) is observed by introducing a model where the splitting rate function $\gamma(s,\xi)$ depends jointly on the size and on the age of the cells and by best fitting the function to the data.

IV. THEORY: THE Δ MODEL OF CONTROL

The Δ model that we propose and discuss in the main text posits that the mechanism of control involves a single variable, Δ, the size added to the cells since their birth. The purpose of this Section is to provide further details on the model and its comparison to experimental data.

As in the previous Section, we shall denote by s the size of cells along the elongating axis of the rod. We shall assume $ds/dt = g(s)$ and the width of the cell is roughly constant. We indicate by s_b and s_d the size of cells at birth and division. If $s(t)$ is the size of a cell at the current time t, its added size is denoted $\Delta(t) = s(t) - s_b$. The density of cells $n(s,\Delta,t)$ having size s and added size Δ obeys the following continuity equation

$$\partial_t n(s,\Delta,t) + \partial_s\left[g(s)n(s,\Delta,t)\right] + \partial_\Delta\left[g(s)n(s,\Delta,t)\right] = -\gamma(\Delta)g(s)n(s,\Delta,t)\,; \tag{13}$$

$$g(s)n(s,0,t) = 4\,g(2s)\int_0^\infty \gamma(x)n(2s,x,t)\,dx\,. \tag{14}$$

The left-hand side in (13) is the total time-derivative and the two drift terms are due to the elongation of the cells, i.e. $ds/dt = g(s)$ and $d\Delta/dt = g(s)$. As in the previous Sections on the sizer and the timer, the right hand side accounts for the division of the cells. The Poissonian splitting rate function $\gamma(\Delta)$ depends now on the added size Δ. Proceeding as for (6) and (8), we obtain the relation

$$\rho_{\Delta_d}(\Delta) = \gamma(\Delta)\,e^{-\int_0^\Delta \gamma(x)\,dx} \;\Rightarrow\; \gamma(\Delta) = \frac{\rho_{\Delta_d}(\Delta)}{1 - \int_0^\Delta dx\,\rho_{\Delta_d}(x)}\,, \tag{15}$$

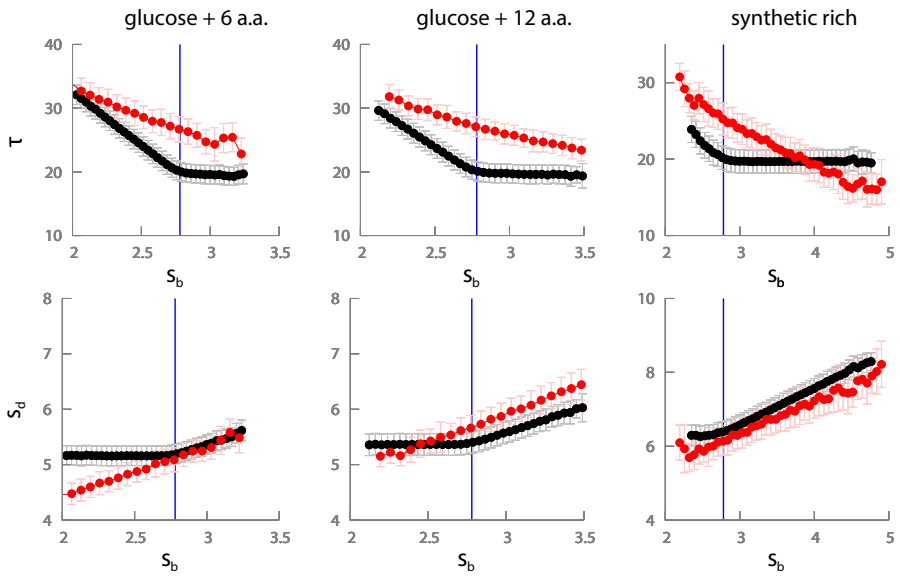


Figure S17. Comparison of bilinearity (black points) with *E. coli* experimental data (red points) for the fast growth conditions glucose + 6 a.a., glucose + 12 a.a. and synthetic rich. The behavior of the size s_d and the generation time τ_d at division *vs* the initial size s_b are presented. A change of behavior should take place [18] at the size s_Λ (indicated by the vertical blue line) but data do not support any such transition. In particular, for $s_b < s_\Lambda$ the division size s_d should be independent of s_b and for $s_b > s_\Lambda$ the generation time τ_d should be independent of s_b. Neither one of these predictions is observed.

between $\gamma(\Delta)$ and the distribution $\rho_{\Delta_d}(\Delta)$ for the size added at division ($\Delta_d = s_d - s_b$) of individual cells. By individual cells we mean that cells should be weighted equally, tracking them individually and avoiding known bias effects related to the speed of reproduction [14]. As in the equation (1), the conversion of the rate of division to unit time involves the Jacobian $|d\Delta/dt| = g(s)$, that appears then in the right hand side of (13). Finally, (14) is the boundary condition that accounts for cells having all $\Delta = 0$ at birth, irrespective of their size $2s$ that gets halved. The integral $\int \gamma(x) n(2s, x, t)\, dx$ represents the total number of cells that divide in the unit time; as in (1), the factor 4 is the product of the factor 2 resulting from the $2\, ds$ width of the range of dividing cells and the factor 2 due to division producing two newborn daughters cells.

A series of relations analogous to those obtained for the sizer model can be derived from (13) for a general splitting rate γ. Integrating (13) over s and Δ and using the boundary condition (14), we obtain

$$\partial_t \ln \int n(s,\Delta,t) ds d\Delta = \int g(s)\gamma(\Delta) \frac{n(s,\Delta,t)}{\int n(s,\Delta,t) ds\, d\Delta} ds\, d\Delta \equiv \langle g\gamma \rangle\,. \tag{16}$$

Multiplying (13) by s and integrating over s and Δ, we derive for the average size

$$\partial_t \langle s \rangle = \langle g(s) \rangle - k\, \langle s \rangle, \tag{17}$$

where the rate of growth of the population $k = \langle g(s)\gamma(\Delta) \rangle$ from (16). For a linear elongation, $g(s) = u$, (17) gives for the size at equilibrium $\langle s \rangle = 1/\langle \gamma(\Delta) \rangle$. For an exponential elongation, $g(s) = \alpha s \ln 2$, (17) yields $\alpha \ln 2 = k$, i.e. the value $\langle s\gamma(\Delta) \rangle = 1$ for the correlation between the size and the splitting rate function.

Finally, we can obtain the following series of relations for higher order moments at the steady state:

$$(1-m)\langle s^m \Delta^q \rangle - q \langle s^{m+1} \Delta^{q-1} \rangle = 2^{1-m} \delta_{q,0} \langle s^{m+1} \gamma(\Delta) \rangle - \langle s^{m+1} \Delta^q \gamma(\Delta) \rangle\,, \qquad m \geq 0\,, q \geq 0\,, \tag{18}$$

where we specified relations for the case $g(s) = \alpha s \ln 2$ and we used $k = \alpha \ln 2$ and $\langle s\gamma \rangle = 1$ derived previously.

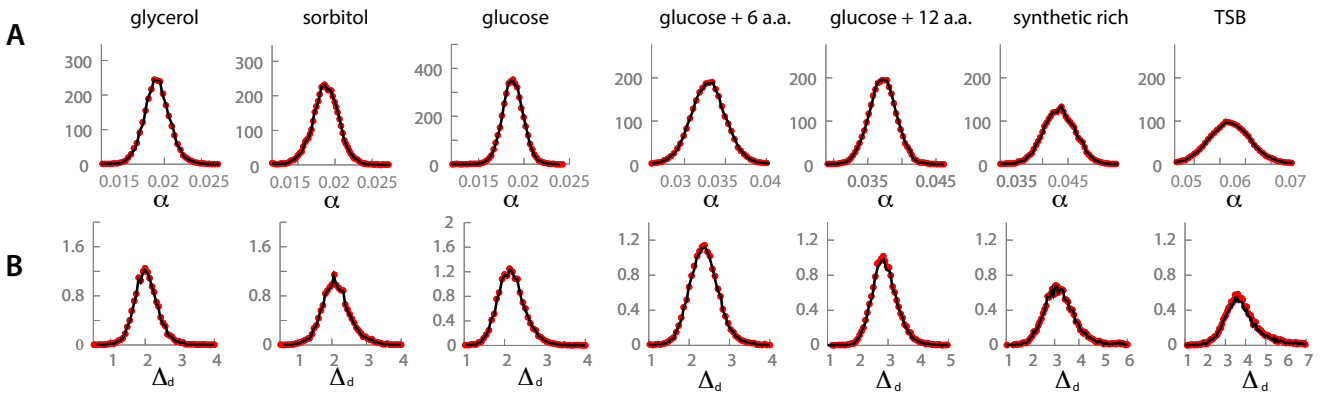


Figure S18. Calibration of the Δ model for the control of the cell size. The elongation rate of the cells is an iid random variable drawn from the *E. coli* experimental distribution, $\rho_\alpha^{ex}(\alpha)$ (red dots) in the plots of panel (A) for the seven different growth conditions presented in Figure S13. Correlations among elongation rates of mother and siblings are weak and thus not taken into account. We verified that results are not modified by including them. The black curves are the results of the numerical simulations. The splitting rate $\gamma(\Delta)$ is computed as detailed in the text (see (15)) from the distribution $\rho_{\Delta_d}^{ex}(\Delta)$ of the increments at division $\Delta_d = s_d - s_b$. In panel (B) we show the experimental distributions (red dots) for the same growth conditions as in panel (A). The curves in black are the results of numerical simulations of the model detailed in the text. Their agreement with the experimental curves confirms that the parameters of the model are appropriately calibrated. Similar curves are obtained for *B. subtilis*.

A. Comparison and consistency with experiments

The comparison with experimental data for the Δ model proceeds as for the sizer mechanism. The elongation rate $g(s)$ and the splitting function $\gamma(\Delta)$ are extracted from experimental data, namely from the distribution of the sizes at division and the distribution of the elongation rates, and we use them to simulate the cell size control process at the level of individual cells. We then compare statistical observables alternative to those used for calibration, in order to assess the validity of the model.

The calibration of the model goes as follows. The elongation rate for individual bacteria is obtained by exponential fits of the experimental curves of size versus time. The probability densities $\rho_\alpha^{ex}(\alpha)$ of the resulting elongation rates in the various growth conditions are shown in Figure S18. In our numerical simulations we draw values of α randomly from $\rho_\alpha^{ex}(\alpha)$, neglecting (as for the sizer case) weak correlations between the elongation rates of mother and daughter cells. The splitting rate $\gamma(\Delta)$ is computed using (15) with the probability of the size increments at division $\Delta_d \equiv s_d - s_b$ read directly from the experimental data $\rho_{\Delta_d}^{ex}(\Delta)$ (see Figure S18).

The distribution of the elongation rates α and the splitting function $\gamma(\Delta)$ extracted from the experimental data as we just described are used to simulate the dynamics of a bacterial colony. Each cycle of elongation of a cell proceeds at the constant (random) rate α and division occurs with the Poissonian rate $\gamma(\Delta)$, which depends on the size increment Δ only. After an initial transient, distributions for the various observables reach a stationary form and the resulting numerical distributions for the added size at division $\Delta_d = s_d - s_b$ and α (used to calibrate the model) compare to the experimental ones as shown in Figure S18.

Distributions of size and age at division. A first test for the Δ model is provided by the curves in Figure S19 showing the agreement of the distributions for the final size s_d and for the generation time $\tau_d = \log_2(s_d/s_b)/\alpha$ (where the values of the various quantities are those of individual bacteria). The corresponding results for *B. subtilis* are shown in Fig. S20. Most importantly, the model by its very definition agrees with the s_b-independent curves of the conditional probability $P(\Delta_d|s_b)$ shown in the main text.

Correlations of the size across generations. An additional test for the model comes from the correlations of the size among genealogically related cells. Some of the correlations below have been calculated in Ref. [22] by a slightly different procedure. Correlations between the added length of the mother Δ_d^M and the daughter, Δ_d^D are experimentally found to be small so we shall neglect them hereafter.

Let us start by the correlation $\mathcal{C}_{dd}^{(p)} \equiv \langle s_d^M s_d^{D(p)} \rangle$ between the size at division of the mother and its p-th generation descendant. For example, $p = 1$ gives the correlation between the sizes at division of mother and daughter

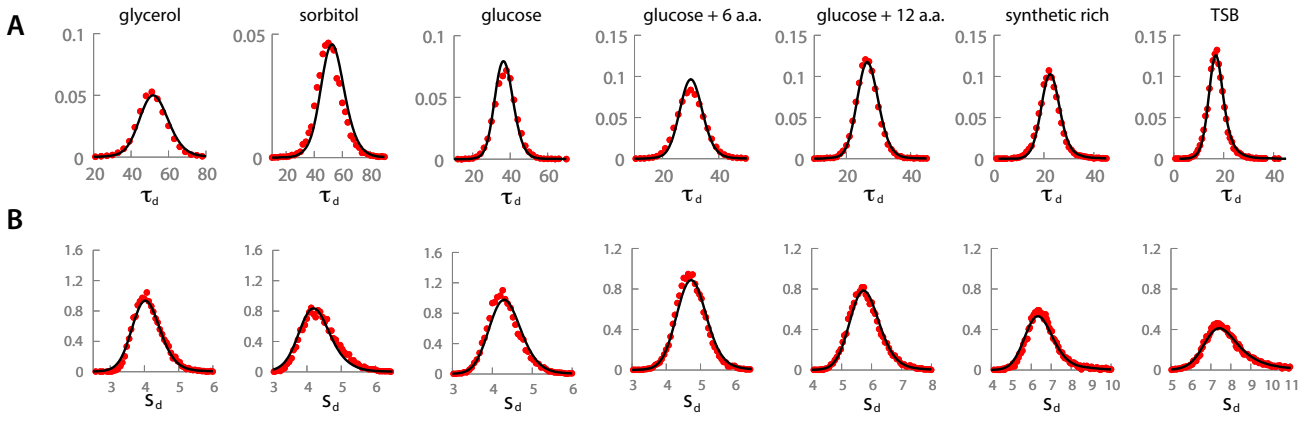


Figure S19. Test of the Δ model for the control of the cell size. The model calibrated as in Figure S18 is simulated numerically and the *E. coli* distributions of the generation time $\tau_d = \log_2(s_d/s_b)/\alpha$ and the size s_d at division of the cells are reported in panels (A) and (B), respectively, for the various growth conditions. Red dots refer to experiments while black curves are the numerics. The agreement of theoretical predictions with experimental data confirms the validity of the Δ mechanism for the control of the cell size.

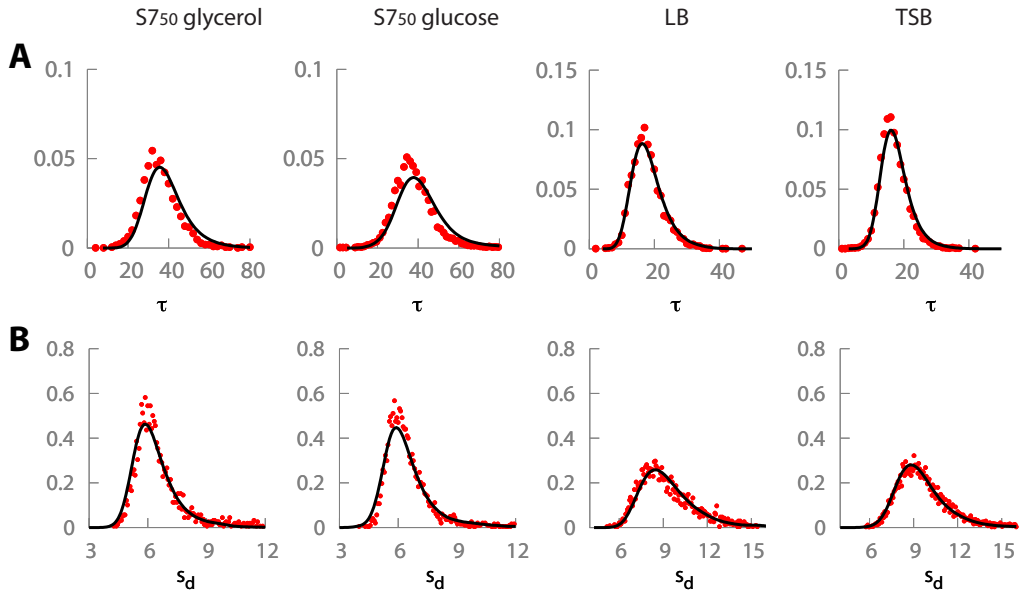


Figure S20. Test of the Δ model for *B. subtilis*. As in Fig. S19, the distributions of the generation time $\tau_d = \log_2(s_d/s_b)/\alpha$ and the size s_d at division of the cells are reported in panels (A) and (B), respectively, for the various growth conditions. Red dots refer to experiments while black curves are the numerical predictions.

cells. The decay of the correlation function is computed by the formula

$$\mathcal{C}_{dd}^{(p+1)} = \left\langle s_d^M s_d^{D(p+1)} \right\rangle = \left\langle s_d^M \left(\frac{s_d^{D(p)}}{2} + \Delta_d^{(p+1)} \right) \right\rangle = \frac{1}{2} \left[\mathcal{C}_{dd}^{(p)} + \langle s_d \rangle^2 \right] . \tag{19}$$

where we have used that $\langle \Delta_d \rangle = \langle s_b \rangle = \langle s_d \rangle/2$, as can be easily derived from $s_d = s_b + \Delta_d$ and $\langle s_b \rangle = \langle s_d \rangle/2$. For the connected part of the correlation function $C_{dd}^{(p)} \equiv \mathcal{C}_{dd}^{(p)} - \langle s_d \rangle^2$, we obtain then

$$C_{dd}^{(p)} \equiv \langle s_d^M s_d^{D(p)} \rangle - \langle s_d^M \rangle \langle s_d^{D(p)} \rangle = \sigma_{s_d}^2 / 2^p \quad \text{where} \quad \sigma_{s_d}^2 = \langle s_d^2 \rangle - \langle s_d \rangle^2 . \tag{20}$$

The comparison with the numerical simulations is shown in Figure S21. By similar arguments we can show that the correlation function for the size at birth $C_{bb}^{(p)} = \sigma_{s_b}^2/2^p$ where the variance of the size at birth $\sigma_{s_b}^2 = \langle s_b^2\rangle - \langle s_b\rangle^2 \simeq \sigma_{s_d}^2/4$ because the size is (roughly) halved at division. The scaling is confirmed in Figure S21. Finally, the mixed correlation $C_{bd}^{(p)} = C_{bb}^{(p)}$ because the added size is statistically independent of the initial size.

The positive correlation between the sizes across generations is quite intuitive to grasp. An ancestor cell bigger than the average will generate progeny that statistically relaxes to the average size, as we showed in the main text and the Figure S9. The fact that the added length is independent of the initial size makes that progeny will inherit only part of the ancestral size, which is successively halved as generations proceed. That accounts for the positive correlations and its 1/2 rate of decay across generations.

Notice that the previous results give for the Pearson correlation coefficient between the size at birth and at division :

$$\frac{\langle s_b\, s_d\rangle - \langle s_b\rangle\langle s_d\rangle}{\sigma_{s_b}\sigma_{s_d}} = \frac{C_{bd}^{(0)}}{\sigma_{s_b}\sigma_{s_d}} = \frac{\sigma_{s_b}}{\sigma_{s_d}} \simeq \frac{1}{2}\,, \tag{21}$$

which accounts for the behavior of experimental data observed in Figure 1D of the main text. It also follows from $s_d = s_b + \Delta_d$ and $s_b \simeq s_d/2$ that for the Δ model

$$\sigma_{s_d}^2 = 4\sigma_{s_b}^2 = \frac{4}{3}\sigma_{\Delta_d}^2\,; \quad \langle s_d\rangle = 2\langle s_b\rangle = 2\langle\Delta_d\rangle\,. \tag{22}$$

We conclude that the coefficients of variation CV of the three quantities are

$$CV_{s_d} \equiv \frac{\sigma_{s_d}}{\langle s_d\rangle} \simeq CV_{s_b} \simeq \frac{1}{\sqrt{3}} CV_{\Delta_d}\,. \tag{23}$$

Predictions are in excellent agreement with the experimental data presented in the main text. The coefficient of variation of s_b is actually slightly larger than CV_{s_d} before s_b is also affected by the noise in the positioning of the septum. We neglected here that source of noise because it is small; when included, it leads to CV_{s_b} being slightly larger than CV_{s_d}, as observed in the main text.

Correlations of generation times. Generation times $\tau_d = 1/\alpha \log_2(s_d/s_b)$ involve a logarithm, which a priori is not obvious to treat. However, we can circumvent the problem using that the average of logarithms involve typical fluctuations around the mean and the observation that the coefficient of variation of s_b and Δ_d are relatively small. We can then derive approximate expressions for those correlations. Neglecting for simplicity the small noise in α, i.e. $\alpha = \langle\alpha\rangle$ below, we expand τ_d and $1/\tau_d$ as

$$\tau_d = \frac{1}{\alpha \ln 2}\ln\left(1+\frac{\Delta_d}{s_b}\right) \simeq \frac{1}{\alpha}\left[1 + \frac{1}{2\ln 2}\frac{\delta\Delta_d}{\langle\Delta_d\rangle} - \frac{1}{2\ln 2}\frac{\delta s_b}{\langle s_b\rangle}\right]\,; \quad \frac{1}{\tau_d} \simeq \alpha\left[1 - \frac{1}{2\ln 2}\frac{\delta\Delta_d}{\langle\Delta_d\rangle} + \frac{1}{2\ln 2}\frac{\delta s_b}{\langle s_b\rangle}\right]\,, \tag{24}$$

where $\delta\Delta_d$ and δs_b denote fluctuations with respect to their respective mean values and we used $\langle s_b\rangle = \langle\Delta_d\rangle$ (see (22)). Second-order terms will not be needed as they cancel out from the correlations we shall compute below.

From (24) we obtain for the mean, the variance and the coefficient of variation of τ_d :

$$\langle\tau_d\rangle \simeq \frac{1}{\alpha}\,; \quad \langle\tau_d^2\rangle - \langle\tau_d\rangle^2 \simeq \frac{1}{4\alpha^2\ln^2 2}\left[\frac{\sigma_{\Delta_d}^2}{\langle\Delta_d\rangle^2} + \frac{\sigma_{s_b}^2}{\langle s_b\rangle^2}\right] = \frac{1}{\alpha^2\ln^2 2}\frac{\sigma_{s_b}^2}{\langle s_b\rangle^2} \quad \Rightarrow \quad CV_{\tau_d} = \frac{CV_{s_b}}{\ln 2} = \frac{CV_{\Delta_d}}{\sqrt{3}\ln 2}\,, \tag{25}$$

where we used (23) in the second equation. The ratios of the coefficients of variation are in excellent agreement with the experimental data presented in Fig. 4B of the main text.

Similarly, we can compute the mean value and the variance for $1/\tau_d$, which read :

$$\left\langle\frac{1}{\tau_d}\right\rangle \simeq \alpha\,; \quad \left\langle\frac{1}{\tau_d^2}\right\rangle - \left\langle\frac{1}{\tau_d}\right\rangle^2 \simeq \frac{\alpha^2}{\ln^2 2}\frac{\sigma_{s_b}^2}{\langle s_b\rangle^2}\,, \tag{26}$$

and derive then

$$\frac{\left\langle\frac{s_b}{\tau_d}\right\rangle - \langle s_b\rangle\left\langle\frac{1}{\tau_d}\right\rangle}{\sigma_{1/\tau_d}^2} \simeq \frac{\langle s_b\rangle\ln 2}{2\alpha}\,. \tag{27}$$

The expression (27) was used in Figure 1C of the main text to fit the dependency of the initial size versus the number of divisions per hour, in given growth conditions. The fitting curve was written as $s_b = c_1 \frac{1}{\tau_d} + c_2$ and it is verified that the constant c_1 is then equal to the correlation (27).

We can also derive the correlation between s_b and the generation time τ_d as

$$\langle s_b \tau_d \rangle - \langle s_b \rangle \langle \tau_d \rangle \simeq -\frac{1}{2\alpha \ln 2} \frac{\sigma^2_{s_b}}{\langle s_b \rangle} \quad \Rightarrow \quad \frac{\langle s_b \tau_d \rangle - \langle s_b \rangle \langle \tau_d \rangle}{\sigma_{s_b} \sigma_{\tau_d}} \simeq -\frac{1}{2}\,, \tag{28}$$

where we used (25). The reason for the negative sign is quite intuitive. Indeed, the elongation at a size s proceeds at the rate $\alpha s \ln 2$, i.e. the longer the cell, the faster it grows. Therefore, if s_b is larger/smaller than the mean it will take less/more time to complete the addition of the size Δ_d (independent of s_b).

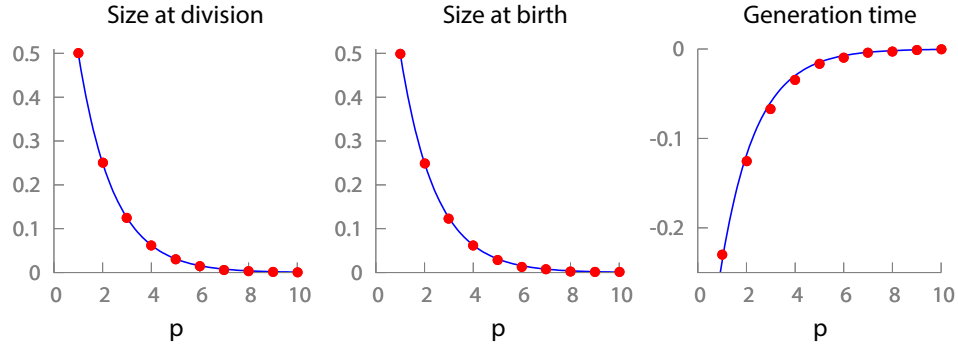


Figure S21. Correlations in the Δ model. We simulate the process under the Δ model and assume no correlation between the Δ_d of the mother and its siblings. The connected correlation function $\langle s_d^M s_d^{D(p)} \rangle - \langle s_d^M \rangle \langle s_d^{D(p)} \rangle$, divided by its value $\sigma^2_{s_d}$ for $p = 0$, is plotted as a function of the generation p in panel (A). The line is the prediction derived in the text 2^{-p} while the dots are numerical values. The corresponding correlation for the newborn size s_b is shown in panel (B). Finally, the connected correlation of the generation time defined in (29) as a function of the generations p is shown in panel (c). The best fit for the decay is the exponential behavior -0.43×2^{-p}, confirming the behavior derived in the text.

Finally, we can explicitly compute the decay of the correlations among the division times across generations. We indicate by τ_d^M the generation time of the mother and by $\tau_d^{D(p)}$ the generation time of a p-th generation descendent. For instance, daughters correspond to $p = 1$. We are interested in the behavior of the correlation

$$C^{(p)}_{\tau_d \tau_d} \equiv \frac{\left\langle \tau_d^M \tau^{D(p)} \right\rangle - \left\langle \tau_d^M \right\rangle \left\langle \tau_d^{D(p)} \right\rangle}{\sigma^2_{\tau_d}}\,, \tag{29}$$

Using (24), we can approximate the correlation by

$$C^{(p)}_{\tau_d \tau_d} \simeq -\frac{1}{\left[CV^2_{\Delta_d} + CV^2_{s_b}\right]} \left\langle \left(\frac{\delta \Delta_d^M}{\langle \Delta_d^M \rangle} - \frac{\delta s_b^M}{\langle s_b^M \rangle} \right) \frac{\delta s_b^{D(p)}}{\langle s_b^{D(p)} \rangle} \right\rangle\,, \tag{30}$$

where we used that $\delta \Delta_d^{D(p)}$ is independent of all other fluctuations in the Δ model and the expression (25) of the variance $\sigma^2_{\tau_d}$ in terms of the coefficients of variation. Note that the mean $\langle s_b^{D(p)} \rangle$ is the usual value of the mean (unaffected by the mother's fluctuations) as corrections would yield higher-order corrections.

The size at birth of a p-th generation descendant is $s_b^{D(p)} = \frac{\left(s_b^M + \Delta_d^M\right)}{2^p} + \sum_{k=1}^{p-1} \frac{\Delta_d^{D(k)}}{2^{p-k}}$. Since the increments $\Delta_d^{D(k)}$ are independent of the mother's fluctuations in the Δ model, we conclude that

$$C^{(p)}_{\tau_d \tau_d} \simeq -\frac{CV^2_{\Delta_d} - CV^2_{s_b}}{CV^2_{\Delta_d} + CV^2_{s_b}} \times \frac{1}{2^p}\,. \tag{31}$$

Using (23), we obtain for the correlation between mother and daughters $C^{(1)}_{\tau_d \tau_d} \sim -\frac{1}{4}$. The behavior (31) is consistent with the experimental data in the Figure S12 and in the main text and with the results of numerical simulations in Figure S21 (small corrections to the constant prefactor are due to the noise in the positioning of the septum and fluctuations in the elongation rate α; we explicitly verified that in their absence the constant agrees indeed with (31)).

Anticorrelations between the generation times of the mother and its descendants are intuitive. Let us consider a given initial size for the mother. As its generation time becomes longer, the mother will divide bigger and will then tend to have descendants with a bigger newborn size. The size at birth and the generation time of a cell tend to be anticorrelated. This follows intuitively from the definition $2^{\alpha\tau_d} = s_d/s_b = 1 + \Delta_d/s_b$ and the fact that the added size is independent of the initial size, i.e. it takes less to add the fixed amount Δ_d if the cell elongates faster (see (28) and (32) for a more formal proof). Combining the two statements above we conclude that the division times of the mother and its descendants are anticorrelated.

Correlations involving exponentials of the generation time. We conclude this Section by computing some correlations involving exponentials of the generation time. The reason is that since $2^{\alpha\tau_d} = s_d/s_b$, correlations do not involve any logarithm and we can demonstrate anticorrelations without any hypothesis on the strength of the fluctuations. We first show that the initial size and $2^{\alpha\tau_d}$ are anticorrelated in the Δ model:

$$\langle s_b 2^{\alpha\tau_d}\rangle - \langle s_b\rangle \langle 2^{\alpha\tau_d}\rangle = \langle s_d\rangle - \langle s_b\rangle \left(1 + \langle\Delta_d\rangle \left\langle \frac{1}{s_b}\right\rangle\right) = \langle s_b\rangle \left(1 - \langle s_b\rangle \left\langle \frac{1}{s_b}\right\rangle\right) \leq 0\,. \tag{32}$$

Here, we used $\langle\Delta_d\rangle = \langle s_b\rangle = \langle s_d\rangle/2$, and the inequality $\langle x\rangle\langle 1/x\rangle \geq 1$ holding for any positive-definite random variable x (as it can be proved for example by the Cauchy-Schwartz inequality $1 = \left\langle \sqrt{x}\,\sqrt{1/x}\right\rangle \leq \langle x\rangle^{1/2}\,\langle 1/x\rangle^{1/2}$).

The exponentials of the division times of the mother and its daughters are also anticorrelated. Indeed:

$$\left\langle 2^{\alpha^M\tau_d^M} 2^{\alpha^D\tau^D}\right\rangle = \left\langle \frac{s_d^M}{s_b^M}\frac{s_d^D}{s_b^D}\right\rangle = 2\left\langle \frac{s_d^D}{s_b^M}\right\rangle = 2\left\langle \frac{s_d^M/2 + \Delta_d^D}{s_b^M}\right\rangle = 1 + 3\,\langle\Delta_d\rangle \left\langle \frac{1}{s_b}\right\rangle, \tag{33}$$

where we have used again that in the Δ model, the added size Δ_d is statistically independent of s_b. Subtracting then the disconnected contribution

$$\left\langle 2^{\alpha^M\tau_d^M}\right\rangle \left\langle 2^{\alpha^D\tau_d^D}\right\rangle = \left\langle \frac{s_d}{s_b}\right\rangle^2 = \left(1 + \langle\Delta_d\rangle \left\langle \frac{1}{s_b}\right\rangle\right)^2, \tag{34}$$

and using again $\langle\Delta_d\rangle = \langle s_b\rangle$, we finally obtain

$$\left\langle 2^{\alpha^M\tau_d^M} 2^{\alpha^D\tau_d^D}\right\rangle - \left\langle 2^{\alpha^M\tau_d^M}\right\rangle \left\langle 2^{\alpha^D\tau_d^D}\right\rangle = \langle s_b\rangle \left\langle \frac{1}{s_b}\right\rangle \left(1 - \langle s_b\rangle \left\langle \frac{1}{s_b}\right\rangle\right) \leq 0\,. \tag{35}$$

V. COLLAPSE OF THE PROBABILITY DISTRIBUTIONS AND SCALING FORMS

It was recently observed in [11] that intra- and inter-species distributions of body sizes have a universal form across many species. The universal form seems to be uniquely determined by the mean of the distribution, i.e., when the various distributions are rescaled by their mean, they tend to collapse onto a unique curve. This recent remark generalizes analogous, classical observations made for different bacterial populations and growth conditions [23; 24], that we also reported in the main text for our experiments on *E. coli* and *B. subtilis*. The aim of this final Section is to present theoretical results showing that the property of scale invariance is common to all size distributions, viz. s_b, s_d and Δ_d. In other words, if one of the three distributions is scale-invariant, the others as well inherit that property.

The distribution $\rho_Z(z)$ of a (generic) random variable Z is scale invariant if it has the form:

$$\rho_Z(z) = \frac{1}{\langle Z\rangle}\phi\left(\frac{z}{\langle Z\rangle}\right), \tag{36}$$

where ϕ is an arbitrary non-negative normalized function. The statistics of Z is supposed to change with conditions, e.g. of the environment or nutrients. The non-trivial content of the scaling form (36) is that when conditions are varied, the distribution will be modified, yet its shape remains invariant when properly rescaled by the new mean value of Z. The form (36) also implies that, when conditions are varied, the normalized moments $\langle Z^p\rangle/\langle Z\rangle^p$ will remain constant. Finally, the scaling form (36) is also equivalent to the statement that the Laplace transform $\mathcal{L}_Z(u)$ of the distribution ρ_Z has the form $\mathcal{L}_Z(u) = \psi(u\langle Z\rangle)$, where u is the Laplace transform variable and ψ is an arbitrary function (respecting the general constraints for the Laplace transform of a probability distribution).

Let us start by showing that if the scaling form (36) holds for the distribution of the cell size either at division s_d or at birth s_b, then it holds also for the other quantity and for the distribution of the added size at division Δ_d.

Indeed, if noise in the halving of the sizes at division is neglected, $s_d = 2s_b$. The distributions for s_d and s_b are then related as $\rho_b(s_b) = 2\rho_d(2s_b)$ and the scaling form of either one of the distributions clearly implies scale-invariance for the other. Moreover, since $s_d = s_b + \Delta_d$ and Δ_d is independent of s_b, we have

$$\mathcal{L}_{s_d}(u) = \mathcal{L}_{s_b}(u) \times \mathcal{L}_{\Delta_d}(u)\,, \tag{37}$$

where $\mathcal{L}$ indicates the Laplace transform of the respective probability distributions. Using $s_d = 2s_b$ we have $\mathcal{L}_{s_b}(2u)/\mathcal{L}_{s_b}(u) = \mathcal{L}_{\Delta_d}(u)$ and therefore the distribution of the added size inherits the scale invariance of s_b (if the latter has it).

If noise in the halving at division of the size of daughters is included, the argument is slightly more involved. We use $s_d = \lambda s_b$, where λ is a random variable centered around 2. We shall assume that the distribution of λ does not change as the means $\langle s_b\rangle$ and $\langle s_d\rangle$ vary with growth conditions. Taking the logarithm of $s_d = \lambda s_b$, we have again a sum and the Laplace transforms of the logarithms of the three variables are therefore related as in (37). The scale-invariant form (36) implies for the Laplace transform of the distribution of $\ln Z$ that $\mathcal{L}_{\ln Z} = \langle Z\rangle^{-u}\psi(u)$, where ψ is arbitrary yet it does not contain $\langle Z\rangle$. Using that $\langle s_d\rangle = 2\langle s_b\rangle = \langle\lambda\rangle\langle s_b\rangle$ (which is valid in any growth condition), one can verify that if either s_b or s_d is scale-invariant, the other variable will inherit that property.

Finally, we show that scale invariance (if present) is also inherited by the size distribution $n(s,\Delta,t)$. The equation for its dynamics is (13), which reduces to

$$kn(s,\Delta) + \partial_s(g(s)n(s,\Delta)) + \partial_\Delta(g(s)n(s,\Delta)) = -\gamma(\Delta)g(s)n(s,\Delta)\,, \tag{38}$$

in the stationary state with the growth rate k defined by (16). Taking $g(s) = \alpha s$ and using $\langle s\rangle \propto \langle\Delta\rangle$, one can verify that a scaling form for $n(s,\Delta)$ is indeed compatible with (38).

A more explicit way to relate $n(s,\Delta)$ to the distributions of Δ and s_b involves the integration of (38) along the characteristics and the tracking of cells from the current time t back to their last division. For that purpose, it is convenient to introduce the age ξ of a cell, as in Section III.B, so that

$$\frac{ds}{d\xi} = g(s)\,, \qquad \frac{d\Delta}{d\xi} = g(s)\,, \tag{39}$$

during the elongation of the cell. The initial size $s_b = s(0)$ and $\Delta(0) = 0$ are the initial conditions. The equation (38) is rewritten as

$$\frac{dn(s(\xi),\Delta(\xi))}{d\xi} = -F(s(\xi),\Delta(\xi))n(s(\xi),\Delta(\xi))\,, \tag{40}$$

where $F(s,\Delta) = k + \partial_s g(s) + \gamma(\Delta)g(s)$. We can then track each cell back to its birth:

$$\begin{aligned} n(s,\Delta) &= n(s-\Delta,0)e^{-\int_0^\xi d\xi' F(s(\xi'),\Delta(\xi'))} = n(s-\Delta,0)e^{-\int_0^\Delta d\Delta' \frac{1}{g(s(\Delta'))}F(s(\Delta'),\Delta')} \\ &= n(s-\Delta,0)e^{-\int_0^\Delta d\Delta'\gamma(\Delta')}e^{-\int_0^\Delta d\Delta' \frac{k+\partial_s g(s(\Delta'))}{g(s(\Delta'))}}\,. \end{aligned} \tag{41}$$

For the exponential elongation rate $g(s) = \alpha s\ln 2$, equation (17) gives $k = \alpha\ln 2$ and we have that

$$n(s,\Delta) \propto \rho_b(s-\Delta)\left(1 - \int_0^\Delta dx\rho_{\Delta_d}(x)\right)e^{-\int_0^\Delta dx\frac{2}{s(0)+x}} = \rho_b(s-\Delta)\left(1 - \int_0^\Delta dx\rho_{\Delta_d}(x)\right)\left(1-\frac{\Delta}{s}\right)^2\,, \tag{42}$$

where we have used (15) to express $e^{-\int_0^\Delta d\Delta'\gamma(\Delta')}$ in terms of ρ_{Δ_d}. it is immediate to verify that if ρ_b and ρ_{Δ_d} are scale-invariant, so will $n(s,\Delta)$ be. Integrating (42) over Δ, we obtain for the marginal $n(s)$ at the steady state:

$$n(s) \propto \int_0^s dx\ \rho_b(s-x)\left(1 - \int_0^x dy\rho_{\Delta_d}(y)\right)\left(1-\frac{x}{s}\right)^2\,. \tag{43}$$

VI. FURTHER DETAILS ON THE RELATION BETWEEN FIGURE 1C AND THE CONSTANCY OF THE CELL MASS AT INITIATION OF REPLICATION

W.D. Donachie proposed in [25] that the mass of the cells (per origin of replication) at the time of initiation of the chromosome replication, is independent of the growth rate. Mass and volume will be used interchangeably here as the density is known to be roughly constant. The argument in [25] is based on two premises:

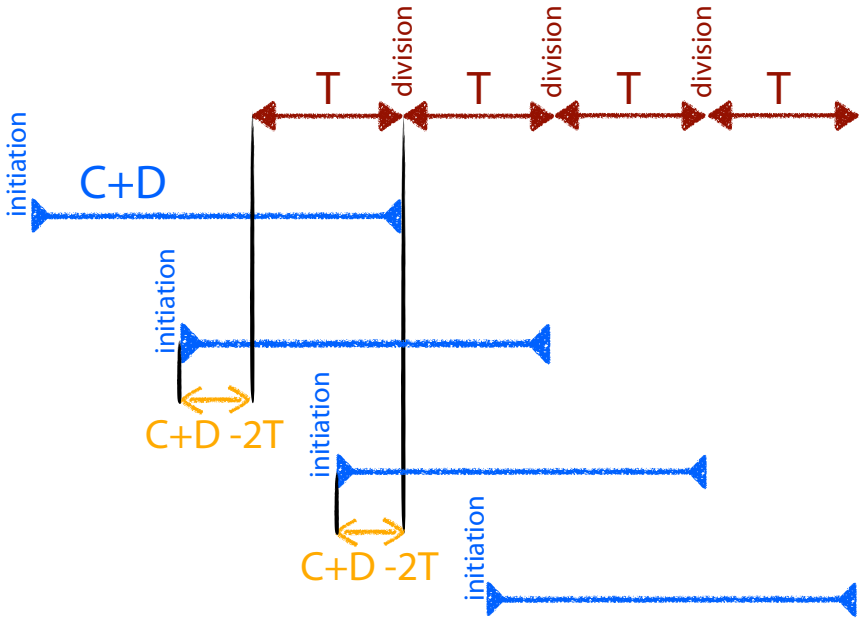


Figure S22. Scheme of the events of initiation of replication and cell division. C denotes the average time for chromosome replication, D is the average time for cell division and the cell average doubling time is denoted here by T. The sum $C+D$ is supposed to be constant for different T. The figure shows that the time separating chromosome initiation to the closest successive cell division is $C+D-2T$. The scheme drawn here refers to those situations where $\lfloor \frac{C+D}{T} \rfloor = 2$, i.e. $2T \leq C+D < 3T$. In the general case where the integer $\lfloor \frac{C+D}{T} \rfloor = r$, i.e. $rT \leq C+D < (r+1)T$, the corresponding time is $C+D-rT$.

(i) The sum of the average time C for chromosome replication and the average time D for cell division is constant in diverse growth conditions. Growth conditions are supposed to be fast, i.e. the cell average doubling time, denoted here by T, satisfies $T \leq C+D$. The doubling time controls the average volume of cells of age ξ as

$$\langle V_\xi \rangle = \langle V_b \rangle \, 2^{\xi/T} \,, \tag{44}$$

where $\langle V_b \rangle$ is the average volume at birth. The constancy of $C+D$ was supported by the experiments in Ref. [26] (with $C+D \simeq 60$ minutes).

(ii) The average volume (mass) of newborn cells scales with T as $\langle V_b \rangle \propto 2^{\frac{\mu}{T}-1}$, where μ is a constant and T is the doubling time defined above. Such behavior of $\langle V_b \rangle$ was obtained in Refs. [4; 5] (with $\mu \simeq 60$ minutes).

Ref. [25] derives the consequence of (i) and (ii) for the time of the initiation of chromosome replication. As shown in Figure S22, the time separating chromosome initiation to the closest successive cell division is $C+D-rT$, where r is the floor of $\frac{C+D}{T}$, i.e. the integer satisfying $rT \leq C+D < (r+1)T$. Therefore, the age of cells at the initiation time is $T-(C+D-rT) = (r+1)T-(C+D)$. It follows from (44) and $\langle V_b \rangle \propto 2^{\frac{\mu}{T}-1}$ that the average volume of cells at the initiation time is

$$\langle V_I \rangle = \langle V_b \rangle \, 2^{r+1-\frac{C+D}{T}} \propto 2^{r-\frac{C+D-\mu}{T}} \,. \tag{45}$$

The remark in Ref. [25] is that the expression (45) reduces to 2^r if $\mu = C+D$, which is precisely the equality obtained in the experiments. Furthermore, it is easy to verify from Figure S22 that the number of origins of replication present at the time of initiation is 2^r. The major ensuing conclusion is that if hypothesis (i) holds and if $\mu = C+D$, i.e. if the average volume of newborn cells scales as

$$\langle V_b \rangle \propto 2^{\frac{C+D}{T}-1} \,, \tag{46}$$

then the average mass per origin of replication at the time of initiation is a constant independent of T. Notice that the average volume $\langle V_b \rangle$ of newborn cells is equal to the average volume $\langle \Delta V_d \rangle$ added at division, which is the key quantity in the Δ model. The average volume added at division is related to the size added at division Δ_d discussed in previous sections by the relation $\Delta V_d = \text{const.} \times w^2 \Delta_d$, where w is the width of the cell (which stays roughly constant over the cycle of growth) and the constant reflects the rod-like shape of *E. coli* cells. The equality $\langle V_b \rangle = \langle \Delta V_d \rangle$ is generally valid since $\langle V_d \rangle = 2\langle V_b \rangle = \langle V_b \rangle + \langle \Delta V_d \rangle$.

Assuming the validity of Donachie's arguments, equation (46) can be used to estimate the value of $C+D$ from the scaling of $\langle V_b \rangle$ *vs* T (see, e.g., [21]). This procedure yields the estimation $C+D \simeq 66$ minutes for *E. coli* data in Fig. 1C of the main text.

VII. CONSTANCY OF TOTAL MASS OF CONSTITUTIVE PROTEINS WITH RESPECT TO GROWTH CONDITION

The starting point of Scott *et al.* [27] is two fundamental correlations between the fraction of ribosomal mass of the proteome, $\phi_{\rm R}$, and the average growth rate, λ, for a given growth condition. The two correlations are

$$\langle \phi_{\rm R} \rangle = \left\langle \phi_{\rm R}^{\rm min} \right\rangle + \alpha \langle \lambda \rangle \tag{47}$$

$$\langle \phi_{\rm R} \rangle = \langle \phi_{\rm R}^{\rm max} \rangle - \beta \left(\langle \lambda \rangle \right) \langle \lambda \rangle , \tag{48}$$

where $\left\langle \phi_{\rm R}^{\rm min} \right\rangle \approx 0.07$ and $\langle \phi_{\rm R}^{\rm max} \rangle \approx 0.55$ for *E. coli* MG1655 and B/r and $\left\langle \phi_{\rm R}^{\rm min} \right\rangle \approx 0.04$ and $\langle \phi_{\rm R}^{\rm max} \rangle \approx 0.49$ for *E. coli* NCM3722. The slope $\alpha = 0.18 \pm 0.1$ is a positive constant for the three different *E. coli* strains. The slope β depends on the growth condition (namely, $\langle \lambda \rangle$), but is constant under a given growth condition. Asymptotically, β can vary between 0 and ∞ as the growth condition changes from rich to poor growth media, respectively.

The above two correlations were obtained from the following experiments. For the first correlation (Eq. 47), $\langle \phi_{\rm R} \rangle$ of a steady-state population was measured for each growth condition $\langle \lambda \rangle$, adding additional information to the growth law (Ref. [4]; Eq. 44; Figure 1C in the main text). For the second correlation (Eq. 48), multiple experiments with varying sub-lethal dosage of chloramphenicol (Cam) were performed to measure $\langle \phi_{\rm R} \rangle$ and $\langle \lambda' \rangle$ for a given growth condition $\langle \lambda \rangle$. Each set of experiments were repeated for other growth conditions.

The invariance of $\phi_{\rm R}^{\rm max} < 1$ is important. Scott *et al.* [27] interpreted this constancy as that the rest of the proteome fraction, $1 - \phi_{\rm R}^{\rm max}$, consists of house-keeping proteins, Q, but in general it can be a basal level of any non-ribosomal proteins. Importantly, since $\phi_{\rm R}^{\rm max}$ is constant, $\phi_{\rm Q} = 1 - \phi_{\rm R}^{\rm max}$ is also constant under all growth conditions. Once we accept this and that the proteome consists of only three types of proteins [ribosomal (R), non-ribosomal (P) and their basal level (Q)], we can predict their relative mass fractions for any growth conditions.

Cell cycle proteins and nutrient transporters belong to P sector, and their total mass fraction of the total proteome can be obtained by solving Eqs. 47 and 48 for $\langle \phi_{\rm P} \rangle = \langle \phi_{\rm R}^{\rm max} \rangle - \langle \phi_{\rm R} \rangle$, leading to

$$\langle \phi_{\rm P} \rangle = \frac{\beta \left(\langle \lambda \rangle \right)}{\alpha + \beta \left(\langle \lambda \rangle \right)} \left(\langle \phi_{\rm R}^{\rm max} \rangle - \left\langle \phi_{\rm R}^{\rm min} \right\rangle \right) \tag{49}$$

$$= \langle \phi_{\rm R}^{\rm max} \rangle - \left\langle \phi_{\rm R}^{\rm min} \right\rangle - \alpha \lambda. \tag{50}$$

The total mass of the P-sector proteins per cell with respect to the per-cell total proteome can be estimated by multiplying the average volume $\langle V \rangle$ of the cells for each growth condition and $\langle \phi_{\rm P} \rangle$. The growth law is given by $\langle V \rangle = 0.24 \cdot \langle V_0 \rangle \, e^{0.81 \langle \lambda \rangle}$ for NCM3722 (this work), $\langle V \rangle = 0.41 \cdot \langle V_0 \rangle \, e^{0.62 \langle \lambda \rangle}$, $\langle V \rangle = 0.33 \cdot \langle V_0 \rangle \, e^{0.70 \langle \lambda \rangle}$, $\langle V \rangle = 0.29 \cdot \langle V_0 \rangle \, e^{0.77 \langle \lambda \rangle}$, and $\langle V \rangle = 0.52 \cdot \langle V_0 \rangle \, e^{0.67 \langle \lambda \rangle}$ for B/r H266, B/r K, B/r A, and BilvA thyA, respectively [28]. $\langle \phi_{\rm P} \rangle$ with the experimental parameters $\alpha, \beta \left(\langle \lambda \rangle \right), \phi_{\rm R}^{\rm min}, \phi_{\rm R}^{\rm max}$ were reported in Scott *et al.* [27].

For the range of the average growth rate $\langle \phi_{\rm P} \rangle$ was measured, the total number of proteins in P sector is remarkably constant in *E. coli* (Figure 4, main text).

REFERENCES

[1] E. Soupene, W. C. van Heeswijk, J. Plumbridge, V. Stewart, D. Bertenthal, H. Lee, G. Prasad, O. Paliy, P. Charernnoppakul and S. Kustu. Physiological studies of Escherichia coli strain MG1655: growth defects and apparent cross-regulation of gene expression. *J Bacteriol* **185**, 5611-5626 (2003).

[2] F. C. Neidhardt, P. L. Bloch, D. F. Smith. Culture medium for enterobacteria. *J. Bact.*, **119(3)**, 736-747 (1974).

[3] P. Wang *et al*. Robust growth of Escherichia coli. *Curr. Biol*. **20**, 1099-1103 (2010).

[4] M. Schaechter, O. Maaløe, N.O. Kjeldgaard. Dependency on medium and temperature of cell size and chemical composition during balanced grown of Salmonella typhimurium. *J Gen Microbiol.*, **19**, 592, (1958).

[5] N.O. Kjeldgaard, O. Maaløe, M. Schaechter. The transition between different physiological states during balanced growth of Salmonella typhimurium. *J Gen Microbiol.*, **19**, 607, (1958).

[6] O. Diekmann, H. Lauwerier, T. Aldenberg, and J. Metz. Growth, fission and the stable size distribution. *J Math Biol.*, **18**, 2, 135–148, (1983).

[7] J. F. Collins and M. H. Richmond, Rate of growth of bacillus cereus between divisions. *J Gen Microbiol*, **28**, 15–33, (1962).

[8] J. J. Tyson and O. Diekmann. Sloppy size control of the cell division cycle. *J Theor Biol*, **118**, 405–426, (1986).

[9] B. Perthame. *Transport Equations in Biology*. Springer Basel AG, (2007).

[10] A. E. Wheals. Size control models of saccharomyces cerevisiae cell proliferation. *Mol Cell Biol*, **2**,(4), 361–368, (1982).

[11] A. Giometto, F. Altermatt, F. Carrara, A. Maritan, and A. Rinaldo. Scaling body size fluctuations. *P Natl Acad Sci*, **03** (2013).

[12] M. Osella, E. Nugent, and M. Cosentino Lagomarsino. Concerted control of escherichia coli cell division. *P Natl Acad Sci*, **02** (2014).

[13] L. Robert, M. Hoffmann, N. Krell, S. Aymerich, J. Robert, and M. Doumic. Division in Escherichia coli is triggered by a size-sensing rather than a timing mechanism. *BMC Biology*, **2**, (1), 17, (2014).

[14] E. O. Powell. Growth rate and generation time of bacteria, with special reference to continuous culture. *J Gen Microbiol*, **15**, (3), 492–511, (1956).

[15] J. J. Tyson. The coordination of cell growth and division —intentional or incidental? *BioEssays*, **2**, (2), 72–77, (1985).

[16] E. Trucco and G. I. Bell. A note on the dispersionless growth law for single cells. *Bull Math Biophys*, **32**, (4), 475-483, (1970).

[17] O. Diekmann, H. J. A. M. Heijmans, and H. R. Thieme. On the stability of the cell size distribution. *J. Math. Biology*, **19**, (2), 227–248, (1984).

[18] W. D. Donachie, K. J. Begg, and M. Vicente. Cell length, cell growth and cell division. *Nature*, **264**, 328–333, (1976).

[19] H.E. Kubitschek. Linear Cell Growth in Escherichia coli. *Biophys. J.*, **8**, 792, (1968).

[20] H.E. Kubitschek. Bilinear Cell Growth in Escherichia coli. *J. Bacteriol.*, **148**, 730, (1981).

[21] S. Cooper. Bacterial Growth and Division. *Academic Press Inc.*, (1991).

[22] A. Amir. Cell size regulation in rod-shaped bacteria. *Phys. Rev. Lett.*, **112**, 208102, (2014).

[23] F.J. Trueba, O.M. Nejassel, and C.L. Woldringh. Generality of the growth kinetics of the average individual cell in different bacterial populations. *J. Bacteriol.*, **150**, 1048, (1982).

[24] H.E. Kubitschek and C.L. Woldringh. Cell elongation and division probability during the Escherichia coli growth cycle. *J. Bacteriol.*, **153**, 1379, (1983).

[25] W.D. Donachie. Relationship between cell size and time of initiation of DNA replication. *Nature*, **219**, 1077, (1968).

[26] S. Cooper, C.E. Helmstetter. Chromosome replication and the division cycle of Escherichia coli B/r. *J Mol Biol.*, **31**, 519, (1968).

[27] M. Scott, C. W. Gunderson, E. M. Mateescu, Z. Zhang, and T. Hwa. Interdependence of Cell Growth and Gene Expression: Origins and Consequences. *Science*, **330** 1099-1102 (2010).

[28] N. Nanninga, C. L. Woldringh. Cell Growth, Genome Duplication, and Cell Division (Ch.9 in Molecular Cytology of Escherichia coli), Academic Press, London (1985).

[29] A. Giometto, F. Altermatt, F. Carrara, A. Maritan, A. Rinaldo. Scaling body size fluctuations. *Proc. Nat. Acad. Sci. USA*, **110(12)**, 4646-4650 (2013).

[30] W. J. Voorn, L. J. H. Koppes and N. B. Grover. Mathematics of cell division in *Escherichia coli*: Comparison between sloppy-size and incremental-size kinetics. *Current Topics in Mol. Gen.*, **1**, 187-194 (1993).

[31] W. J. Voorn, L. J. H. Koppes. Skew or third moment of bacterial generation times. *Arch. Microbiol.* **169(1)**, 43-51 (1997).

[32] S. Jun, S. Taheri-Araghi, Cell-size maintenance: universal strategy revealed. *Trends Microbiol.*, http://dx.doi.org/10.1016/j.tim.2014.12.001 (2014).

VIII. APPENDIX

A. Background: development of the adder model in this work (by Suckjoon Jun)

The central thesis of our work is that bacterial cells grow by a constant size Δ between birth and division. What initially led us to this discovery was the correlations analysis of our published data in [3]. Specifically, we studied the relationship between the newborn cell size (s_b) and the dividing cell size (s_d). Our expectation was that the dividing size would be roughly 2 times the newborn size ($s_d = 2s_b$). However, the linear fit of the correlations systematically deviated from the expected $s_d = 2s_b$. Instead, we found $s_d = s_b + \Delta$ (Figure S23). For almost one year from 2011 to 2012, we repeated our analysis investigating any possible errors in image and data analysis, and always reached the same result, $s_d = s_b + \Delta$. In the mean time, we were extending our experiments in [3] to other growth conditions and also other organisms, especially, *B. subtilis*.

On August 21, 2012, Johan Paulsson visited our lab then at the FAS Center for Systems Biology at Harvard University to discuss cell size control. During our meeting, Taheri-Araghi presented his $s_d = s_b + \Delta$. Upon seeing this, Paulsson made a profound remark that a relationship like "$y = x + \Delta$" exists in his field of plasmid copy-number control. Importantly, its interpretation is that cells control the average plasmid copy number by adding a constant number N_c of plasmids in each generation, irrespective of the number of plasmids the cell is born with. In the dynamical process, it is straightforward to show that this ensures copy-number homeostasis with the average copy number being N_c itself. In the case of cell size control, $s_d = s_b + \Delta$ would then mean that *cells sense neither size nor time*, but add a constant mass Δ irrespective of the initial cell size s_b.

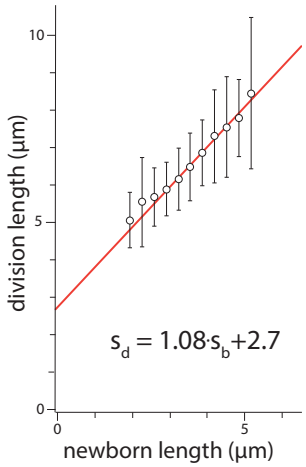


Figure S23. $s_d = s_b + \Delta$.

After our meeting with Paulsson, we had to make a decision. On the one hand, we had a potential text-level discovery that virtually no one in bacterial physiology and cell size control were aware of (with one exception of Koppes and his closest colleagues; see Sec. VIII.B). On the other hand, correlations can suffer from degeneracy of different models that result in the same correlation coefficients, mask important causal relationships, but, more important, we only had data for only one growth condition for two *E. coli* strains (K12 MG1655 and B/r). Due to the limited data, we realized we cannot make a general claim of the constancy of Δ. For this reason, we concluded we needed more data from a wide range of physiological conditions as well as different organisms and strains, although it was clear that a complete investigation was going to take several years of major efforts.

Here the experimental difficulty was that *E. coli* changes its dimensions in response to different growth conditions. This is the growth law established in the 1950s, which states that the average size of the cells increases exponentially with respect to the nutrient-imposed growth rate. In practice, what this means is that each growth condition was a project of its own, because each growth condition required a new microfluidic device optimized for the new cell dimensions. By persevering for more than two years, Taheri-Araghi managed to obtain the complete data of quality and quantity that met our standard for *E. coli*. In the mean time, Sauls was extending our experiments to Gram-positive *B. subtilis* and *E. coli* size mutants in collaboration with Taheri-Araghi and Levin's lab.

Starting in late 2012, an intriguing feature emerged from our new data. That is, all probability distribution functions from different growth conditions collapsed when rescaled by their respective means. This is a highly non-trivial result. Fortunately, in early 2013 Vergassola pointed out an important reference by Rinaldo and colleagues [29] on scaling properties of body size, which I consider as important an insight as the one by Paulsson. We immediately started serious collaborations and divided our main responsibilities between experiment and modeling. This Extended Supplementary Information presents the theoretical description by Vergassola and his post-doc Serena Bradde in its final form (see Sec. III, IV, and V), which goes far beyond our preliminary modeling attempt in 2012 (Sec. VIII.C).

To us, obtaining rigorous and relevant experimental data that can pin down a model was the real issue. The idea of constant Δ was originally and explicitly proposed by Koppes and colleagues in 1993 [30] and 1997 [31], who called it an "incremental model" (see Sec. VIII.B). In the 1997 article, Koppes wrote

> *Ideally, what is required is a comprehensive set of data containing distributions of both size and time between successive cellular divisions. Unfortunately, such measurements still have to be made.*

This is precisely what we have done in our work.

Finally, size homeostasis requires neither perfect adder nor symmetric division. This explains why different organisms such as yeast or size mutants achieve size homeostasis without having to be perfect symmetric adder. We described this general quantitative adder principle in a recent reference [32].

B. WJ Voorn, LJH Koppes, and NB Grover, Mathematics of cell division in Escherichia coli, Curr Top Mol Genetics 1, 187-194, 1993

To our knowledge, this is the first paper in which Koppes and his colleagues fully developed and presented what they called an *incremental model* (which is identical to our adder principle). For example, they explicitly calculated the correlation coefficients between division and newborn size of *E. coli* B/r strains, and obtained 0.52 and 0.56 for two different growth conditions.

Unfortunately, this reference was published in a journal (Curr Top Mol Genetics) that immediately and completely disappeared right after its first volume from the libraries in the United States as well as online databases. Early this year, one of us (SJ) contacted Voorn, who kept a reprint of this paper and kindly sent a high-resolution picture of every page. For its historical importance, and for the reasons explained above, we include the reference here.

Note that the abstract of this paper reads

Two statistical hypotheses for the occurrence of cell division were tested by observed distributions of cell size during steady-state growth. The thirty year old so-called sloppy-size model could be rejected, whereas the newly-developed incremental-size model was accepted (by lack of alternative).

E.H. 1969. Symp. Soc. Gen. Microbiol. 19:237-259
Painter, P.R. 1975. J. Gen. Microbiol. 88:217-220
Painter, P.R. & Marr, A.G. 1967. J. Gen. Microbiol. 48:155-159
Painter, P.R. & Marr, A.G. 1968. Ann. Rev. Microbiol. 22:519-548
Plank, L.D. & Harvey, J.D. 1979. J. Gen. Microbiol. 115:69-77
Powell, E.O. 1955. Biometrika 42:16-44
Powell, E.O. 1956. J. Gen. Microbiol. 15:492-511
Powell, E.O. 1958. J. Gen. Microbiol. 18:382-417
Powell, E.O. 1964. J. Gen. Microbiol. 37:231-249
Powell, E.O. & Errington, F.P. 1963. J. Gen. Microbiol. 31:315-327
Rahn, O. 1932. J. Gen. Physiol. 15: 257-277
Ron, E.Z., Grossman, N. & Helmstetter, C.E. 1977. J. Bacteriol. 129: 569-573
Schaechter, M., Williamson, J.P., Hood, J.R. & Koch, A.L. 1962. J. Gen. Microbiol. 29:421-434
Shields, R. 1978. Nature 273:755-758
Skarstad, K., Boye, E. & Steen, H.B. 1986. EMBO J. 5:1711-1717
Sompayrac, L. & Maaløe, O. 1973. Nature New Biology 241:133-135
Takahashi, M. 1968. J. theor. Biol. 18:195-209
Tyson, J.J. 1987. J. theor. Biol. 126: 381-391
Tyson, J.J. & Diekmann, O. 1986. J. theor. Biol. 118:405-426
Tyson, J.J. & Hannsgen, K.B. 1985. J. theor. Biol. 113:29-62
Verwer, R.W.H. & Nanninga, N. 1980. J. Bacteriol. 144:327-336
Woldringh, C.L., de Jong, M.A., van den Berg, W. & Koppes, L.J.H. 1977. J. Bacteriol. 131:270-279
Zoelen, E.J.J., Saag, P.T. & de Laat, S.W. 1981. J. theor. Biol. 93:1037-1045

Current Topics in Mol. Gen. 1 (1993)

Mathematics of cell division in *Escherichia coli:* Comparison between sloppy-size and incremental-size kinetics

W.J. Voorn*, L.J.H. Koppes and N.B. Grover*****

*Klinical Epidemics & Biostatistics, Meibergdreef 9, NL-1105 AZ Amsterdam, The Netherlands. **Department of Molecular Cytology, Plantage Muidergracht 14, NL-1018 TV Amsterdam, The Netherlands. ***Hubert H. Humphrey Center for Experimental Medicine & Cancer Research, The Hebrew University, Faculty of Medicine, P.O. Box 1172, Jerusalem 91010, Israel

Two statistical hypotheses for the occurrence of cell division were tested by observed distributions of cell size during steady-state growth. The thirty year old so-called sloppy-size model could be rejected, whereas the newly-developed incremental-size model was accepted (by lack of alternative).

INTRODUCTION

Koch and Schaechter (1962) put forward a model for the statistics of cell division according to which a growing bacterial cell divides when it has attained a critical size; this size is considered to be the same, on the average, for all the cells in the culture regardless of their size at birth (Koch and Schaechter 1962). The hypothesis was refined by Powell (1964) and by Harvey et al. (1967), the latter calculating that it predicts quite a broad symmetrical distribution of interdivision times. Marr, Painter and Nilson (1969) then showed that this so-called sloppy-size model fits their synchronous growth data rather poorly.
There is another way of viewing the control of cell division, the incremental-size model. This model states that a growing cell divides after having increased its size by a critical amount; here it is the increment that is considered to be the same, on the average, for all the cells in the culture regardless of their size at birth (Kubitschek and Woldringh 1983). This model predicts a narrow, positively-skewed distribution of interdivision times.
Ideally what is required is a comprehensive data set containing both size and interdivision-time distributions. Unfortunately, no such measurements exist. Instead we have used the most extensive measurements of cell-size distributions available (Cooper 1991, page 657); since the latter were made under the electron microscope on fixed cells, time measurements were of course not possible. Our cell-size measurements were then used to predict various parameters of interdivision time chosen to be as universal as possible. The literature was then searched for reliable interdivision-time distributions. Here we came upon the sixteen experiments of Powell, the problem with the kurtosis (Koch 1962 and 1980), and the three additional measurements of Schaechter et al. (1962).
From measurements of the cell-surface areas of dividing cells in both rapidly and slowly growing cultures of *Escherichia coli* B/r A (Grover et al. 1987), we were able to compute the corresponding interdivision-time distributions for both models under the assumption of exponential cell-surface extension. It appears that the predicted interdivision-time distribution is compatible with incremental but not with sloppy-size kinetics.

MATERIALS AND METHODS

Cultures of *E. coli* B/r A were grown

to steady state in either nutrient broth (mass doubling time 21 min) or in simple alanine medium (doubling time 125 min), fixed in osmium tetraoxide, and prepared by agar filtration (Woldringh et al. 1977). Length and diameter were measured from electron micrographs (Grover et al. 1987), and cell-surface area computed by assuming an idealized geometry of right circular cylinders with hemispherical polar caps. In what follows, size will be used to denote cell-surface area arrived at in this way.
The relevant equations for the sloppy-size model have been published by Powell (1964):

$$q_\tau = A / \sum_{j=0}^{b-1} (b+j)^{-1}$$

and

$$\gamma = 2 \sum_{j=0}^{b-1} (b+j)^{-3} / A^3$$

where

$$A^2 = 2 q_y^2 + q_w^2$$

τ is the interdivision time, q the coefficient of variation, γ the skewness, y the cell size at division, w the ratio of daughter to mother cell size at division, and:

$$b = \tfrac{1}{2} (q_w^{-2} - 1)$$

is the parameter of the symmetrical beta distribution of w (Koppes and Grover 1992).
We were able to derive analogous expressions for the incremental-size model, but they proved to be intricate (Appendix). Instead, the moments of the distribution were calculated under the assumption of a discrete unimodal distribution of incremental size, $h(z)$, with 47 size classes. Various alternative designations for $h(z)$ appear in the literature: life-length distribution (Kubitschek and Woldringh 1983), distribution of interdivision sizes, and division-probability function (H.J.A.M. Meijsman 1986).
In the incremental-size model the increase in size z between birth and division depends not on the size of the newborn cell x nor on its future size at division y but only on the increment $y-x$. Thus the size distribution of dividing cells $\varphi(y)$ assumes the form:

$$\varphi(y) = 2y^{-1} \int_0^y x.\psi(x).h(y-x).dx ,$$

where $\psi(x)$ is the size distribution of newborn cells, the factor $2x/y$ being a consequence of the fact that cell numbers in steady-state populations increase exponentially (Painter and Marr 1968). This equation complements the one first derived by Powell for the cell-size distribution of newborn cells (Koppes and Grover 1992). For incremental growth in which individual cells increase in size exponentially, then, the distribution of interdivision times τ is given by:

$$f_i(\tau) = \frac{\ln(2)}{\hat{\tau}} . 2^{\tau/\hat{\tau}} \int_0^\infty x.\psi(x).h(x2^{\tau/\hat{\tau}} - x)dx ,$$

where $\hat{\tau}$ ($= \ln(2)/\mu$, the specific growth rate) is the doubling time of the culture.
The first three moments of $f_i(\tau)$ were calculated and from these the skewness γ and form factor β ($= \gamma/q$). Finally, β and γ were compared with the experimental data by means of an exact nonparametric test: the chi-square test of Wilcoxon (Verwer and Nanninga 1980).
Several data sets were used. The first (data set 1) is the sixteen measurements reported by Powell (1958). Three of those (designated by Powell as Ba 4, Cp 1 and Bf 3) show too high a kurtosis (larger than 6), probably due to the presence of anomalous cells in the tails (Koch 1980), and were either dropped (data set 2) or replaced (data set 3) by those from Schaechter et al. (1962).

RESULTS

Table 1 summarizes the data on interdivision-time distributions found in the literature. They are the only published measurements on wildtype bacterial cells in steady-state exponential growth in which the variance is not excessive ($q < 25\%$). It has been

Table 1. Parameters (mean ± standard error) of observed interdivision-time distributions

Data set	n	$\bar{\tau}$[a] (min)	q[b]	skewness	sk/q
1	16	43.1 ± 5.3	0.180 ± 0.007	0.746 ± 0.098	4.08 ± 0.48
2	13	45.2 ± 6.2	0.176 ± 0.007	0.600 ± 0.069	3.41 ± 0.39
3	16	43.2 ± 5.1	0.178 ± 0.006	0.532 ± 0.072	3.02 ± 0.41

[a] average interdivision time
[b] coefficient of variation

Table 2. Parameters of calculated interdivision-size(time) distributions

Sample	n	distribution	method	mean(units)	q	skewness	sk/q
a	769[a]	$\varphi(y)$	moments	4.81 μm^3	0.091	0.673	7.42
		$h(z)$	likelihood	2.42 μm^3	0.147	0.784	5.32
		$h(z)$	moments	2.42 μm^3	0.151	1.137	7.51
		$f_i(\tau)$	likelihood	125.2 min	0.128	0.410	3.21
		$f_i(\tau)$	moments	126.5 min	0.133	0.924	6.96
		$f_s(\tau)$	moments	125.2 min	0.199	0.0081	0.0407
b	1162[b]	$\varphi(y)$	moments	16.94 μm^3	0.117	0.559	4.78
		$h(z)$	likelihood	8.56 μm^3	0.197	0.671	3.40
		$h(z)$	moments	8.57 μm^3	0.201	0.817	4.07
		$f_i(\tau)$	likelihood	21.0 min	0.166	0.377	2.28
		$f_i(\tau)$	moments	21.4 min	0.172	0.834	4.85
		$f_s(\tau)$	moments	21.0 min	0.246	0.0021	0.0085

[a] precision of division into siblings of about equal size: 5.2%
[b] precision of division into siblings of about equal size: 4.3%

suggested that complex media give rise to a greater dispersion of interdivision times than do simple media (Powell and Errington 1963), and in an attempt to correct for such an effect, we defined the dimensionless form factor β as γ/q (Ash and Powell 1962). From q_x, q_w, and q_τ, we were able (to be published) to calculate the correlation coefficient r between size at birth and size at division: 0.52 and 0.56 for interdivision times of 21 min and 125 min, respectively. Such high correlation has been reported before (Errington et al. 1965; Koppes et al. 1980) and argues against sloppy-size control.
Table 2 lists the observed moments of the size distributions at division of two large samples of E. coli B/r A cells growing under steady-state conditions (Grover et al. 1987), together

Table 3. Comparison between incremental-size and sloppy-size models

Model / data set[a]		Incremental-size 1	2	3	Sloppy-size 1	2	3
moment	Sample[b]	Probability[c]			Probability[c]		
skewness	a	+	+	+	-	-	-
	b	+	+	+	-	-	-
sk/q	a	+	-	-	-	-	-
	b	+	+	+	-	-	-

[a] Table 1
[b] Table 2
[c] Probability that difference between predicted and experimental values due to chance, as determined by the chi-square test for symmetry (Verwer and Nanninga, 1980). On the assumption of equivalence of both methods, i.e. moments and maximum likelihood (see Appendix), the average predicted parameter (Table 2) was tested by observed values (Table 1) at $\alpha = 0.005$ significancy.

with the interdivision moments predicted by the two models. As expected (Powell and Errington 1963), q_γ and q_τ are smaller at the slower growth rate, perhaps because the biochemical processes leading to division are more complicated. (The average size at division is slightly less than twice the average size increment owing to the exponentially increasing cell number in steady-state cultures). The hypothesis of Koch and Schaechter (1962) predicts a coefficient of variation for $f_\tau(\tau)$ about twice that for $p(\gamma)$ and a negligible skewness; in the incremental size model, $f_\tau(\tau)$ is less variable and positively skewed. This pronounced difference in the shape of the $f(\tau)$ distributions is reflected in the value of their form factor β.
Finally, Table 3 compares experimental and predicted shapes of $f(\tau)$ for both models using an exact nonparametric test, the chi-square test according to Pearson (Verwer and Nanninga 1980). It is clear that the sloppy-size model can be rejected for all three data sets, in both samples, whereas the incremental size model fares much better especially as regards skewness. As expected, the form factor is more discriminatory for shapes, because it rejects the incremental-size hypothesis in two instances. The latter finding could be due to both the upset quality of data sets 2 and 3, and the slow growth rate of the bacteria in sample a, in contrast to that in samples 2 and 3, partly invalidating the comparison. We conclude, therefore, that interdivision time is not simply derived from cell size at division, but rather from interdivision size.

DISCUSSION

Published data. The interdivision-time distribution has been estimated from steady-state liquid culture and found to be positively skewed (Harvey et al. 1967; Marr et al. 1969). This method appears less sensitive than the earlier, direct microscopic observations of individual cells, however, because the cumulative nature of the data can be expected to distort whatever skewness may be present (Harvey 1972; Plank and Harvey 1979).

The earliest distributions of bacterial interdivision time were obtained by Kelly and Rahn (1932), who recorded the family tree of individual cells under the microscope, primarily to estimate mortality. Their skewed distributions show low kurtosis (less than 3) but high coefficient of variation (more than 30%) owing mainly to the low resolution (5 min intervals). This technique was improved considerably by Powell (1955 and 1956) who constructed a special culture chamber that permitted balanced growth under the microscope. The distributions obtained this way are regarded as the most reliable available, and show both low coefficient of variation and low kurtosis (Powell, 1958). Later measurements by Schaechter et al. (1962) and by Kubitschek (1962) confirm the positive skew.
Distributional shape. In contrast to above, we measured the dimensions of dry, dead, flattened bacteria sampled from steady-state cultures, under the electron microscope. Despite this apparent drawback, our cell sizes were found to be approximately proportional to the in vivo cell sizes (Woldringh et al. 1977). And so whereas the precise form of the in vivo interdivision-time distribution remains unknown, its form factor β probably lies between 2 and 8, allowing us to reject sloppy-size division control (Koch and Schaechter 1962; Tyson and Hannsgen 1985; Tyson and Diekmann 1986). In other words, a bacterial cell cannot gauge its absolute size, as reported before (Ron et al. 1977; Grossman and Ron 1989).
Previous models. In practice, incremental growth implies that many of the old cell-cycle ideas, be they merely descriptive (Ash and Powell 1962; Koch 1966; Kubitschek 1971) or more quantitative, and whether they involve one (Shields 1978; Bremer 1982 and 1986; Domach and Shuler 1984), two (Brooks et al. 1980; van Zoelen et al. 1981), or many random events (Rahn 1932; Kendall 1948 and 1952; Takahashi 1968), are still valid but should be applied to the distribution of interdivision size $h(z)$ rather than time $f(\tau)$. The basic concept in the hypothesis of Koch and Schaechter (1962) was that size, rather than cell age or time, ought to be considered the relevant variable (Powell 1964; Errington et al. 1965; Koch 1980). Here we show that it is not absolute size but relative size, i.e. size corrected for the size of the cell at birth. Thus, while growth rate is independent of cell age, the division probability (Bell 1968; Kubitschek and Woldringh 1983) is a function of both age and size.
Control of cell division presupposes the existence of underlying molecular events in the division pathway. Such hypothetical events could occur either with constant probability, exponential kinetics (Tyson 1987), or with a probability that is inversely proportional to cell size, hyperbolic kinetics (Berg and Blomberg 1977), depending on genotype (Gottesman 1989) In either case, the resulting interdivision-time distribution will not be symmetrical (Painter and Marr 1967) but Yule-like.
Familial correlations. One of the reasons for the survival of Koch and Schaechter's sloppy-size model was its ability to predict with reasonable accuracy the correlation between the interdivision time of parents and that of their progeny, -0.45, and between the siblings themselves, +0.35 (Koch and Schaechter 1962; Powell and Errington 1963; Kubitschek 1966). The correlations predicted by the incremental-size model are lower, -0.20 and +0.20, respectively (unpublished results). A closer look at the experimental data, however, suggests that discrimination between these two sets of values is not at all reliable: microscopic observation artificially raises the correlation because of random errors in the registry of cell separation (Painter 1975) while Gaussian filtering of synchronous-growth curves tends to do the same owing to the inherent assumption of symmetry (Plank and Harvey 1979).
Chromosomal replication. It has recently been demonstrated quite convincingly, by using a chromosomal run-away replication mutant, that the DNA-replication and the cell-division cycles are interdependent, i.e. they are separate consequences of the same growth (Bernander and Nordström 1990). It is not unreasonable, therefore, to assume that the replication of plas-

mid F (Koppes 1992) and of the cell chromosome (Sompayrac and Maaløe 1973) are timed by the same incremental-size mechanism as that proposed here for cell division. Sibling correlation between interinitiation times can then be expected to contribute to the tight synchrony of initiation observed to occur at multiple origins in the same cell (Koppes et al. 1978; Skarstad et al. 1986).

Growth law. In a previous article (Grover et al. 1987), we compared eight growth models for surface extension during unrestricted balanced growth: two exponential and six bilinear. Two in the latter group were rejected on statistical grounds, the fit to the experimental data being inadequate. For theoretical reasons, we can now discard two more: the two-parameter models in which the growth rate doubles by a sloppy-size mechanism. There are thus four surviving one-parameter models, two bilinear and two exponential ones. It may be that individual bacterial cells in continuous cultures grow bilinearly with discrete rate doublings at nuclear division (Kubitschek 1986; Kubitschek and Pai 1988), but the situation in batch cultures remains equivocal. We chose simple exponential extension of the cell-surface area (Kubitschek and Woldringh 1983) for computational convenience, the introduction of the extra parameter not being justified by the data; the difference between it and the bilinear models is very small (Grover et al. 1987) and not expected to detract from the arguments presented here involving the skewness of the interdivision-time distribution.

Continuum model. A commonly held view is that prokaryotes and eukaryotes are quite similar as far as the basic features of their cell cycles are concerned (Cooper 1991). It is tempting, therefore, to consider extending the mechanism of incremental-size growth to simple eukaryotic cells such as Saccharomyces cerevisiae and Schizosaccharomyces pombe; the analysis there can be expected to be much more complicated, however, because of the inequality of cell division and the many subpopulations involved (Fantes 1977).

Synchronous growth. The rejection of the sloppy-size mechanism has implications for the way in which cell-cycle-specific biochemical reactions should be investigated. Apart from possible trauma caused by the synchronization procedure itself, selection by age (membrane elution) is as poorly defined as selection of newborn cells by size (centrifugal elutriation) as far as the cell cycle is concerned. It would appear that the only reliable method is double selection, by both age and size (Koppes et al. 1980).

APPENDIX

Maximum likelihood method. The distribution of the size ratio w=sibling/parent is assumed to be B(b;w), where b denotes the single parameter of the beta division-imprecision distribution (Koppes and Grover 1992). With known coefficient of variation, it is then straightforward to calculate from the incremental-size equation for given function h(z) the birth and division functions $\psi(x)$ and $\varphi(y)$, respectively. Sufficient accuracy was obtained by taking as unit of cell size about one tenth of the standard deviation of the observed distribution of parental size.
The discrete division probabilities h(z) that were tried to maximize the likelihood of the observed parental size frequencies, were allowed to have 47 points with positive probabilities and were only restricted to be unimodal. Several starting points in a (47-1) - dimensional parameter space of h(z)-probabilities were tried, but every time the same point of convergence was found.
Method of moments. Given the moments $\bar{y}_k$ of the distribution of cell size at division. It is then possible to calculate the moments $\bar{z}_k$ of the distribution h(z) of interdivision size z from:

$$\bar{y}_k = 2 \sum_{j=0}^{k-1} \binom{k}{j} \cdot \bar{x}_{k-j} \cdot \bar{z}_j, \quad k = 1,2, \ldots$$

(An elementary proof is by Laplace transformation of the incremental-size equation, derivation and substitution of zero argument). From Powell's equation it follows that:

$$\bar{x}_k = b_k \cdot \bar{y}_k \qquad k = 1, 2, \ldots$$

where $b_1 = \frac{1}{2}$ and

$$b_k = b_{k-1}(b+k-1)/(2b+k-1) \qquad k = 1, 2, \ldots$$

Let g(t) denote the distribution of interdivision-size ratio $t = z/x = \exp(\mu\tau) - 1$. Then:

$$g(t) = \int_0^\infty x \cdot \psi(x) \cdot h(xt) \cdot dx$$

Mellin transformation gives the moments $\bar{t}_k$ as product of those of h(z) and $\varphi(y)$:

$$\bar{t}_k = \bar{x}_{-k} \cdot \bar{z}_k = b_{-k} \cdot \bar{y}_{-k} \cdot \bar{z}_k \qquad k = 1,2$$

where $b_0 = 1$ and

$$b_{-k} = b_{1-k} \cdot (2b-k)/(b-k) \qquad k = 1, 2, \ldots$$

Finally, the approximate moments $\bar{\tau}_k$ are found by linear transformation of $g(t)$: $\tau \approx t/2\ln(2) + 1 - 1/2\ln(2)$. This transformation only affects the mean and standard deviation, but not the coefficients of skewness and kurtosis.

ACKNOWLEDGEMENTS

We wish to express our thanks to dr. Nanninga and to dr. Woldringh for their kindness and warm hospitality, respectivily.
This work was supported in part by grants from the Swedish Natural Science Research Council (no.3634) and the Swedish Cancer Society (no.479) to dr. Nordstrom, and by an EMBO short term fellowship (no.5565) to L.K.

REFERENCES

Ash, R. & Powell, E.O. 1962. Nature 195:770-772
Bell, G.I. 1968. Biophys. J. 8:431-444
Berg, O.G. & Blomberg, C. 1977. J. theor. Biol. 67:523-533
Bernander, R. & Nordstrom, K. 1990. Cell 60:365-374
Bremer, H. 1982. J. Gen. Microbiol. 128 2865-2876
Bremer, H. 1986. J. theor. Biol. 118: 351-365
Brooks, R.F., Bennett, D.C. & Smith, J.A. 1980. Cell 19:493-504
Cooper, S. 1991. Microbiol. Rev. 55: 640-674
Domach, M.M. & Shuler, M.L. 1984. J. theor. Biol. 100:577-585
Errington, F.P., Powell, E.O. & Thompson, N. 1965. J. Gen. Microbiol. 39: 109-123
Fantes, P.A. 1977. J. Cell Sci. 24: 51-67
Gottesman, S. 1989. Ann. Rev. Genet. 23:163-198
Grossman, N. & Ron, E.Z. 1989. J. Bacteriol.171:80-82
Grover, N.B., Woldringh, C.L. & Koppes L.J.H. 1987. J. theor. Biol. 129: 337-348
Harvey, J.D. 1972. J. Gen. Microbiol. 70:99-107
Harvey, R.J., Marr, A.G. & Painter, P.R. 1967. J. Bacteriol.93:605-617
Heijmans, H.J.A.M. 1986. In: Metz, J.A.J. & Diekmann, O. (Eds.) Springer Lecture Notes in Biomathematics 68:185-202
Kelly, C.D. & Rahn, O. 1932. J. Bact. 23:147-153
Kendall, D.G. 1948. Biometrika 35: 316-330
Kendall, D.G. 1952. J. Royal Stat. Soc. (Britain) 14:41-44
Koch, A.L. 1966. J. theor. Biol. 12: 276-290
Koch, A.L. 1980. J. Supramol. Struct. Suppl. 4:154
Koch, A.L. & Schaechter, M. 1962. J. Gen. Microbiol. 29:435-454
Koppes, L.J.H. 1992. J. Bacteriol. 174: 2121-2123
Koppes, L.J.H. & Grover, N.B. 1992. Arch. Microbiol. 157:402-405
Koppes, L.J.H., Overbeeke, N. & Nanninga, N. 1978. J. Bacteriol. 133: 1053-1061
Koppes, L.J.H., Meyer, M., Oonk, H.B., de Jong, M.A. & Nanninga, N. 1980. J. Bacteriol. 143:1241-1252
Kubitschek, H.E. 1962. Exp. Cell Res. 26:439-450
Kubitschek, H.E. 1966. Exp. Cell Res. 43:30-38
Kubitschek, H.E. 1971. Cell Tissue Kinet. 4:113-122
Kubitschek, H.E. 1986. J. Bacteriol. 168:613-618
Kubitschek, H.E. & Pai, S.R. 1988. J. Bacteriol. 170:431-435
Kubitschek, H.E. & Woldringh, C.L. 1983. J. Bacteriol. 153:1379-1387
Marr, A.G., Painter, P.R. & Nilson,

C. Preliminary theory note by Taheri-Araghi, September 28, 2012

For readers interested in the initial development on the theory front, we included an original preliminary theory note written by Taheri-Araghi. This summarizes our understanding of the consequences of constancy of Δ as of September 2012 soon after our meeting with Johan Paulsson. This was primarily at the correlations level based on limited date sets in Wang *et al.* [3]. As explained in Section VIII.A, a complete, independent theoretical framework was established only after Vergassola and Bradde joined with the on-going acquisition of new experimental data.

Growth and Cell Size Control in *Escherichia coli*

Sattar Taheri-Araghi, Ariel Amir, Suckjoon Jun

September 28, 2012

0.1 Introduction

The life cycle of an *Eschirichia coli* cell from an outsiders point of view consist of growth in size, ended by division of the cell in half. Complex molecular mechanisms drive and control this cell cycle, on which our knowledge is still at early stages. A precise quantitative picture of cell size evolution over a cell cycle, and coordination between size and division can shape our fundamental understanding on growth, cell cycle, and the molecular circuits that control this process under the hood.

In this study, we investigate the coordination between cell size and division of *E. coli* during steady state growth. Specifically, we challenge the general belief that cells "*double up*" their size during their life cycle. We analyzed growth of *thousands* of *E. coli* cells in "mother machine", a high throughput continues culture microfluidic device.

Quantitative analysis of cell level parameters, such as size, elongation rate and generation time, reveals a remarkable, yet simple, coordination between division and cell size: the cell mass increase during each cell cycle is constant regardless of the cell size at birth. That is, *E. coli* cells somehow measure the amount of mass they accumulate after birth. Once reaching a critical value, depending on the growth condition, cells divide into two viable daughters. This finding argues against models that propose cells measure their size to trigger division and models that picture a constant timer controlling cell division. However, at a population average, it is consistent with that the cell size at division is twice the size at birth.

0.2 Steady state growth from a single cell perspective

We used mother machine to collect data on the growth and cell cycle of 7,300 single cells in the steady state growth. Images of growing cell, collected by time lapse microscopy, analyzed by custom developed image analysis software and single cell parameters are quantified essentially in terms of four independently measured parameters:

- ℓ: the length of a cell at birth (newborn length)
- L: the length of a cell just before division (division length)
- T: the generation time of a cell
- λ: the elongation rate, as defined for one dimensional growth of a cell below.

In this work we, put an emphasis on precise quantitative analysis of the data obtained from the experiment. Specifically, we investigate not only the average and fluctuations of different parameter, but also the correlation among them

and eventually develop an analytical model to explain the results and provide insights on cell size control mechanism in *E. coli* cells.

0.2.1 How big is the variations in cell size?

A previous study has shown *E. coli* cells have a robust growth mechanism over hundreds of generation where the their elongation rate remain constant regardless of the replication age of cell. In other words, cells do not get slow in growth as they age, even though the probability of fillamentation becomes larger. While cells maintain a constant elongation rate and robust growth, their size and generation time do fluctuate over 20% of the average value. This level of size variations indicates that there should be a mechanism that controls the fluctuations in a populations of the cells. An obvious question in this context is that what happens to those cells who are much larger, or smaller, than the average population size? How do they catch up with the average size?

Fig. 1 shows histograms of cell size at birth (left panel) and division length (right panel). The distributions are not symmetric, later in this study we show that they are approximately lon-normal distributions.

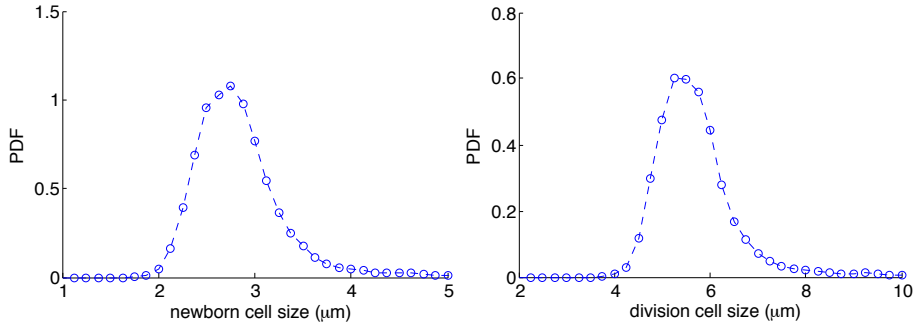


Figure 1: Normalized distribution of newborn cell length (left panel) and cell length at division right panel. Distributions are asymmetric (non Gaussian).

Fig. 2 demonstrates an scatter plot of division length as a function of newborn length. The red line that passes through the scatter points is the best line fit to this data points. There are two points in this graph that deserve some discussion. First, the division length is correlated with the newborn length. This implies that, if a cell is born larger than the average, its size at division will be likely larger than average too, yielding larger than average daughter cells. While this may seem inconsistent with an active mechanism controlling the cell size, the second thing we want to point out is the slope of the fitted line. Consider, hypothetically, the slope of the line was two. That would imply the size of a cell at division is twice the newborn size resulting in daughter cells with same size as previous generation. A slope smaller than two, however, indicates that daughter cells have a size closer to the average size compared to the newborn cells of the previous generation.

What is evidence here is that, cells who deviate from average size do get back to the average size but this does not happen in one generation. In the next section, we show how cells coordinate growth to catch up with an average size of the cell.

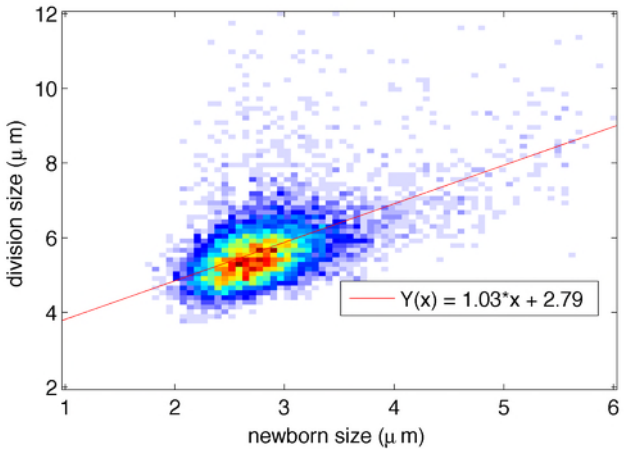


Figure 2: Length of cells right before division as a function of their newborn length for a total number of 7,300 mother cells in steady state growth. The slope of the fitted line, 1.03, is an indication of an active mechanism controlling cell size as opposed to *doubling* cell mass during a cycle.

0.2.2 How do cells control their size?

At a cell level observation, we can quantify the growth in terms of elongation rate and generation time. Elongation rate denotes how *fast* cells grow and generation time denotes *how long* they grow before division. Here, we illustrated the correlation of these two parameters with newborn cell size.

Figure 3 depicts scatter plots of cell elongation rate (left panel) and generation time (right panel) versus newborn cell size. Elongation rate does not show any significant correlation with cell size. On the other hand, generation time shows a strong negative correlation with newborn size.

To further analyze this observation, we have binned cells based on their newborn size and plotted the average elongation rate and average generation time in each bin. Figure 4 shows the average elongation rate (left panel) and average generation time (right panel) of each bin as a function of newborn cell size. The error bars show the standard deviation in each bin. For both elongation rate and generation time, the standard deviations seems similar in different bins.

One important observation here is that normalized standard deviation for generation time is larger than that of elongation rate. For a quantitative com-

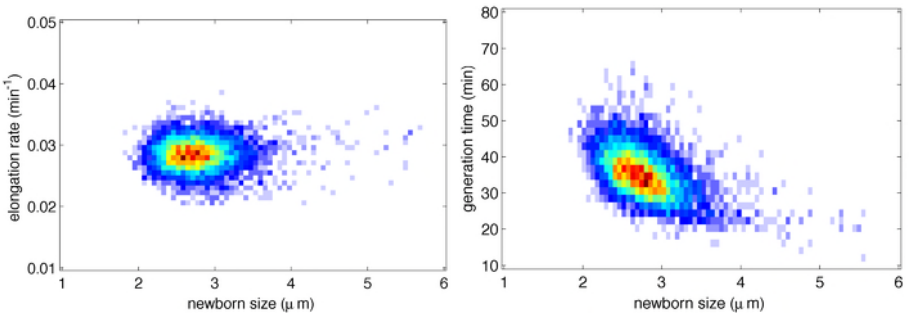


Figure 3: Scatter plot of elongation rate (left panel) and generation time (right panel) as a function of newborn cell size. Elongation rate does not show any correlation with size while cells generation time is negatively correlation with their size at birth.

parison, we have chosen one bin and normalized a histogram of both variables as plotted in figure 5. The generation time has a standard variation of $\sigma = 0.144$ while the elongation rate standard deviation is $sigma = 0.077$. In such stochastic systems, the parameter with a higher fluctuation eventually dominates the systems fluctuation. In the next section, we use this fact to reduce the complexity of our theoretical model considering all the fluctuations in size come from variations in generation times.

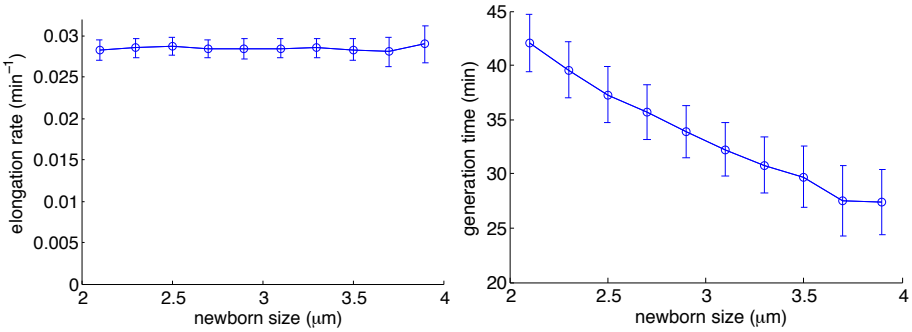


Figure 4: Elongation rate and generation time as a function of newborn cell size.

0.2.3 Experiment condition

The experiment is conducted on SJ288 strain in rich MOPS media at 30C. Mother machine device was made from SJ110 microchannel set. The growth media was continuously injected to the device at 3,000 μl/hr. Microscopy is performed with Nikon Ti using 100X oil immersion lens with bright field phase contrast. Images from the mother machine device acquired at 10 different fields

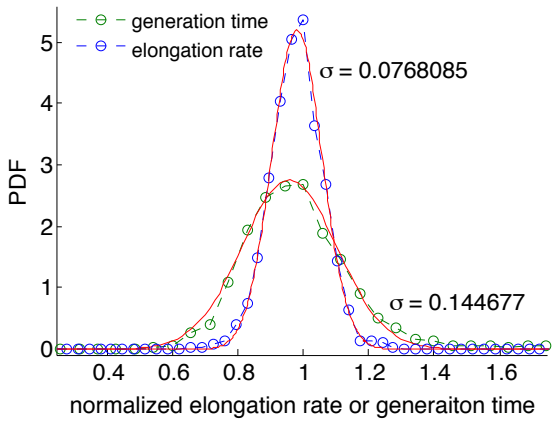


Figure 5: Normalized elongation rate and generation histogram for a bin of figure 4. The generation time has a wider distribution with $\sigma = 0.145$ compared to elongation rate with $\sigma = 0.077$.

of views at 2 min intervals. LED illumination was used at 100% power with exposure time of 70ms for each image acquisition.

0.3 Quantitative model for growth

0.3.1 Quantitative model for in depth understanding of cell size control mechanism

Based on our experimental observations, we put forward an analytical model for growth and cell cycle of *E. coli* cells. Our model essentially describes the strategy that *E. coli* cells apply to control their growth and division in each cell cycle. The model is at a single cell level, yet has the power to predict the distributions and correlations between cell size, elongation rate and generation time. In the following, we present the model and then test the model with the data we have collected from *E. coli* cells.

Assumptions of the model. The model is constructed upon a few assumptions that we have itemized below. This assumptions are based on our single cell observation.

- Growth of *E. coli* in steady state is one dimensional. That is, the cell maintains a constant width during the growth and upon division. The mass of the cell is thus proportional to the length of the cell. Here and in what follows, length and elongation rate is used to characterize "size" and "size increase rate" of the cells.

- Cells are evidenced to grow exponentially during a cell cycle. Quantitatively, the relationship is $L = \ell \times 2^{\lambda T}$ where L the is cell length at division,

ℓ the newborn length, λ the elongation rate, and T is the generation time.

- Elongation rate is constant while generation time is a stochastic variable correlated with cell size. As demonstrated in the previous section, the coefficient of variations of elongation rate is two folds smaller than that of generation time. That is, it's effect on cell size variation is about 4 times smaller than that of generation time. To establish a robust model with the smallest number of fitting parameters, we assume the elongation rate, λ, is constant during a cell cycle and is the same for all the cells in the population. On the other hand, the generation time, T, is a stochastic variable, correlated with newborn cell size. This correlation and the noise and fluctuations of T will be discussed below.

Functional form of generation time. The generation time can be divided into two compartments: the deterministic part and the noise part, $T = T_n + T_d$. The deterministic part is a function of the newborn cell size, yielding the correlation between generation time and cell size, $T_d = T_d(\ell)$. The noise part, T_n, consistent with the observations, is assumed to be a Gaussian variable with a constant coefficient of variations, independent of the cell size. As the first step for developing the model, we work out the functional form of the deterministic part of the generation time, T_d. To keep the generality of the model, we construct the function such that it interpolates between two plausible cell cycle models, constant timer model and constant division size model (see below for details). We, however, argue that non of these limiting models are correct and we conclude that cells behavior is somewhat in between, which is explained quantitatively as a model in which cell mass increase is the same in each cell cycle, regardless of the newborn size. This model predicts that any deviation from average cell size will be corrected in a few generations and cell size fluctuations remain constant over different replicative age. An interpolating parameter, which is eventually determined by fitting, indicates in which model the cells are behaving.

Asymptotic models. The two asymptotic models for cell cycle that we interpolate between are constant generation time and constant cell size at division. In the former, cell adjust the generation time to the average expected value depending on the growth condition regardless of the newborn cell size. In the latter, a cell grows until it reaches a especial size, let's say $2 \times \overline{\ell}$, before division. Generation time for these limiting cases can be formulated, respectively, as:

$$T_d = \frac{1}{\lambda} \tag{1}$$

$$T_d = \frac{1}{\lambda} + \frac{1}{\lambda} \log\left(\frac{\overline{\ell}}{\ell}\right) \tag{2}$$

Interpolation function. These two limiting models may sound appealing, but we can readily argue that non of these may hold for *E. coli* growth; in the constant timer model, the fluctuations in size only arises from noise in the generation time. In such a model, noise in the cell size accumulates over

generations – no mechanism to control the fluctuations. In the constant division size model, cells divide upon reaching an especial cell size appropriate for cell division. in this model, there will be no correlation between cell size at division and newborn cell size. We use an interpolating parameter α to bridge between these two cases. The generation time is then defined as

$$T_d = \frac{1}{\lambda} + \frac{\alpha}{\lambda} \log\left(\frac{\overline{\ell}}{\ell}\right) \tag{3}$$

The interpolation parameter α can be eventually obtained by fitting Eq. 3 to experimental results. The α obtained from fitting indicates the cells behavior in growth and cell size control.

Special case at $\alpha = 0.5$: constant mass increase in each cycle. Before proceeding to fitting the generation time to experimental data. We would like to point out that the mid point of the interpolation, at $\alpha = 0.5$, is an special case that the length increase in each cell cycle is constant $\overline{\ell}$. This can be shown as below:

$$\begin{aligned} &L = \ell \times 2^{\lambda T_d} \\ &\overline{\ell} = \Delta L = \ell \times 2^{\lambda T_d} - \ell \\ &\overline{\ell}/\ell + 1 = 2^{\lambda T_d} \\ &\log\left(1 + \overline{\ell}/\ell\right) = \lambda T_d. \end{aligned} \tag{4}$$

since $\overline{\ell}/\ell \approx 1$, at the first order approximation, $\log\left(1 + \overline{\ell}/\ell\right) \approx 1 + 0.5 \log\left(\overline{\ell}/\ell\right)$. Substituting it back in Eq. 4, we obtain:

$$T_d = \frac{1}{\lambda} + \frac{0.5}{\lambda} \log\left(\frac{\overline{\ell}}{\ell}\right) \tag{5}$$

comparing Eq. 5 with Eq. 3, we see $\alpha = 0.5$ approximately corresponds to the case that cell length increase is constant, $\overline{\ell}$, during each cell cycle, regardless of the newborn size.

Derivation of the distributions. Here we derive an analytical expression for the distributions of newborn cell size and generation time. Starting from the generalized form of T_d in Eq. 3 we can proceed calculating newborn size of a new generation based on the previous generation parameters, assuming precise cells division.

The generation time is given as,

$$\begin{aligned} &T = T_d + T_n \\ &T = \frac{1}{\lambda} + \frac{\alpha}{\lambda} \log\left(\frac{\overline{\ell}}{\ell}\right) + T_n. \end{aligned} \tag{6}$$

Length of a new generation can be written in terms of previous generation parameters:

$$L = \ell \times 2^{\lambda T}$$

$$\ell_{new} = \frac{1}{2}\ell_{prev} \times 2^{\lambda T}. \tag{7}$$

Substituting Eq. 6 in Eq. 7, and reordering the results we obtain:

$$\begin{aligned}
\log\left(\frac{\ell_{new}}{\overline{\ell}}\right) &= \log\left(\frac{\ell_{prev}}{\overline{\ell}}\right) + \lambda T_d + \lambda T_n - 1 \\
&= \log\left(\frac{\ell_{prev}}{\overline{\ell}}\right) + \alpha \log\left(\frac{\overline{\ell}}{\ell_{prev}}\right) + \lambda T_n \\
&= (1-\alpha)\log\left(\frac{\ell_{prev}}{\overline{\ell}}\right) + \lambda T_n
\end{aligned} \tag{8}$$

From stationarity, $P\left[\log\left(\ell_{new}/\overline{\ell}\right)\right] = P\left[\log\left(\ell_{prev}/\overline{\ell}\right)\right] = P\left[\log\left(\ell/\overline{\ell}\right)\right]$. Since the T_n is a Gaussian variable, $P\left[\log\left(\ell/\overline{\ell}\right)\right]$ is Gaussian too. Variance of the distribution can be related as follows:

$$\begin{aligned}
\mathrm{Var}\left[\log\left(\ell/\overline{\ell}\right)\right] &= (1-\alpha)^2 \mathrm{Var}\left[\log\left(\ell/\overline{\ell}\right)\right] + \lambda^2 \mathrm{Var}\left[T_n\right], \\
\alpha(2-\alpha)\mathrm{Var}\left[\log\left(\ell/\overline{\ell}\right)\right] &= \lambda^2 \mathrm{Var}\left[T_n\right]
\end{aligned}$$

$$\mathrm{Var}\left[\log\left(\ell/\overline{\ell}\right)\right] = \frac{\lambda^2 \mathrm{Var}\left[T_n\right]}{\alpha(2-\alpha)} \equiv \sigma_\ell^2. \tag{9}$$

Assuming

$$P\left[T_n\right] = \frac{1}{\sigma_n\sqrt{2\pi}} \exp\left(\frac{-T_n^2}{2\sigma_n^2}\right), \tag{10}$$

with $\sigma_n^2 = \mathrm{Var}\left[T_n\right]$, distribution of ℓ, the newborn cell size, can be derived as follows:

$$\begin{aligned}
P\left[\log\left(\ell/\overline{\ell}\right)\right] &= \frac{1}{\sigma_\ell\sqrt{2\pi}} \exp\left(\frac{-\log\left(\ell/\overline{\ell}\right)^2}{2\sigma_\ell^2}\right) \\
P\left(\ell\right) &= \frac{1}{\ell\sigma_\ell\sqrt{2\pi}} \exp\left(\frac{-\log\left(\ell/\overline{\ell}\right)^2}{2\sigma_\ell^2}\right),
\end{aligned} \tag{11}$$

with $\sigma_\ell^2 = \lambda^2\sigma_T^2/\alpha(2-\alpha)$. (see Eq. 9).

The distribution of the generation time can also be calculated in a similar way. Recalling Eq. 6, considering the Gaussian distribution of T_n and $\log(\ell/\overline{\ell})$, the generation time, T, should be a Gaussian variable. The variance is determined from Eq. 6:

$$\begin{aligned}
\mathrm{Var}\left[T\right] &= \left(\frac{\alpha}{\lambda}\right)^2 \mathrm{Var}\left[\log\left(\frac{\overline{\ell}}{\ell}\right)\right] + \mathrm{Var}\left[T_n\right], \\
\mathrm{Var}\left[T\right] &= \frac{2\alpha}{\alpha(2-\alpha)}\mathrm{Var}\left[T_n\right] \equiv \sigma_T^2.
\end{aligned} \tag{12}$$

The distribution of generation time, T, is thus given,

$$P[T] = \frac{1}{\sigma_T \sqrt{2\pi}} \exp\left[-\frac{(T - 1/\lambda)^2}{2\sigma_t^2}\right]. \tag{13}$$

0.3.2 Testing the model

In this section we test, the model that we have developed against our experimental data. Here are the steps we take to test the mode: *i*) finding α by fitting the Eq. 3 to the average generation time versus newborn cell size (Fig. 4, right panel). *ii*) Finding the variance of T_n from results plotted in figure 5. *iii*) comparing the distribution of generation time and cell size, prediction of the model in Eqs. 13 and 11.

Finding the value of interpolating parameter α. Figure 6 shows the deterministic part of the generation time as a function of newborn cells size for a few different cases including the best fit for α. The green dashed line in this figure corresponds to constant timer where $T_d = 1/\lambda$, regardless of the cell size. The green dashed line corresponds to fixed division size where $T_d = 1/\lambda + (1/\lambda)\log\left(\bar{\ell}/\ell\right)$, with $\bar{\ell}$ the half of the division size. The blue line in the figure with error bars are experimental results and the solid red line is the best fir of Eq. 3 to the experimental results.

The value of α we have obtained from this fitting is $\alpha = 0.48$ which is very close to the special case $\alpha = 0.5$ in which the cell size increase in each cell cycle is constant.

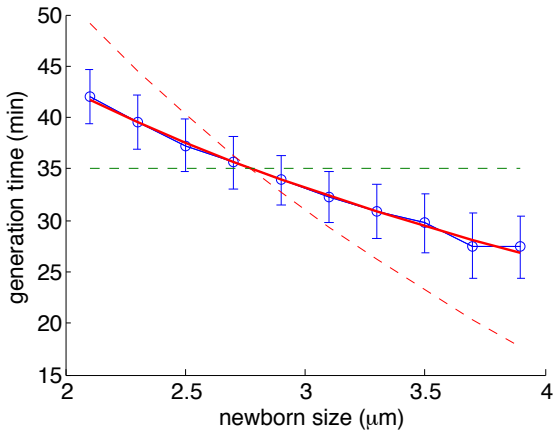


Figure 6: The deterministic part of generation time as a function of newborn size. This graph to to show how to finding the value for the interpolating parameter α. The green dashed line depicts constant timer model, $\alpha = 0$. The red dashed line depicts constant division size, α. The blue line show the experimental results. The solid red line shows the best fit corresponding to $\alpha = 0.48$.

Finding the standard variation of T_n. To find the standard deviation of T_n, σ_n, we used data similar to that plotted in figure 5 for different bins and took the average. The value we found is $\sigma_n = 5.38\,\text{min}$.

comparing the model versus experiment. Now that we have the two parameters that we had to obtain by fitting, $\alpha = 0.48$ and $\sigma_n = 5.38\,\text{min}$, we can compare the distribution of newborn cell length and generation time as predicted in Eqs. 11 and 13, respectively, with experimental results. Figure 7 depicts both theoretical predictions and experimental results for these distributions. The agreement between theory and experiment is remarkable.

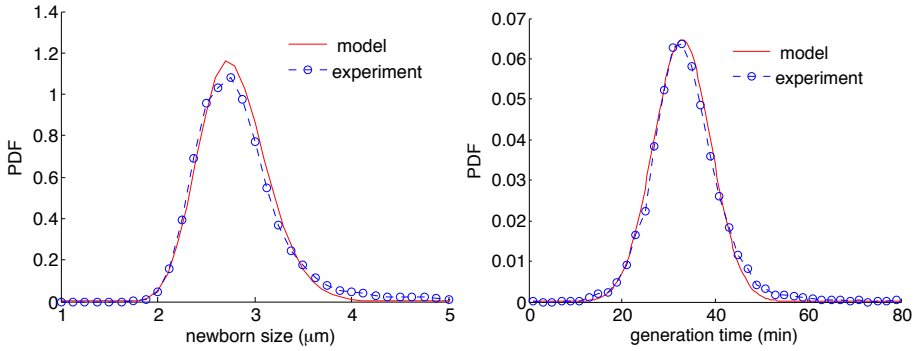


Figure 7: Comparison of model predictions (solid red lines) with experimental data for the distributions of newborn cell size (left panel) and generation time (right panel).

0.3.3 Autocorrelation functions

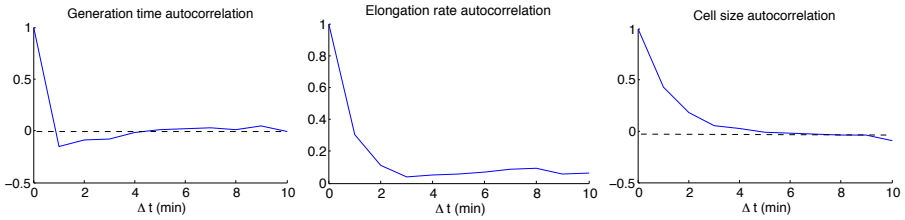


Figure 8: Autocorrelation function for generation time, elongation rate and cel size.

0.3.4 Analytical solution for generation time autocorrelation

Here we find the autocorrelation function of generation time for $\alpha = 0.5$, as it corresponds to constant cell mass (or, alternatively, length) increase per cell

cycle. The normalized autocorrelation function is defined as

$$A(t) = \frac{\left\langle (T_i - \overline{T}).(T_{i+t} - \overline{T}) \right\rangle}{\left\langle (T_i - \overline{T})^2 \right\rangle} \tag{14}$$

where T_i is the generation time of a cell with replicative age of i and $\overline{T} = 1/\lambda$ is the average generation time of all cells. T_{i+t} denotes the generation time of the cell born at replicative age of $i + t$ in the same lineage as the cell with T_i generation time. The average $\langle\,\rangle$ runs over all cells with all replicative ages. In this section we calculate the $A(1)$ and $A(2)$.

To begin with, let's choose a random cell, whose replicative age and newborn length are respectively i and ℓ_i. ℓ_i is a stochastic variable with log-normal distribution as calculated in Eq. 11. The generation time for this cell, T_i, is given based on Eq. 6:

$$T_i = \frac{1}{\lambda} + \frac{1}{2\lambda} \log\left(\frac{\overline{\ell}}{\ell_i}\right) + T_n^i. \tag{15}$$

where T_n^i refers to the noise in generation time for this cell. Note that T_n^i is a Gaussian variable independent of i and ℓ_i. To calculate $A(1)$, we need to go one step forward and find the generation time for the next generation cell, whose replicative age is $i + 1$. First, considering constant mass increase in each cell cycle, the newborn length for the next generation cell is given by

$$\ell_{i+1} = \frac{1}{2}\left(\overline{\ell} + \ell_i\right) \times 2^{\lambda T_n^i} \tag{16}$$

Then, the generation time corresponding to this newborn length is

$$\begin{aligned} T_{i+1} &= \frac{1}{\lambda} + \frac{1}{2\lambda} \log\left(\frac{2\overline{\ell}}{\ell_i + \overline{\ell}} \times 2^{-\lambda T_n^i}\right) + T_n^{i+1} \\ &= \frac{1}{\lambda} - \frac{1}{2\lambda} \log\left(\frac{1}{2} + \frac{\ell_i}{2\overline{\ell}}\right) - \frac{1}{2} T_n^i + T_n^{i+1}. \end{aligned} \tag{17}$$

Since $\ell_i/\overline{\ell} \approx 1$, at the first order approximation we can replace $\log\left(1/2 + \ell_i/2\overline{\ell}\right)$ with $\frac{1}{2}\log\left(\ell_i/\overline{\ell}\right)$. T_{i+1} then reads

$$T_{i+1} = \frac{1}{\lambda} - \frac{1}{4\lambda} \log\left(\frac{\ell_i}{\overline{\ell}}\right) - \frac{1}{2} T_n^i + T_n^{i+1}. \tag{18}$$

Substituting Eqs. 15 and 18 in autocorrelation function we obtain

$$A(1) = \frac{\left\langle \left(\frac{1}{2\lambda} \log\left(\overline{\ell}/\ell_i\right) + T_n^i\right) . \left(-\frac{1}{4\lambda} \log\left(\ell_i/\overline{\ell}\right) - \frac{1}{2} T_n^i + T_n^{i+1}\right) \right\rangle}{\left\langle \left(\frac{1}{2\lambda} \log\left(\overline{\ell}/\ell_i\right) + T_n^i\right)^2 \right\rangle}. \tag{19}$$

Here the averages are on over all possible ℓ_i, T_n^i, T_n^{i+1}. Mathematically, $\langle X \rangle = \int X P\left[\ell_i\right] P\left[T_n^i\right] P\left[T_n^{i+1}\right] d\ell_i dT_n^i dT_n^{i+1}$. Since $\log(\ell_i/\overline{\ell})$, T_n^i, T_n^{i+1} are independent stochastic variables with Gaussian distribution around zero, many terms

in the above equation are zero once the averaging is carried out. With some simplification we oobtain

$$A(1) = \frac{\left\langle \frac{1}{8\lambda^2} \log^2\left(\ell_i/\overline{\ell}\right) - \frac{1}{2}\left(T_n^i\right)^2 \right\rangle}{\left\langle \frac{1}{4\lambda^2} \log^2\left(\overline{\ell}/\ell_i\right) + \left(T_n^i\right)^2 \right\rangle} = \frac{\frac{1}{8\lambda^2}\left\langle \log^2\left(\ell_i/\overline{\ell}\right)\right\rangle - \frac{1}{2}\left\langle \left(T_n^i\right)^2 \right\rangle}{\frac{1}{4\lambda^2}\left\langle \log^2\left(\overline{\ell}/\ell_i\right)\right\rangle + \left\langle \left(T_n^i\right)^2 \right\rangle} \tag{20}$$

The second moment of the Gaussian variables are given as

$$\left\langle \left(T_n^i\right)^2 \right\rangle = \sigma_n^2 \tag{21}$$

$$\left\langle \log^2\left(\ell_i/\overline{\ell}\right)\right\rangle = \sigma_\ell^2 = \frac{\lambda^2\sigma_n^2}{\frac{1}{2}(2-\frac{1}{2})} \tag{22}$$

Substituting these moments in previous equation we get

$$A(1) = \frac{\frac{1}{4(2-1/2)}\sigma_n^2 - \frac{1}{2}\sigma_n^2}{\frac{1}{2(2-1/2)}\sigma_n^2 + \sigma_n^2} = -\frac{1}{4}. \tag{23}$$

To find $A(2)$ we should go one step forward. A similar approach yields

$$A(2) = -\frac{1}{8} \tag{24}$$

0.4 conclusions