UC San Diego
UC San Diego Electronic Theses and Dissertations

Title

Escherichia coli aging from a single cell

Permalink

https://escholarship.org/uc/item/7jb1k8sc

Author

Shi, Chao

Publication Date

2022

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

Escherichia coli aging from a single cell

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy

in

Biology

by

Chao Shi

Committee in charge:

        Professor Lin Chao, Chair
        Professor Ronald S. Burton
        Professor Terry Hwa
        Professor Justin Raymond Meyer
        Professor Scott A. Rifkin

2022

The Dissertation of Chao Shi is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2022

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

I would like to first thank Professor Lin Chao for his extensive mentorship in scientific research and communication.His great passion for science and continuous desire to understand the biology have been a great motivation for my academic research.

I would like to thank mycommittee members: Professor Terry Hwa, Professor Justin Meyer, Professor Scott A. Rifkin, Professor Ronald Burton, for their helpful suggestions and support.

I would like to thank my colleagues, Ulla Camilla Rang, Audrey Proenca, Christen Buetz, Andrew Qiu, Xiyu Liu, and Travis Chan.

I would like to thank all past and current members in quantitative biology specialization.

Last but not least, I would like to thank my parents for they give me life.

Chapter 1, in full, is a reprint of material published in: Chao Shi, Lin Chao, Audrey Menegaz Proenca, Andrew Qiu, Jasper Chao and Camilla U. Rang. 2020. "Allocation of gene products to daughter cells is determined by the age of the mother in single Escherichia coli cells".*Proc. R. Soc. B.28720200569.20200569*. The dissertation author was the primary investigator and author of this paper.

Chapter 2, in full, is currently being prepared for submission for publication of the material: Chao Shi, Lin Chao, Audrey Menegaz Proenca, Andrew Qiu and Camilla U. Rang "An aging landscape results from damage dynamics in single cell". *In preparation*. The dissertation author was the primary investigator and author of this material.

Chapter 3, in full, is currently being prepared for submission for publication of the material: Chao Shi, Lin Chao, Audrey Menegaz Proenca, and Camilla U. Rang. "Optimal bacterial size results from multi-linear elongation." *In preparation*. The dissertation author was the primary investigator and author of this material.

VITA

EDUCATION

2015   Bachelor of Science, Sun Yat-sen University
            Biological Science

2022   Doctor of Philosophy, University of California San Diego
            Biology

PUBLICATIONS

Chao Shi, Lin Chao, Audrey Menegaz Proenca, Andrew Qiu, Jasper Chao and Camilla U. Rang. 2020. "Allocation of gene products to daughter cells is determined by the age of the mother in single Escherichia coli cells".*Proc. R. Soc. B.28720200569.20200569.*

Audrey Menegaz Proenca, Camilla Ulla Rang, Andrew Qiu, **Chao Shi**, Lin Chao. 2019. "Cell aging preserves cellular immortality in the presence of lethal levels of damage." PLoSBiol 17(5): e3000266. https://doi.org/10.1371/journal.pbio.3000266.

Audrey Menegaz Proenca, Camilla Ulla Rang, Christen Buetz, **Chao Shi** & Lin Chao. 2018"Age structure landscapes emerge from the equilibrium between aging and rejuvenation in bacterial populations." *Nat Commun* 9, 3722 (2018). https://doi.org/10.1038/s41467-018-06154-9.

Chao Shi, Jian Yin, Zhe Liu, Jian-Xin Wu, Qi Zhao, Jian Ren and Nan Yao. 2015. "A systematic simulation of the effect of salicylic acid on sphingolipid metabolism." Front Plant Sci. 2015; 6: 186.doi: 10.3389/fpls.2015.00186

Tie-Jun Sun, Yun Lu, Mari Narusaka, **Chao Shi**, Yu-Bing Yang, Jian-Xin Wu, Hong-Yun Zeng, Yoshihiro Narusaka and Nan Yao. 2015. "A Novel Pyrimidin-Like Plant Activator Stimulates Plant Disease Resistance and Promotes Growth" PLoS ONE 10(4): e0123227. https://doi.org/10.1371/journal.pone.0123227

ABSTRACT OF THE DISSERTATION

Escherichia coli aging from a single cell

by

Chao Shi

Doctor of Philosophy in Biology

University of California San Diego, 2022

Professor Lin Chao, Chair

E.coli are prokaryotes that show aging and rejuvenation. Evidences show that damage allocation among daughter cells can explain the aging and rejuvenation pattern in an E.coli lineage. However, population aging originated from cellular aging. And in E.coli, there is not enough study on the molecular dynamics that are necessary for a single cell to create different daughter cells. This is the focus of this thesis. Namely,

single cell experiments and single cell biophysic models are built to explain cellular aging in E.coli. And these studies are aimed to provide informative connections between cellular aging and population aging.

The age difference between cell poles is established physically by the division pattern of E.coli. With one pole being synthesed anew from most recent division, two poles of E.coli is inevitably different by the time they each exist. And proteins that harbor at the poles will have different age and wear and tear. On the other hand, given the division pattern, natural selection is optimizing the physical parameters and dynamics of proteins inside cell, so their intracellular distribution will result in the optimal intercellular distribution among the daughters and become beneficial to population fitness. Two categories of proteins are of greatest interest. First is damaged protein, usually in the form of protein aggregate, with its consequence being studied but dynamics obsure. Second is fresh functional protein and repair protein. Up to the starting of this thesis, no study has contributed to this topic.

The chapter 1 will investigate the fresh protein distribution in line with cellular aging. Chapter 2 is a model tackling the damage dynamics inside single cell. Chapter 3 is discussing more conceptually and fundamentally, the evolutionary origin of cell division.

In chapter 1, a constitutively expressing green fluorescent protein (GFP) gene is inserted into E.coli genome under a stable promoter, and its intracellular and intercellular distribution is observed and discussed. We also invented a protocol of deconvoluting fluorescent image of E.coli colony. GFP here is representing a group of fresh proteins that are free of damage. The observation of GFP enriches in new pole

and new daughters are in consistent with the higher growth rate of new pole and new lineage. The representation of GFP is also appreciated by its relative small molecular weight and fast diffusion. This makes its localization most likely be a passive visualization of free intracellular space instead of an active and protein-specific arrangement. The fact of lack of physiological function of GFP in E.coli also supports the enrichment being passive, which indicates a baseline enrichment level of functional but invisible proteins that contribute to real fitness. In more detailed study, we find the free cellular space revealed by GFP can be predicted qualitatively by the age of the pole. The fact that younger pole has more free space leads to the speculation of accumulation of aged or damaged protein on the other end of cell. And therefore, fresh proteins are spatially excluded by damaged protein, from which polarity of cell is established and physiological distinction between daughters are maximized. Cellular polarity will subsequently be magnified into standing fitness variation of population, and we found the age can explain 37% of total variance of GFP distribution. This polarity is created by self-organisation based on size of molecules, and cannot be realized without damage aggregation. Given the possible energy cost of sorting and rearranging each specific damage or fresh protein of all sizes, the aggregation of damage allows damage and fresh protein to be organised physically by their size with no energy involved. The damage aggregation- disaggregation dynamics will be explored in Chapter 2.

In chapter 2, classical view of damage distribution causinig cellular aging is examined by a biophysical and individual based finite population model. Damage is hypothesed to have been actively transported to old pole, or have been excluded from nucleoid because of aggregation. The first hypothesis is rejected because the resulting

exponential-shape distribution is conflicting the observed damage enrichment of both poles. However, nucleoid exclusion hypothesis have not been tested and parameterized in an elongating E.coli cell. The key parameters in this picture, and are certainly parameters under natural selection, are the size of damage and the elongation regime of E.coli.

Damage particles originated from wear and tear of fresh and functional protein. When protein is damaged, the size will most likely keep similar. In this way, the high diffusion rate of protein will cause damage to have uniform distribution inside cell, and no polarity can be built without extra mechanism or energy. The evolution of protein aggregation solves this problem by allowing damaged protein to bind and form larger aggregate. The larger size will be physically excluded by nucleoid and eventually reside in the cell poles. And when cell divides, new pole is created damage free by the division of nucleoid, and will take time for new aggregate to colonize. And damage polarity is established automatically. Therefore, it is important that the size of aggregate is larger than the mesh size of nucleoid and larger than the space between nucleoid and cell wall. Once resides into the pole, the aggregate will keep growing by merging with more and more free flowing new damages, and will eventually cause old lineage die from damage overload, if not saved by disaggregation or elongation. Disaggregation will efficiently control the size of damage aggregate by releasing repaired protein out from aggregate, and once aggregation –disaggregation equilibrium is reached, the size of aggregate will keep constant. In this way, disaggregation saves the old lineage by compromising the polarity of cell. And a moderate level of disaggregation is necessary. In our model, the free damage size after parameterization is close to the size of a ribosome. Therefore

free damage is small enough to travel through nucleoid between poles. And it takes aggregation of 3 free damages to be excluded from nucleoid and establish damage polarity. The disaggregation probability from model is 0.1. Since damage polarity is established passively, it is not robust under changing level of external damage, which occurs randomly across a cell. It explains why both experiment and model produce symmetry when external damage increases.

Elongation, in general, is an automatic dilution of existing damage, and will further promote elongation by alleviating detrimental effect of damage. And fast elongation will in turn lead to less reception of new external damage. There is a prominent positive feedback between fast/slow elongation and less/more amount of damage. Subcellular elongation pattern has been measured by transient cell wall labeling and monitoring local dilution. Our observation of new pole elongates significantly fast than old pole means elongation happened locally and the elongation –dilution feedback will also apply subcellularly. This results in more polarized damage distribution of a cell once initial damage/fresh protein polarity is established. In our model, each cell is represented by multiple one dimension compartments. And each compartment has its unique elongation probability, due to local damage, to reproduce local elongation. After parameterization, the model can reproduce the elongation rate relation between mother and old and new daughters, including stochastic attractor behavior at the equilibrium point where mother's doubling time equals one of her daughter's.

This model is developed on the assumption of exponential cell growth from self-replicating compartments (growth units). In chapter 3, an in-detailed analysis is performed to quantitatively describe their behavior and consequences.

When measured in better time resolution, E.coli elongation performs accerlerating multi-linear phases of growth. Each linear phase can be realized as constant biomass production of a certain growth units, with the shift to a faster phase being the event of a new unit start functioning. To keep up with cell size, the growth unit number needs to double as the cell length doubles. Considerinig most cells perform two linear phases during doubling, the growth unit should start from number of 2, advance to 3, and divide right at the time it reaches 4. The fact of observing the third unit functioning also indicates the mixing of products from 2 existing units (cooperation). This mechanism gives cell that contains more growth units fitness advantage for their elongation span accomadates more rounds of cooperation. And a cell with short length has lower fitness. However, our elongation rate data at long cell length indicates slowing down of accerleration across phase shift, which can be explained by inefficient cooperation at new cell pole (c.f. chapter 1) caused by long distance between poles. Therefore, growth units not only result in multi-linear cell elongation, but their cooperation produces an optimal cell size and explains cell division.

Chapter 3 now has complete conceptual projection, and there are still controls and tests need to be performed.

Chapter1: Allocation of gene products to daughter cells is determined by the age of the mother in single Escherichia coli cells.

1.1 Abstract

Gene expression and growth rate are highly stochastic in E. coli. Some of the growth rate variations result from the deterministic and asymmetric partitioning of damage by the mother to its daughters. One daughter, denoted the old daughter, receives more damage, grows more slowly, and ages. To determine if expressed gene products are also allocated asymmetrically, we compared the levels of expressed green fluorescence protein in growing daughters descending from the same mother. Our results show that old daughters were less fluorescent than new daughters. Moreover, old mothers, which were born as old daughters, produced daughters that were more asymmetric when compared to new mothers. Thus, variation in gene products in a clonal E. coli population also has a deterministic component. Because fluorescence levels and growth rates were positively correlated, the aging of old daughters appears to result from both the presence of both more damage and fewer expressed gene products.

1.2 Background

Gene expression and protein levels in individual cells is highly variable in clonal populations (1–10). Because many gene-regulating elements have low copy numbers (11), the variation is attributed to stochastic sampling. For example, if the elements are Poisson distributed, they will have a mean of μ and a variance of $[\![\sigma]\!]^{\wedge}2$, where $[\![\mu=\sigma]\!]^{\wedge}2$, but the coefficient of variation for the relative difference between cells is σ/μ = $1/\sqrt{\mu}$, which increases with decreasing values of μ. Given that the amount of expressed gene products is an important component of cellular function and fitness, the amount of stochasticity is at first glance puzzling. A possible explanation is that cellular metabolism constrains the total pool of gene products and some genes are limited to a smaller, and perhaps suboptimal, number of regulatory elements. An additional explanation is that the variation is a form of bet hedging (12,13). If the environment is changing or variable, a variant cell could have by chance the gene product level that is appropriate for that instance. Alternatively, the apparent stochasticity could result from yet uncovered deterministic causes (14–16).

Recent results have shown that the growth rate of single and clonal bacterial cells is also highly stochastic (17–19). However, the growth rates were found to have a significant deterministic component that is controlled by the asymmetrical partitioning of non-genetic damage, such as oxidized or mistranslated proteins, by a mother bacterium to its two daughters. The allocation of more damage to one daughter by a mother bacterium is associated with the age of the maternal cell poles. Because a rod shaped bacterium such as E. coli divides at the midplane of its long axis, the poles formed at the midplane are new while the distal poles are older (figure 1). As a result, all E. coli

cells have an old and a new pole. When a mother cell divides, one of its daughters receives the maternal old pole and the other the new pole. The daughters are denoted, respectively, the old and new daughters. The evidence for the asymmetrical partitioning of damage is manifold. Old daughters have been shown to have a slower growth rate than new daughters (17,20–26). The old pole and old daughters are more likely to harbor aggregates of damaged and synthetically misfolded proteins, and aggregate size is negatively correlated with cell growth rate (25,27). If dnaK, the gene responsible to dismantling aggregates for repair, is knocked out in E. coli lineages of new daughters survive while lineages of old daughters perish (18). Although early investigations reported that damage rates in standard laboratory culture were too low to generate an asymmetry between old and new daughters (23,28), follow up studies have shown that a difference is detectable with improved microscopy and larger sample sizes (17,18,26).

The difference between growth rate of old and new daughters increases the variation between single cells. However, because the asymmetry is a deterministic component of bacterial cell division, the growth rate variation observed in a bacterial population, even in a clonal one, cannot be explained entirely by stochasticity. Because cell growth rates and ribosome number are positively correlated (29), it follows that levels of expressed gene products could also be similarly correlated. We therefore investigated whether levels of expressed proteins could also be asymmetrically distributed between old and new daughters. Previous studies of the stochasticity of expressed proteins in single cells did not look for possible differences between old and new daughters and pooled them as independent replicates. We found that new daughters overall contained higher levels of expressed proteins than old daughters. The

difference between a pair of old and new daughters was greatest when the mother was born as an old daughter. Because old daughters have more damage, their lower level of expressed proteins could be explained by a competition model in which damage and proteins compete for space. Moreover, the level of expressed proteins correlated positively with the growth rate of the cells. From an evolutionary perspective, a growth rate difference between old and new daughters is beneficial because the resulting variation increases the efficiency of selection (24). Thus, the variation in growth rates and expressed gene products in a clonal population of E. coli has a deterministic component that is evolutionarily advantageous.

1.3 Results

Protein levels are biased towards new poles and asymmetrical between old and new daughters

To investigate whether variation in expressed protein levels between E. coli cells has a deterministic component, we first analyzed the GFP levels of micro-colonies of 1-4 cells and compared old and new daughter pairs descending from the same mother (hereafter old and new daughters; figure 1) shortly after division. Despite fluorescence levels varying considerably between cells (consistent with earlier studies (2,4,6)), we noted less fluorescence in old poles than the new poles rendering old daughters overall dimmer as we followed single cells dividing into two and four cells (figure 2a). To explore this further, we plotted the fluorescent profile of new and old daughters along the cells and normalized the daughters into the same cell lengths for comparison (figure 2b). Fluorescence along the cells was biased towards the new poles and not uniformly distributed.

To quantify the bias towards new poles, we compared the old and new pole difference between old and new daughters (figure 1). We found that the new pole was significantly brighter than the old pole in both new and old daughters ($p = 3.41 \times 10^{-48}$ and 2.58×10^{-97}, respectively) (figure 3a). However, the fluorescence ratio of new to old poles was 1.2 ± 0.06 (SEM) in old daughters and 1.11 ± 0.006 in new daughters, and the pole ratio of old daughters was significantly larger than that of old daughters ($p = 2.8 \times 10^{-19}$). If the old and new poles were pooled to obtain the total fluorescence for single cells of new and old daughters, new daughters were significantly brighter ($p = 2.57 \times 10^{-65}$) and the daughter fluorescence ratio (new/old) was 1.08 ± 0.004.

5

Deterministic asymmetry of GFP fluorescence is higher in daughters from old mothers than from new mothers

Because pole fluorescence was more similar in new daughters than in old ones, we hypothesized that fluorescence between daughters from new mothers should be more similar than between daughters from old mothers. To identify old and new mothers, which were mothers born respectively as old or new daughters, we tracked the divisions for an additional generation beyond lineages presented in figure 1. As before (see above), the new daughters from both new and old mothers were brighter ($p = 1.78 \times 10^{-5}$ and $p = 1.68 \times 10^{-21}$, respectively). However, old mothers produced significantly ($p = 2.75 \times 10^{-9}$) more different daughters than new mothers, as demonstrated by the respective daughter fluorescence ratios of 1.12 ± 0.009 and 1.04 ± 0.008 (figure 3b). Note that the average of the 1.12 and 1.04 ratios replicates closely the value of 1.08 that was obtained for the pooled daughters in the above section.

Deterministic asymmetry accounts for a large component of the variance of expressed gene products in a population

To determine how much of the variation of expressed gene products in single cells is explained by deterministic asymmetry, we estimated the total variance (V_T), the subcomponents attributable to deterministic asymmetry (V_A) and error or unexplained factors (V_E), and $V_T = V_A + V_E$. In the absence of more information, V_E can be interpreted to represent the stochastic component ((15).

V_T, V_A, and V_E were estimated separately for old and new mothers depicted in figure 3a. To obtain estimates for one mother type, the variance of fluorescence levels in its new and old daughters, V_{New} and V_{Old}, was first determined (figure 4). V_T was then estimated from a pool consisting of all the old and new daughters. If deterministic asymmetry is absent, V_{Old} and V_{New} are the sole components, $V_A = 0$, and $V_T = V_E = (V_{Old} + V_{New}) / 2$. If deterministic asymmetry renders the fluorescence level of new daughters higher, the difference $D = M_{New} - M_{Old} > 0$, where MNew and MOld are the mean fluorescence levels of the new and old daughters. Because $D > 0$ pushes apart the old and new daughter distributions and inflates V_T (figure 4), $V_T = (V_{Old} + V_{New}) / 2 + D^2/4$ (19). The deterministic component of variance V_A is the term $D^2/4$, in which case $V_A = V_T - (V_{Old} + V_{New}) / 2$. The contribution of deterministic asymmetry expressed as a percentage is

$$h^2 = V_A / V_T$$
$$= 1 - (V_{Old} + V_{New}) / 2 \, V_T$$
$$= 1 - V_E / V_T$$

Our results showed that h2 = 10.1 and 40.1% for new and old mothers (figure 4a,b). The higher h2 for older mothers is consistent with our results that old poles have less fluorescence (figure 3a) and that old mothers have daughters that are more different (figure 3b). An old pole in an old mother is older than the old pole in a new mother and therefore has lesser fluorescence than the old pole in the new mother. In other words – the more different the poles, the more different the daughters. Thus, a substantial proportion of the variation of expressed gene products previously attributed to stochasticity in single E. coli cells results from the deterministic process by which

more expressed gene products are allocated to new daughters. The amount $V_E = (V_{Old} + V_{New}) / 2$ that remains unexplained and attributed to stochasticity could be further reduced if other deterministic processes are uncovered.

Deterministic component of variance for elongation rates of single cells

Group Because elongation rates in E. coli have a deterministic asymmetric component of variance (17,18), the cells quantified for fluorescence in Figure 4 were further examined to determine whether they manifested a similar asymmetry and variance pattern for elongation rates and whether GFP and elongation rates were correlated. Elongation rates were measured (see Material and Methods) for all daughters and grouped by old and new mothers. The effect of deterministic asymmetry on elongation rates was clear. If the old and new mothers were pooled, the new daughters had a significantly higher elongation rate ($p = 1.96 \times 10^{-23}$), and the ratio of the rates was 1.07 ± 0.006 in favor of new daughters. Moreover, new daughters had a significantly higher rate than old daughters regardless of whether they came from an old or new mother ($p = 1.12 \times 10^{-16}$ and $p = 1.07 \times 10^{-8}$, respectively; ratios of 1.083 ± 0.009 and 1.059 ±0.009) (figure 3c). The ratios of 1.0833 and 1.0584 were significantly different by a one-tail t-test ($p = 0.03$), and the higher ratio suggested again that difference between daughters was bigger in old mothers. Thus, we partitioned the variance of elongation rates to estimate the deterministic fraction. Using the same approach followed for fluorescence (cf. figure 4) we estimated that the variance component due to deterministic asymmetry was $h^2 = 16.9$ and 32.2% for elongation rates of daughters from new and old mothers (figure 5). Thus, just as GFP fluorescence

in single E. coli cells, elongation rates are highly stochastic but a large fraction is deterministic.

Correlation between GFP fluorescence and elongation rate of single cells

Because the patterns of fluorescence and elongation rate in old and new daughters and mothers were similar, we tested next whether the two traits could be related. A test for correlation revealed that fluorescence and elongation rates were positively correlated, however, only significant when coming from old mothers. Daughters from old mothers showed a strong and significant correlation ($r = 0.26$, $p = 0.0066$) (figure 6a), while daughters from new mothers exhibited a weaker and not significant association ($r = 0.041$, $p = 0.14$) (figure 6b). The correlations from old mothers were also significantly larger than the correlations from new mothers ($p = 0.012$). The correlation between elongation rate and fluorescence is consistent with early reports that show positive correlations between ribosome levels, which produce proteins, and growth rate (29,34). It is also consistent with figure 4,5 that show that asymmetry between daughter cells is always higher when originating from old mothers, which allocate more of her damaged proteins to her old daughter (18) and more newly synthesized proteins to her new daughter.

1.4 Discussion

Our results show that expressed gene products, much like cell growth or elongation rates, have both a stochastic and a deterministic component. The deterministic components are manifested in single cells as a consistently higher level of reported GFP fluorescence in the new poles relative to the old poles of a bacterium. If the data, originally reported as ratios (figure 3a), are re-expressed as percent differences, new poles showed 5.2 and 9.1% more fluorescence in new and old daughters, respectively. The difference between the poles is unlikely explained by differential rates of localized gene expression. The mean length of the E. coli cells in this study was 3.6 μm and the mean doubling time was 40.2 min. The GFP mut3b variant used for our study has a maturation time of t50 = 4 min (30) and a diffusion rate of 9 μm2 sec-1 (31). With such a high diffusion rate, mature and fluorescent GFP molecules should effectively have a uniform distribution in the absence of any interfering factors. The presence of aggregates in the old poles of cells (25) suggests that an interfering factor could be limiting space. With more damage in the old pole, newly expressed gene products would find more space to occupy in the new pole. Because the maternal old pole is allocated to the old daughter (figure 1), space limitation due to the damage in the old pole is anticipated to restrict the abundance of expressed GFP in the old daughter. This prediction is supported by our observation that new daughters always were more fluorescent than old daughters, although difference was greater when daughter pairs came from old mothers (2.0 and 5.7% difference; new vs. old mothers; figure 3b). Our reported differences between new and old poles in new and old daughter (5.2 vs 9.1%) and between new and old daughters from new and old mothers

10

(2.0 vs 5.7%) also demonstrate a consistency that supports a possible role for damage. The difference is smaller for new daughters and for new mothers because they have less damage to create the difference. This is further shown by our estimates of the deterministic variance component explained by asymmetry. Deterministic asymmetry accounts for 10.1% of the total variance of GFP fluorescence in daughter pairs of new mothers, but 40.1% in old mothers (figure 4a).

Because there is a strong correlation between elongation rate and the ribosome levels (29,34–36), it follows that the expressed level of gene products could also correlate with elongation rates. Our results confirmed this correlation between elongation rate and protein levels by using GFP production as a proxy for expressed proteins and measuring elongation rate of the same cells (figure 6a,b). The strength of these correlations is noteworthy for two reasons. First, because the production of GFP can be costly to cell growth (37), the correlation shows that it can, at the levels observed in our study, serve as a proxy for expressed gene products. The correlation may in fact be stronger because it may have been attenuated by the cost. Second, the stronger correlation in daughters produced by old mothers (figure 6a) supports again our suggestion that GFP levels, and now growth rates, could result from the space limitation. Because old mothers have more damage, their old and new daughters have more divergent levels of damage, gene products, and growth rates. New daughters, compared to old daughters, had 2.9 vs 4.0% higher elongation rates when produced, respectively, by new and old mothers (figure 3c). Space limitation is additionally supported by similarity between the effects of deterministic asymmetry on variation in expressed GFP levels and elongation rates. Deterministic asymmetry accounted for

32.2 vs 16.9% of total variance when old and new daughter pairs were produced by old vs. new mothers (figure 5a,b). Thus, our results show that the amount of expressed gene products and elongation rates in single cells are highly variable, but cannot be entirely attributed to stochasticity. When the age of a mother cell is considered, deterministic components account for a large fraction of single cell variability.

A tempting hypothesis at this juncture is that the space occupied by aggregates limits the amount of ribosomes in a cell and therefore reduces the level of expressed proteins and the final elongation rate. We recognize that the observed relationships we use to formulate this hypothesis could be correlational and not causal. The hypothesis is only meant for stimulating discussion and future testing. However, because it has been shown that the addition of an external damage agent to cells decreases elongation rates (18), damage may be the trigger that starts the process. Because aggregates of damaged proteins tend to accumulate more in old poles and old daughters (25), and less in new poles and new daughters, the resulting asymmetry has a marked effect on the bacterial population (17,24). Lineages of old daughters accumulate increasingly more damage and the new daughters decreasingly less damage. While the old daughter lineage grows more slowly and ages, the new daughter lineage grows more rapidly and rejuvenates. However, because the total damage a cell has at birth is diluted by growth and increasing cell size, the accumulation of damage in the old daughter lineage increases only until the point at which the rate of increase is cancelled by the dilution. At this point, the growth rate of the old daughters stabilizes at an equilibrium value. Likewise, an equilibrium is achieved by the new daughter lineage. However, because new daughters are allocated less damage, the equilibrium growth rate for

those cells is higher than for old daughters. Thus, old daughters that are grown under standard laboratory conditions age but only until the equilibria. There is no death and the lineages are immortal. However, if the damage rate is increased by introducing external damage agents, such as phototoxicity, antibiotics, or heat, the old daughter equilibrium can be destabilized and the lineage dies. With the death, aging makes the lineage mortal.

The asymmetry between the growth rate of old and new daughters (figures 3c, 5a,b) is in principle evolutionarily advantageous to the bacteria. The growth rate difference between the old and new daughters creates fitness variation that increases the efficiency of natural selection for eliminating the damage from the population. A more intuitive explanation comes from a banking analogy comparing one account started with $1,000 and an interest rate of 8% yr-1 and two accounts with $500 at 6% and $500 at 10%. Splitting the $1,000 into the two $500 accounts yields more returns after one year because of the 10% returns. A bacterial lineage that allocates damage asymmetrically to its daughters likewise gains higher growth rate or fitness returns from the new daughters. The analogy is not perfect because bacterial lineages split the damage every generation. However, in both cases the outcomes are predicted by Jensen's Inequality (38). If the returns are generated by a greater than linear process, such as exponential population growth or interest rates, they are increased by increasing the variance of the initial states.

Cell growth rate variance created by random stochasticity alone can generate the advantage provided by Jensen's Inequality (19). Lineages that by chance received more damage would also age, attain stable equilibrium states, and become mortal with high

rates of damage. However, although these lineages, along with their aging and rejuvenation, could be readily tracked by time-lapse microscopy, they would show no association with old poles and old daughters. Damage could accumulate equally in either old or new poles or daughters. The fact that damage aggregates and aging are associated with old poles and daughters in E. coli (17,18,20,21,25,27,39–42) has led us to suggest that the partitioning may have been polarized by anchored, and thus not diffusible or movable, damage (19). By virtue of being older, old poles most likely harbored the initial anchored damage. Allocating damage to the new pole would have countered the damage anchored to the old pole and the variance between old and new daughters would have been decreased. On the other hand, the variance is increased by polarizing the allocation to the old pole. Because aggregates in E. coli are sticky (39), an aggregate at the old pole grows as other aggregates adhere to it. Whether aggregate stickiness is inherent property of damaged proteins, or a trait evolved to concentrate damage to the old pole, is debatable and not known. Regardless, the asymmetric distribution of damage, and now also of expressed gene products (figures 2a,b, 3a,b, 4a,b), is a deterministic process that increases the variance of single cells in an E. coli population.

1.5 Availability of data and materials

Data are available on Dryad at doi:10.6075/J0542M0K
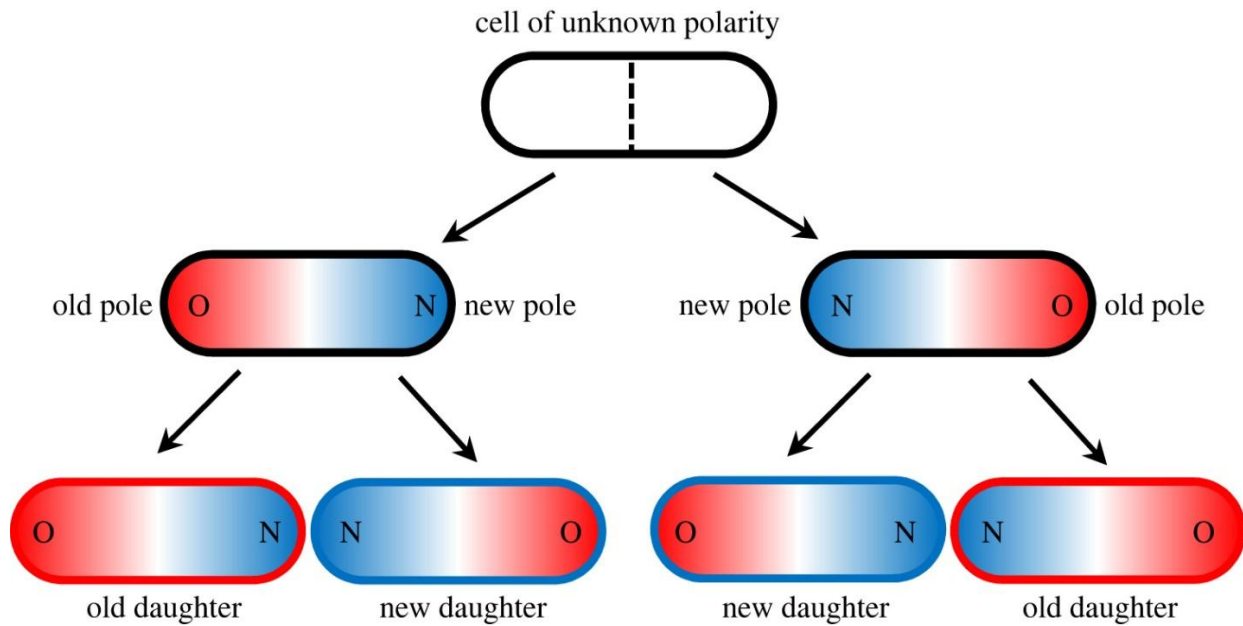
Figure 1.1 | Assignment of old (red) and new (blue) poles and daughters in E. coli. Vertical, dotted line shows the middle and axial plane of the cell. Because the division plane cuts E. coli at the midpoint of the long axis, the poles formed at the division point are new and the distal poles are old. Note that if the polarity of the first cell is unknown, two divisions are required to determine old and new daughter. The outlines of the bottom four daughters in the figure are colored red and blue to identify them as old and new daughters, while the intracellular red and blue colors identify the old and new poles, also designated as O and N.
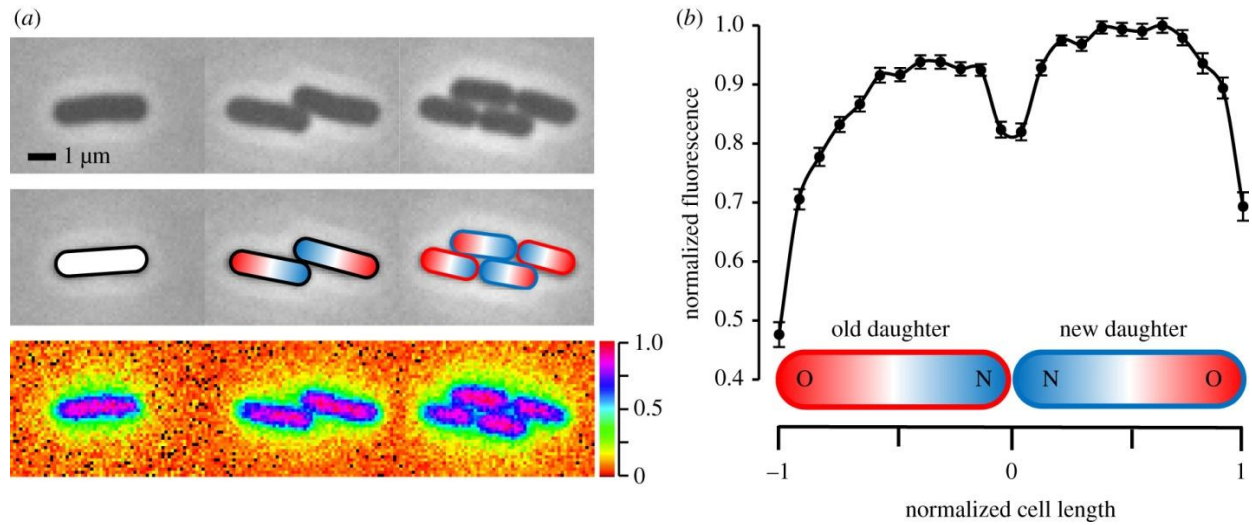
Figure 1.2 | Intracellular fluorescent difference of single cells.(a) Time-lapse images of an E. coli bacterium dividing into two and four cells. Top row: Phase contrast. Middle row: Assignment of old (red) and new (blue) poles from the top row cells. Bottom row: Heat-map of fluorescent images of the top row cells, showing lesser intensity by the old poles (blue color spots) than the new poles and inside the cells (pink color). Scale on the right goes from highest intensity = pink, to lowest = orange. (b) Fluorescence profile along the cells of new and old daughter pairs with color designation as in figure 1 for old (red) and new (blue) poles and daughters. The length of the cells is normalized for comparison (n = 40 pairs). Error bars show Standard Error of the Mean (SEM).
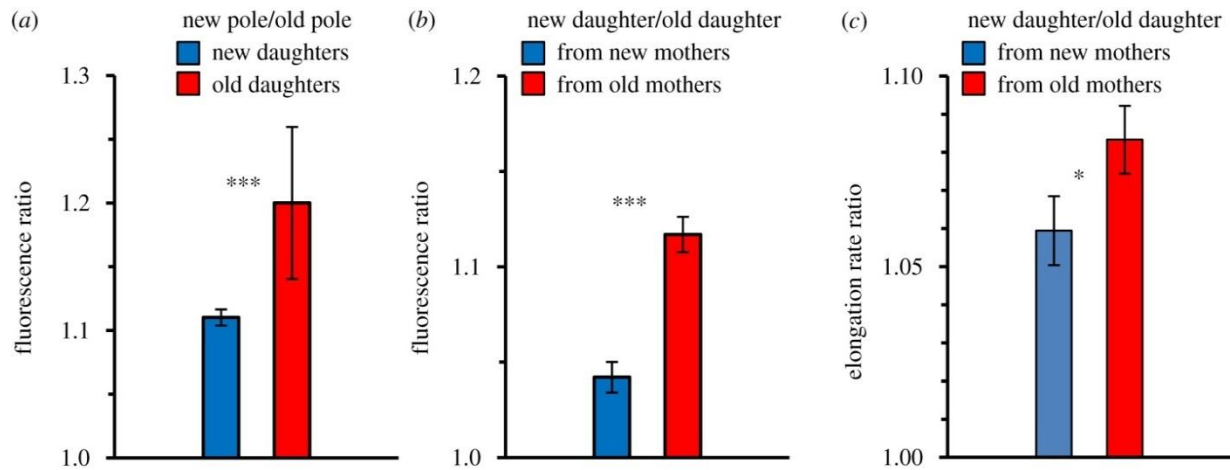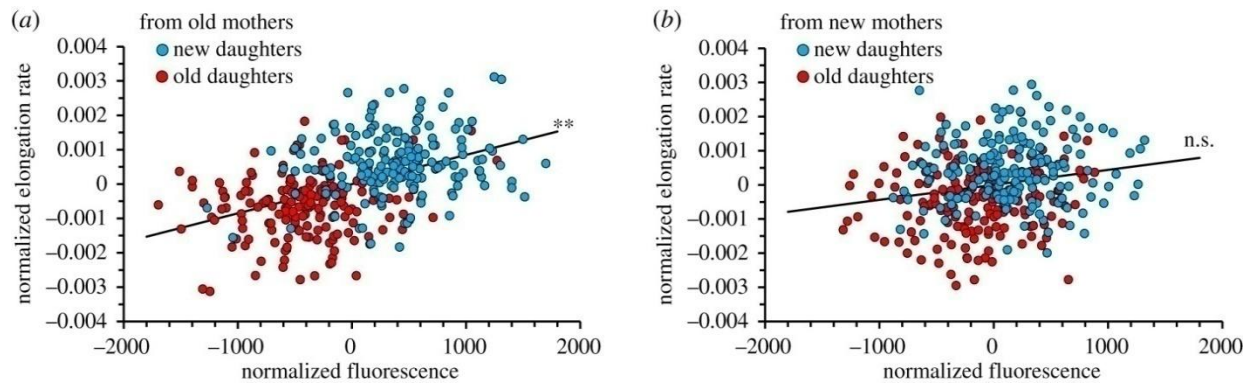
Figure 1.3 | Ratios of fluorescence and elongation rates within and between cells. (a) Fluorescence ratio (new/old) of the two polar halves from old (red) and new (blue) daughters (n = 404 pairs). Old daughters, show a higher asymmetry between the two polar halves than do new daughters. The ratio of fluorescence from the cell half that contains the oldest pole is 1.2 ± 0.06 (p = 2.58 x 10^{-97}, 2-tailed paired t-test) in favor of the half that contains the newest pole in old daughter, and 1.1± 0.006 (p = 3.41 x 10^{-48}, 2-tailed paired t-test) in new daughters. The difference between the two ratios was significant (p = 2.8 x 10^{-19}, 2-tailed non-paired t-test). New daughters in general had a higher fluorescence than old daughters with a ratio of 1.08 ± 0.004 (p = 2.57 x 10^{-65}, 2-tailed paired t-test). (b) Fluorescence ratio (new/old) of daughters from old (red) and new (blue) mothers. The ratio from old mothers was 1.12 ± 0.009 (p = 1.68 x 10^{-21}, 2-tailed paired t-test, n = 178 pairs) and from new mothers 1.04 ± 0.008 (p = 1.78 x 10^{-5}, 2-tailed paired t-test, n = 177 pairs). The difference between the two ratios was significant (p = 2.75 x 10^{-9}, 2-tailed non-paired t-test). Consistent with the same finding mentioned in figure 3a, new daughters in general had a higher fluorescence than old daughters with an average ratio of 1.08 ± 0.006 (p = 4.67 x 10^{-23}, 2-tailed paired t-test). (c) Ratio of elongation rate (see Material and Methods for details) of new over old daughters from new (blue) and old (red) mothers. The ratio from old mothers was 1.0833 ± 0.009 (p = 1.12 x 10^{-16}, 2-tailed paired t-test, n = 178 pairs) and from new mothers 1.0584 ± 0.009 (p = 1.07 x 10^{-8}, 2-tailed paired t-test, n = 177 pairs). The two ratios were significantly different from each other (p = 0.03, 1-tailed non-paired t-test), showing a higher asymmetry between the daughters coming from old mothers than from new mother. New daughters in general had a higher elongation rate than old daughters with an average ratio of 1.07 ± 0.006 (p = 1.96 x 10^{-23}, 2-tailed paired t-test), consistent with the fluorescence ratios mentioned in figure 3a,b.
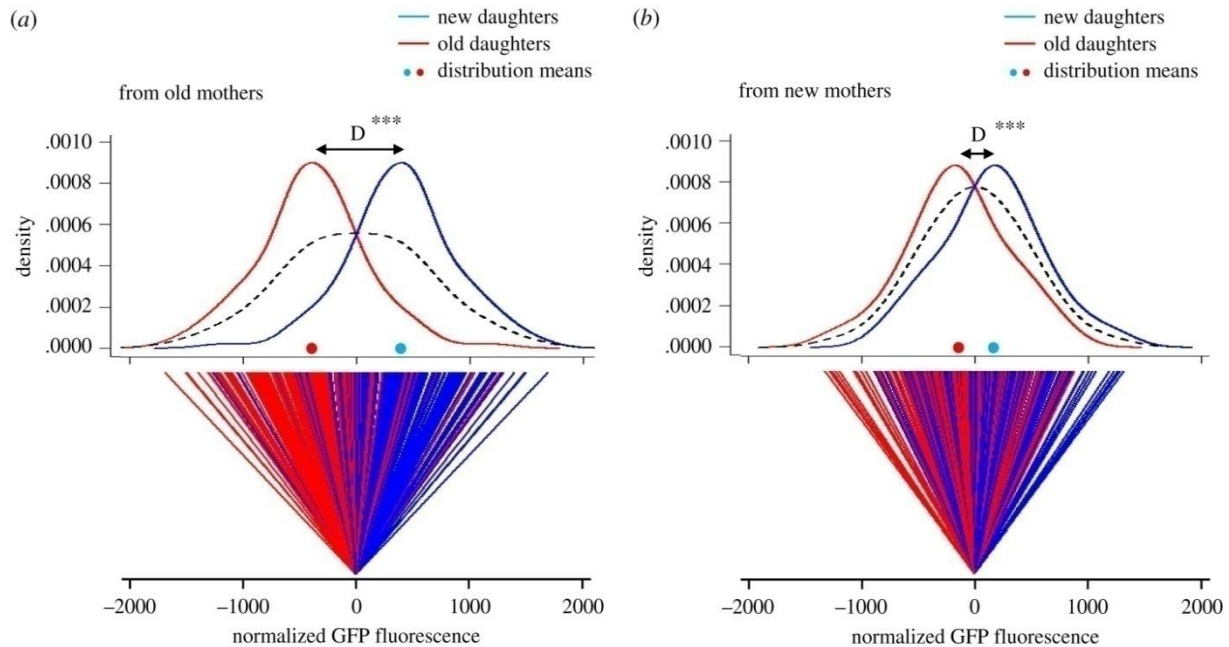
Figure 1.4 | Deterministic variance of fluorescence between new and old daughters from old and new mothers.(a) Top panel: Normalized density – distribution– of fluorescence of new (blue) and old (red) daughters from old mothers (n = 178 daughter pairs). Dots on the x-axis indicate the average fluorescence for each distribution. "D" (black arrow) stands for the distance between peaks of new and old daughter curves. Significance for "D" as determined from ratios between new and old daughters in figure legend 3. The average population density, combining new and old daughters, is indicated by dashed lines. Bottom panel: Normalized fluorescence of each new (blue) and old (red) daughter pairs from old mothers. The zero point is the average fluorescence between each pair. As can be seen, the old daughter in each pair more often ends up on the minus-side of the pair's zero-point, i. e. having less fluorescence. The deterministic asymmetry when the daughters come from old mothers was calculated to constitute 40% of what normally is reported as stochasticity. (b) Same as (a), but from new mothers. n = 177 daughter pairs. The deterministic asymmetry when the daughters originate from new mothers (bottom panel) was calculated to constitute 10% of what normally is reported as stochasticity.
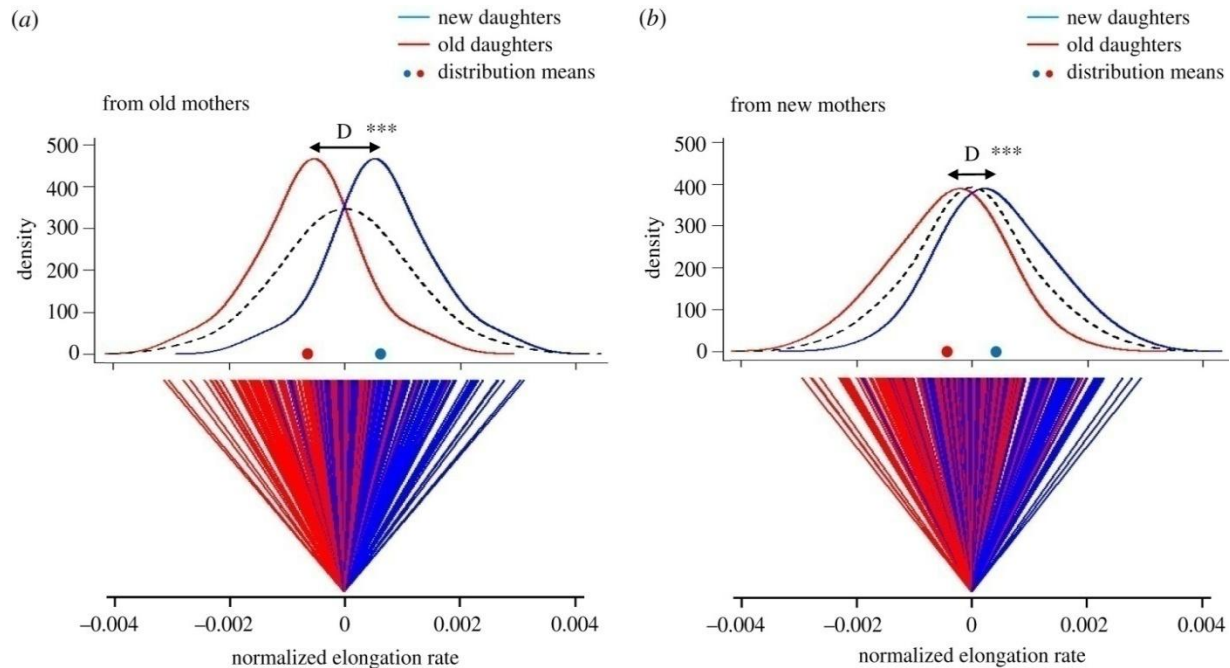
Figure 1.5 | Deterministic variance of elongation rate between new and old daughters from old and new mothers. (a) Top panel: Density –distribution– of elongation rate of new (blue) and old (red) daughters from old mothers (n = 178 daughter pairs). Dots on the x-axis indicate the average fluorescence for each distribution. "D" (black arrow) stands for the distance between peaks of new and old daughter curves. Significance for "D" as determined from ratios between new and old daughters in figure legend 3. The average population elongation rate, combining new and old daughters, is indicated by dashed lines. Bottom panel: Normalized elongation rate of new (blue lines) and old (red lines) daughter pairs from old mothers. The zero point is the average elongation rate between each pair. As can be seen, the old daughter in each pair more often ends up on the minus-side of the pair's zero-point, i. e. having a slower elongation rate. The deterministic asymmetry when the daughters come from old mothers was calculated to constitute 32.2% of the total variance. (b) Same as as (a), but from new mothers. n = 177 daughter pairs. The deterministic asymmetry of daughters coming from new mothers was calculated to be 16.9% of the total variance.

Figure 1.6 | Correlation between normalized GFP fluorescence and normalized elongation rate showing a higher rate when daughters come from old mothers than from new mothers. Because the normalized values did not conform to a standard Gaussian distribution, the correlation statistics and comparisons were conducted by randomizing the data and obtaining a null distribution of correlations. The p-values reported represent the probability that the observed correlation is exceeded by the correlations of the null distribution.(a) Normalized fluorescence versus normalized elongation rate between new and old daughters from old mothers. Correlation r = 0.24, p = 0.0066 (**), n = 178 daughter pairs. (b). Normalized fluorescence versus normalized elongation rate between new and old daughters from new mothers. Correlation r = 0.041, p = 0.14 (n.s.), n = 177 daughter pairs. A comparison of observed correlation from old and new mothers (figure 6a vs 6b) was found to be significant (p = 0.012).
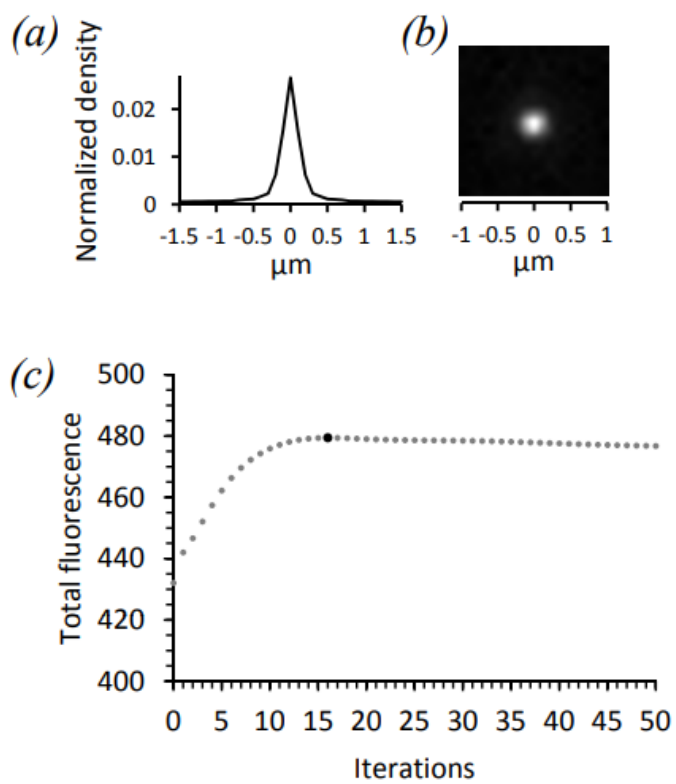
Supplementary Figure 1.1 | Point spread function and criteria of iterative deconvolution

1.7 Methods

Bacterial strains, growth media, and GFP reporter

Growth experiments were performed using E. coli K12 (NCM3722 ΔmotA:frt, chromosomal:T:ptet-GFP:frt) (43), which has a chromosomal insert of constitutively expressed native green fluorescent protein (gfp), unfused to any protein and thereby with no deterministic spatial placement in the cell as a mature protein. Cells were grown in M9 minimal media (44) supplemented with 0.02 mg ml-1 of thiamine, and 0.18 mg ml-1 of glucose as the carbon source. Protein levels were quantified by using GFP as a reporter. Because native GFP is estimated to have a diffusion rate of 9 $\mu m^2 sec^{-1}$ (31) and E. coli cell has a mean cross sectional area of about 3 μm^2, the protein is rapidly dispersed throughout a cell in less than 1 s. Because our fluorescence images were taken at 20 min intervals, the distribution of GFP densities in a mother cell, and consequently also in the daughters, is not diffusion limited. Rather, the different densities result from differential production or gene expression within the cells. The strain was kindly provided by Minsu Kim (Emory University).

Cell growth and microscope slides

Cells from -80° C glycerol stock were streaked onto agar plates. A single colony was inoculated into M9 media and grown at 37° C overnight. The following day the culture was diluted 1:100 in M9 and grown for 2 hours. One μl of the culture was then pipetted onto a 10 μl M9 agarose pad. The agarose pad was then was then flipped with the bacterial side down onto a 24 x 60 mm cover glass and placed over a 25 x 75 mm single depression slide sealed with vaseline (modified from earlier methods described in

(20,21,23) to fit inverted microscope). Individual cells from two different movies were followed through time lapse microscopy at 37˚ C until each grew into a micro-colony of 64 cells.

Time-lapse microscopy

Cells were imaged with an inverted microscope (Nikon Eclipse Ti-S), equipped with Nikon NIS-Elements AR control software, 100X objective (CFI Plan APO NA 1.4), external phase contrast rings for full intensity fluorescence imaging (FITC), fluorescence light source (Prior Lumen 200) with motorized shutter (Lambda 10-B Sutter SmartShutter), and camera (Retiga 2000R FAST 1394, mono, 12 bit). Phase contrast and fluorescence images were recorded every 2 and 20 minutes, respectively.

Image quantification and analysis

Fluorescence measurements were collected by tracing cell outlines on the phase contrast images, transferring the outlines to the corresponding fluorescence frame, and quantifying density of fluorescence inside the outline. Outlines were traced manually. Blind replicate outlines, made without any awareness of cell polarity, reproduced the same results. All fluorescence images were corrected by removing outliers, subtracting background, and deconvoluted to correct for diffraction scattering. The software ImageJ (NIH) was used for quantifying fluorescence densities, outlier removal, and background subtraction. Fluorescence measurements were collected by first tracing cell outlines on the phase contrast images and the corresponding fluorescent frame was processed as following: The background of each frame was subtracted using "rolling ball" algorithm in ImageJ with ball radius 20 pixels. Noise created by heat overflow of single pixel was

corrected by "remove outliers" algorithm in ImageJ with threshold intensity difference 1000 and threshold radius 0.5 pixel. Deconvolution was accomplished by the Lucy-Richardson method in Matlab 2017b (The MathWorks, Inc., Natick, MA), (see Supplemental Material for details). Fluorescence measurements for pairs of old and new daughters were normalized by subtracting the mean of the pair's values. To calculate elongation rates, lengths of individual bacterial cells were extracted manually from recorded time-lapse images with ImageJ. From lengths compiled over time, the elongation rate r was estimated as the slope of a linear regression of (log/length) over time. A log transformation was used because elongation rates are known to be exponential (20). All lengths were measured immediately after division and prior the next division.

Statistical tests

All comparisons were evaluated by either t-tests or randomized designs. Details of sample sizes and choices of paired, unpaired, one- and two-tailed comparisons are provided in the figure legends. Randomized designs were used when data did not conform to standard Gaussian requirements. When appropriate, values are presented as mean ± SEM (standard error of the mean).

1.8 Acknowledgments

We thank Kevin Chi and Xiyu Liu for assistance.

1.9 References

1.   Thattai M, van Oudenaarden A. Intrinsic noise in gene regulatory networks. Proc Natl Acad Sci U S A [Internet]. 2001 Jul 17 [cited 2019 Jun 10];98(15):8614–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11438714

2.   Elowitz MB, Levine AJ, Siggia ED, Swain PS. Stochastic gene expression in a single Cell. Science (80-) [Internet]. 2002 [cited 2019 Jun 10];297:1183–6. Available from: www.sciencemag.org

3.   Raser JM, O'Shea EK. Noise in gene expression: origins, consequences, and control. Science (80-) [Internet]. 2005 Sep 23 [cited 2019 Jun 10];309(5743):2010–3. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16179466

4.   Rosenfeld N, Young JW, Alon U, Swain PS, Elowitz MB. Gene regulation at the single-cell level. Science (80-) [Internet]. 2005 Mar 25 [cited 2019 Jun 10];307(5717):1962–5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15790856

5.   Wang Z, Zhang J. Impact of gene expression noise on organismal fitness and the efficacy of natural selection. Proc Natl Acad Sci U S A [Internet]. 2011 Apr 19 [cited 2019 Jun 10];108(16):E67-76. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21464323

6.   Kiviet DJ, Nghe P, Walker N, Boulineau S, Sunderlikova V, Tans SJ. Stochasticity of metabolism and growth at the single-cell level. Nature [Internet]. 2014 Oct 3 [cited 2019 Jun 10];514(7522):376–9. Available from: http://www.nature.com/articles/nature13582

7.   Kærn M, Elston TC, Blake WJ, Collins JJ. Stochasticity in gene expression: from theories to phenotypes. Nat Rev Genet [Internet]. 2005 Jun 10 [cited 2019 Jun 10];6(6):451–64. Available from: http://www.nature.com/articles/nrg1615

8.   Raj A, van Oudenaarden A. Nature, nurture, or chance: stochastic gene expression and its consequences. Cell [Internet]. 2008 Oct 17 [cited 2019 Jun 19];135(2):216–26. Available from: https://www.sciencedirect.com/science/article/pii/S0092867408012439

9.   Sanchez A, Choubey S, Kondev J. Regulation of noise in gene expression. Annu Rev Biophys [Internet]. 2013 May 6 [cited 2019 Jun 27];42(1):469–91. Available from: http://www.annualreviews.org/doi/10.1146/annurev-biophys-083012-130401

10.   Huh D, Paulsson J. Random partitioning of molecules at cell division. Proc Natl Acad Sci [Internet]. 2011 Sep 6 [cited 2019 Jun 27];36(6):15004–9. Available from: https://www.pnas.org/content/108/36/15004

11.   Guptasarma P. Does replication-induced transcription regulate synthesis of the myriad low copy number proteins of Escherichia coli? BioEssays [Internet]. 1995 Nov 1

[cited 2019 Jun 18];17(11):987–97. Available from: http://doi.wiley.com/10.1002/bies.950171112

12.   Carey JN, Mettert EL, Roggiani M, Myers KS, Kiley PJ, Goulian M. Regulated stochasticity in a bacterial signaling network permits tolerance to a rapid environmental change. Cell. 2018 Mar 22;173(1):196-207.e14.

13.   Veening JW, Stewart EJ, Berngruber TW, Taddei F, Kuipers OP, Hamoen LW. Bet-hedging and epigenetic inheritance in bacterial cell development. Proc Natl Acad Sci U S A. 2008 Mar 18;105(11):4393–8.

14.   Huang S. Non-genetic heterogeneity of cells in development: more than just noise. Development [Internet]. 2009 Dec 1 [cited 2019 Jun 10];136(23):3853–62. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19906852

15.   Zernicka-Goetz M, Huang S. Stochasticity versus determinism in development: a false dichotomy? Nat Rev Genet [Internet]. 2010 Nov 28 [cited 2019 Jun 10];11(11):743–4. Available from: http://www.nature.com/articles/nrg2886

16.   Bergmiller T, Andersson AMC, Tomasek K, Balleza E, Kiviet DJ, Hauschild R. Biased partitioning of the multidrug efflux pump AcrAB-TolC underlies long-lived phenotypic heterogeneity. Science (80-). 2017;356(6335):311–5.

17.   Proenca AM, Rang CU, Buetz C, Shi C, Chao L. Age structure landscapes emerge from the equilibrium between aging and rejuvenation in bacterial populations. Nat Commun. 2018;9(1).

18.   Proenca AM, Rang CU, Qiu A, Shi C, Chao L. Cell aging preserves cellular immortality in the presence of lethal levels of damage. Kaeberlein M, editor. PLOS Biol [Internet]. 2019 May 23 [cited 2019 Jun 11];17(5):e3000266. Available from: http://dx.plos.org/10.1371/journal.pbio.3000266

19.   Chao L, Rang CU, Proenca AM, Chao JU. Asymmetrical damage partitioning in bacteria: a model for the evolution of stochasticity, determinism, and genetic assimilation. PLoS Comput Biol. 2016;12(1):1–17.

20.   Stewart EJ, Madden R, Paul G, Taddei F, Burland V. Aging and death in an organism that reproduces by morphologically symmetric division. Kirkwood T, editor. PLoS Biol [Internet]. 2005 Feb 1 [cited 2017 Jul 10];3(2):e45. Available from: http://dx.plos.org/10.1371/journal.pbio.0030045

21.   Rang CU, Peng AY, Chao L. Temporal dynamics of bacterial aging and rejuvenation. Curr Biol [Internet]. 2011;21(21):1813–6. Available from: http://dx.doi.org/10.1016/j.cub.2011.09.018

22.   Rang CU, Proenca A, Buetz C, Shi C, Chao L. Minicells as a damage disposal mechanism in Escherichia coli. 2018; Available from: https://doi.org/10.1128/mSphere

23.   Rang CU, Peng AY, Poon AF, Chao L. Ageing in Escherichia coli requires damage by an extrinsic agent. Microbiol (United Kingdom). 2012;158(6):1553–9.

24.   Chao L. A model for damage load and its implications for the evolution of bacterial aging. PLoS Genet. 2010;6(8).

25.   Lindner AB, Madden R, Demarez A, Stewart EJ, Taddei F. Asymmetric segregation of protein aggregates is associated with cellular aging and rejuvenation. Proc Natl Acad Sci [Internet]. 2008;105(8):3076–81. Available from: http://www.pnas.org/cgi/doi/10.1073/pnas.0708931105

26.   Urszula Ł, Glover G, Capilla-lasheras P, Young AJ, Pagliara S, Young AJ. Bacterial ageing in the absence of external stressors. Phil Trans R Soc B. 2019;374:20180442.

27.   Govers SK, Mortier J, Adam A, Aertsen A. Protein aggregates encode epigenetic memory of stressful encounters in individual Escherichia coli cells. Laub M, editor. PLOS Biol [Internet]. 2018 Aug 28 [cited 2019 Jun 19];16(8):e2003853. Available from: https://dx.plos.org/10.1371/journal.pbio.2003853

28.   Lele UN, Baig UI, Watve MG. Phenotypic plasticity and effects of selection on cell division symmetry in Escherichia coli. Kaeberlein M, editor. PLoS One [Internet]. 2011 Jan 10 [cited 2019 Dec 17];6(1):e14516. Available from: https://dx.plos.org/10.1371/journal.pone.0014516

29.   Schaechter M, Maaløe O, Kjeldgaard NO. Dependency on medium and temperature of cell size and chemical composition during balanced growth of Salmonella typhimurium. J Gen Microbiol [Internet]. 1958 [cited 2019 Jun 12];19:592–606. Available from: www.microbiologyresearch.org

30.   Balleza E, Kim JM, Cluzel P. Systematic characterization of maturation time of fluorescent proteins in living cells. Nat Methods. 2018 Jan 3;15(1):47–51.

31.   Mullineaux CW, Nenninger A, Ray N, Robinson C. Diffusion of green fluorescent protein in three cell environments in Escherichia coli. J Bacteriol. 2006 May;188(10):3442–8.

32.   Cormack BP, Valdivia RH, Falkow S. FACS-optimized mutants of the green fluorescent protein (GFP). In: Gene. Elsevier B.V.; 1996. p. 33–8.

33.   Sundararaj S. The CyberCell Database (CCDB): a comprehensive, self-updating, relational database to coordinate and facilitate in silico modeling of Escherichia coli. Nucleic Acids Res. 2004 Jan 1;32(90001):293D – 295.

34.   Kjeldgaard NO, Kurland CG. The distribution of soluble and ribosomal RNA as a function of growth rate. J Mol Biol [Internet]. 1963 Apr 1 [cited 2019 Jun 12];6(4):341–8. Available from: https://www.sciencedirect.com/science/article/pii/S0022283663800935

35.   Poulsen LK, Licht TR, Rang C, Krogfelt K a, Molin S. Physiological state of Escherichia Coli BJ4 growing in the large intestines of streptomycin-treated Mice. J Bacteriol [Internet]. 1995;177(20):5840–5. Available from: isi:A1995RZ77000018

36.   Rang CU, Licht TR, Midtvedt T, Conway PL, Chao L, Krogfelt KA, Cohen PS and Molin S. Estimation of growth rates of Escherichia coli BJ4 in streptomycin-treated and previously germfree mice by in situ rRNA hybridization. Clin Diagn Lab Immunol [Internet].   1999;6(3):434–6.   Available   from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=103738&tool=pmcentrez&ren dertype=abstract

37.   Rang C, Galen JE, Kaper JB, Chao L. Fitness cost of the green fluorescent protein in gastrointestinal bacteria. Can J Microbiol [Internet]. 2003;49(9):531–7. Available from: http://www.nrcresearchpress.com/doi/abs/10.1139/w03-072

38.   Perlman MD. Jensen's inequality for a convex vector-valued function on an infinite-dimensional space. J Multivar Anal. 1974;4(1):52–65.

39.   Coquel A-SS, Jacob J-PP, Primet M, Demarez A, Dimiccoli M, Julou T,Moisan L, Lindner AB and Berry H. Localization of protein aggregation in Escherichia coli is governed by diffusion and nucleoid macromolecular crowding effect. Shvartsman S, editor. PLoS Comput Biol [Internet]. 2013 Apr 25 [cited 2017 Jul 7];9(4):e1003038. Available from: http://dx.plos.org/10.1371/journal.pcbi.1003038

40.   Mortier J, Tadesse W, Govers SK, Aertsen A. Stress-induced protein aggregates shape population heterogeneity in bacteria. Curr Genet [Internet]. 2019 [cited 2019 Mar 4];65(1):11–6. Available from: https://doi.org/10.1007/s00294-019-00947-1

41.   Tyedmers J, Mogk A, Bukau B. Cellular strategies for controlling protein aggregation. Nat Rev Mol Cell Biol. 2010;11(11):777–88.

42.   Winkler J, Seybert A, König L, Pruggnaller S, Haselmann U, Sourjik V, Weiss M, Frangakis AS, Mogk A and Bukau B . Quantitative and spatio-temporal features of protein aggregation in Escherichia coli and consequences on protein quality control and cellular ageing. EMBO J. 2010 Mar 3;29(5):910–23.

43.   Kim M, Zhang Z, Okano H, Yan D, Groisman A, Hwa T. Need-based activation of ammonium uptake in Escherichia coli. Mol Syst Biol [Internet]. 2012 Sep 25 [cited 2019 Jun 18];8(1):616. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23010999

44.   Miller JH. Experiments in molecular genetics. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press; 1972.

Chapter 2: An aging landscape results from damage dynamics in single cell

2.1 Abstract

Aging researches started from multicellular organism as a physiological and populational phenomenon with underlying obscurity. With the rise of molecular and single cell approaches, data of cellular aging has accumulated. Taking this advantage, our model, for the first time, connects single cell damage dynamics to characteristics of an aging population. The model is built on damage dynamics of E.coli cell, and can precisely recapitulated features of an aging E.coli population under natural selection. The model also predicts populational asymmetry under different level of external damage and verified the single cell biophysical theory of damage dynamics simultaneously.

2.2 Background

Aging has been observed and studied as the progressive loss of function at the macromolecule, tissue, organ, or individual level, driven by the deterioration of intracellular processes [12]. Single cell organisms, among which E.coli and budding yeast are the most extensively studied under the topic of aging, are unique in that all aging processes occur in one cell, and the lifetime of an individual can be monitored by scale of hours. In addition, in single cell organisms, cellular aging is directly connected with aging-related molecular dynamics [13,14]. Within these two decades, the discovery of E.coli aging has particularly broadened the area by the observation of coexistence between immortality and aging, revealing an ancestral type of the evolution of aging [1-2,10,15].

In E.coli, aging is reported by experiments that show asymmetrical physiology among two daughter cells of the mother cell, leading to an aging landscape in E.coli population that directly determines population fitness [1,2,4,5,10]. Mechanistically, this is originated by asymmetrical distribution of macromolecules, especially damaged protein in form of aggregates, in the mother cell [3,6-9]. Protein aggregates have been shown significant correlation to the age and elongation rate of E.coli [8]. Due to the large size of aggregates, experiments and steady-state simulations proposed that aggregates are largely excluded from E.coli nucleoid and show passive enrichment and free Brownian motion in both cell poles (nucleoid exclusion model) [6,9]. When cell divides, the new pole is born with very little aggregates whereas old pole contains aggregates from mother cell. This is the source of asymmetrical damage distribution in

single cell, and ultimately spreading to the whole population throughout successive cell divisions.

However, two points are missing in the steady-state single cell aging model. First, elongation is an automatic dilution of damage and will further promote elongation that could result in less reception of new damage, starting a new dilution-elongation feedback on elongation, and vice versa. Besides single cell-level elongation, researches have tracked down subcellular elongation events by insertion of patches of new cell walls [18-21]. It worth examine if the pattern of insertion correlates to the age of cell pole since the distribution of local dilution-elongation event is important in shaping the damage distribution of the whole cell.

Second, although aging starts from asymmetrical distribution of macromolecules of single cell, aging landscape, as a group trait, evolves in population. Fisher's fundamental theorem provides argument of fitness advantage for evolution of standing fitness variation inside population [27], and aging has shown to contribute up to 37% to E.coli physiology variation measured by GFP intensity [26]. As mother cell divides, the fitness asymmetry of siblings prevail as population builds up by generations, shaped by in-population selection from exponential population expansion, and eventually reaches equilibrated aging landscape as standing fitness variation under selection. Previous numerical simulation has successfully predicted the old and new daughter doubling time equilibrium by assuming uniform single cell damage distribution and a constant segregation factor of damage in all mother cells [11]. Therefore more realistic simulation is needed to reconstruct asymmetrical damage segregation by simulating damage dynamics of singe cell and then rebuild the equilibrated population.

As a significant link between E.coli single cell aging and population aging, our model explored the dynamics of damage aggregates within single cell and examined the influence of internal and external factors by recreating the population aging landscape.

2.3 Results

Single cell elongation and division

A lineage is established from a single cell by successive elongation and division. Considering the difficulty of analytically redistributing damage during elongation, we start from numerical discrete simulation of single cell.

In the model, a single cell is born with seven linearly concatenated growth units, mimicking the uni-directional exponential elongation nature of E.coli. Each growth unit contains certain amount of damage that uniformly distributed within this unit, and all units have the same damage-free duplication probability. The single cell model is simulated by the time step of one minute. As time goes by, each unit makes independent decision on its duplication according to their duplication probability. As a unit duplicate, a new growth unit is added next to the old unit, and the damage in the old unit is shared equally among old and new unit.

It is not clear if each growth unit has its own elongation probability according to their damage content, though reports have confirmed the negative correlation between elongation rate and amount of damage aggregates in single cell level [8]. Therefore a transient labeling of cell wall is performed to quantify the amount of new elongation in a subcellular level (Fig 1). We observed significant higher elongation rate of new pole over old pole in old daughters, and non-significance in new daughters. This result is consistent with the experimental observation of higher damage deposit in the old pole over new pole in old daughters, and similar damage deposit across old and new pole in new daughters [17], demonstrating a similar negative correlation of elongation rate with

35

amount of damage in subcellular level as shown in whole cell [8]. It allows us to apply unique duplication probability to each growth unit being reversely correlated to its own damage content. To prevent negative probability in very high damage in the old poles, this damage-duplication relation is constrained by a first order hill function.

Cell division is set when number of growth unit doubled from birth. At division, growth units are evenly allocated to two daughter cells, from on the position of the division ring. All the damaged proteins stay in their original growth unit, manifesting the maternal effect of damage inheritance. For the sake of tracking cell lineage in a population, the two daughter cells are named according to their parent cell's name, and also according to the pole they inherit.

Single cell damage distribution dynamics

The major cause of aging is external and internal damage. In our model, damage distribution determines fitness of single cell and populational fitness variation. As we model each bacteria cell by one dimensional growth units, each unit carries a certain amount of damaged particles. The damage particles are either inherited from mother cell, or imposed uniformly by time onto every growth unit. We assume that all damage particles are identical hypothetical proteins, and cellular repair of damage is negligible.

Based on nucleoid exclusion model of cellular damage [6,9], a damage particle in the model performs constant diffusion rate across whole cell except much slower outward diffusion from the two growth units representing two cell poles. The behavior of each damage particle is best described by one dimensional random walk with Einstein's

36

diffusion equation [24,25]. Determined by simulation time step, each damage particle inside its growth unit creates a probability distribution. The proportion of distribution that stretches out of the original unit is the unit-crossing probability of one time step, in both directions. To simulate the effect of damage aggregation-disaggregation on cell poles, the outward diffusion rate from both ends of growth unit is reduced from the normal diffusion rate by a fraction.

In each time step, an in silico E.coli cell goes through an iteration that consists of uniform new damage accumulation, damage diffusion, aggregation-disaggregation, cell elongation, damage rearrangement and checkpoint of cell division. Fig 2 shows a typical simulated damage distribution of single cell close to division, with notable enrichment of damage at both poles. The distribution is consistent with experimental data [3].

Population growth and natural selection

No age structure is observed without a lineage of organism. In particular, age structure in E.coli has shown some of the most intriguing features originated from asymmetrical damage segregation [11,22]. As reported by experiment, when following the consecutive old pole inheritance (old daughter lineage) or new pole inheritance (new daughter lineage) from an arbitrary E.coli cell, the old daughter of old daughter lineage or new daughter of new daughter lineage converges to their respective equilibrated doubling time. And these two lineages serve as the low and high limit of fitness of population under physiological damage condition [11,15,22]. This is an important observation of prokaryote aging coexisting with immortality. When following consecutive

old and new daughter lineages in the model, we can successfully reproduce the two equilibrium points of the population. In addition, attractor behavior emerged around equilibrium due to probabilistic elongation of each growth unit, resembling the innate noise of E.coli elongation (Fig 3). Interestingly, experiments have also show similar stochasticity around old and new equilibrium points of doubling time [22].

To parameterize the model, the doubling time relation between mothers and daughters published from microfluidic experiment [22] is used. And cell length of birth in the model is obtained by same experiment. When mapped to published E.coli cytosolic diffusion dynamics [23], the size of damage particle in our model is estimated as 40.31MDa, with radius 39nm assuming comparable density of E.coli ribosome. This estimation is below the low size limit of aggregate that has been observed to be excluded by the nucleoid [9]. It is consistent with its ability of travelling between two cell poles. After parameterization, our model can quantitatively reproduce an aging E.coli population (Fig 4).

Asymmetry of an E.coli population

Aging has been argued having selective advantage as the advantage of standing variation of population fitness. The mechanistic origin of aging lies on physiological asymmetry of each individual. Previous model has assigned same factor of damage asymmetry across population, due to lack of knowledge of damage diffusion process [11]. Here we found the differential damage asymmetry according to the age of E.coli in the same population. Equilibrated new daughters are on average 20% more symmetrical than equilibrated old daughter (Fig 5). And as a cell ages from new

38

daughter equilibrium to old daughter equilibrium, asymmetry increases. The result is most likely due to the positive elongation-dilution feedback in new daughters, and less age difference between the poles of new daughters (1 generation) than old daughters (at least 2 generations). Similar results are reported by analyzing doubling time and elongation rate of siblings from new or old mother cells [26], and from single cell imaging of GFP distribution within new and old cells [26].

External damage is inevitable and is one of the most important sources of aging [3,6,17]. It varies due to environmental fluctuation. As nucleoid exclusion mechanism being a most supported mechanism of damage organization under physiological damage level, it worth examine its response to elevated external damage. To best mimic the random nature of external damage in the model, new damages are imposed uniformly to every growth unit. When new damage rate increases, we observed concurrent decrease of damage asymmetry of every dividing mother cell (Fig 6). This is in consistent with microfluidic experiment result, where external damage is imposed by phototoxicity [17]. The result reveals a limit of damage organization performed by nucleoid exclusion machinery. The inability of organizing all damage into the old pole is also the reason for the immortality of old lineage of E.coli, viewed as an ancestral form of aging.

2.4 Discussion

External and internal damage, provided its fitness harm and inevitability, can be organized to create positive fitness effect by generating standing variation of population. According to Fisher's fundamental theorem [27], a completely asymmetrical damage separation (0 and 100%) will be ideal to maximize population fitness. However, observed damage separation factor is 41 and 59% [11,22], clearly implicating the unknown trade-off on asymmetry, which can be explored mechanistically.

The advantage of damage nucleoid exclusion theory is that there is no energy cost of organizing damage, and the asymmetry is achieved completely by the size of free-diffusing damage and aggregated damage. In this way, the size of damage has direct evolutionary significance. Studies in other organism have shown the major source of aggregate being RNA binding protein [16], whose size varies in a wide range. But most of them, as free-diffusing damage, can shuffle between old and new poles. The advantage of damage shuffling before aggregation is that old pole will have chance to share damage with new pole, preventing damage overload and the death of old lineage (simulated and experimentally observed [17]). It requires the dimension of free-diffusing aggregate not being greater than the continuous free space between old and new poles. On the other hand, too much damage shuffling between poles will promote symmetry and weaken the advantage of asymmetry by generating variance. Therefore, as aggregated damage being impossible to leave the pole they reside, free-diffusing damage size is key to adjust population asymmetry and should be evolved into a moderate range that balances between the pros and cons of damage shuffling. Our

estimation of free-diffusing damage size falls in this range: its molecular weight, as if being a sphere, can tightly fit into the mesh of the nucleoid. In this way, the difference between theoretical and observed damage separation factor can be explained.

As our model has explored, the fluctuation of external damage can change the asymmetry of population dramatically. In the same way, the external damage level could leverage the optimal size and dynamics of free-diffusing and aggregated damage. It will be significant both in theory and clinical studies to predict damage parameters (size, aggregation and disaggregation rate, repair rate, damage sensitivity) computationally by using similar model, and evolve E.coli in different level of external damage and test its evolutionary response.
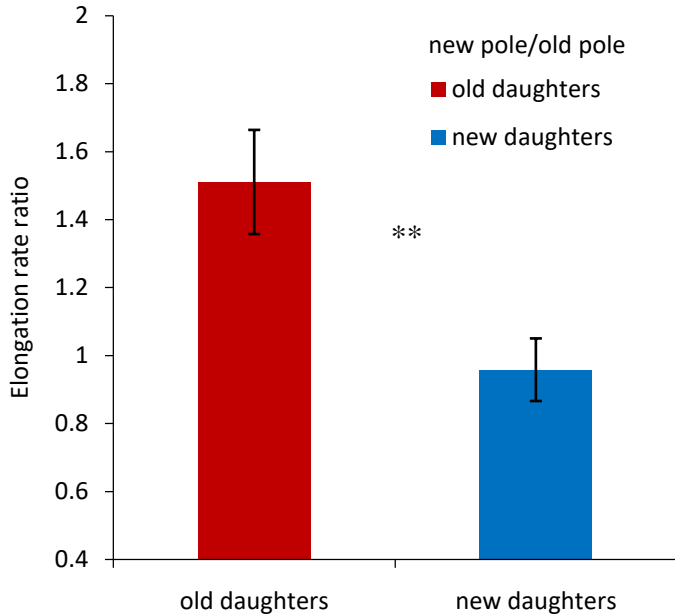
Figure 2.1 | Ratio of elongation rate between poles in old and new daughters. Elongation rate ratio is calculated as elongation rate of new pole versus old pole. Elongation rate of old and new poles are significantly different in old daughters (two-tailed paired t-test, n=28, p=0.00545), with elongation rate ratio 1.510±0.153. Elongation rate of old and new poles are not significantly different in new daughters (two-tailed paired t-test, n=28, p=0.366), with elongation rate ratio 0.957±0.0921. Old daughters show significantly higher asymmetry of elongation than new daughters (two-tailed paired t-test, n=14, p=0.00779). The differential elongation rate asymmetry between old and new daughters is inconsistent with experimental observation of differential aggregate asymmetry across daughters.
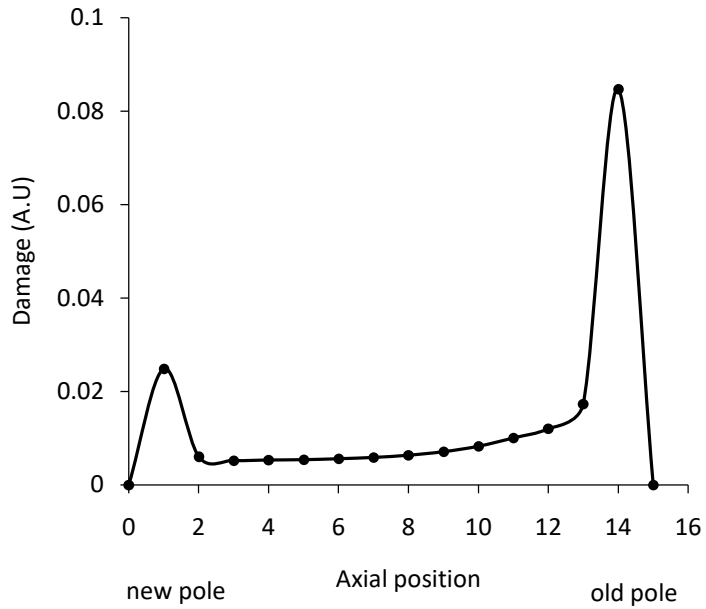
Figure 2.2 | Simulated 1-dimension damage distribution of a single cell. Damages show significant enrichment at both poles of a single cell. Inherited from mother cell, old pole shows higher level of damage enrichment than new pole. This is in consistent with nucleoid exclusion theory that damage aggregation occurs only at both poles.
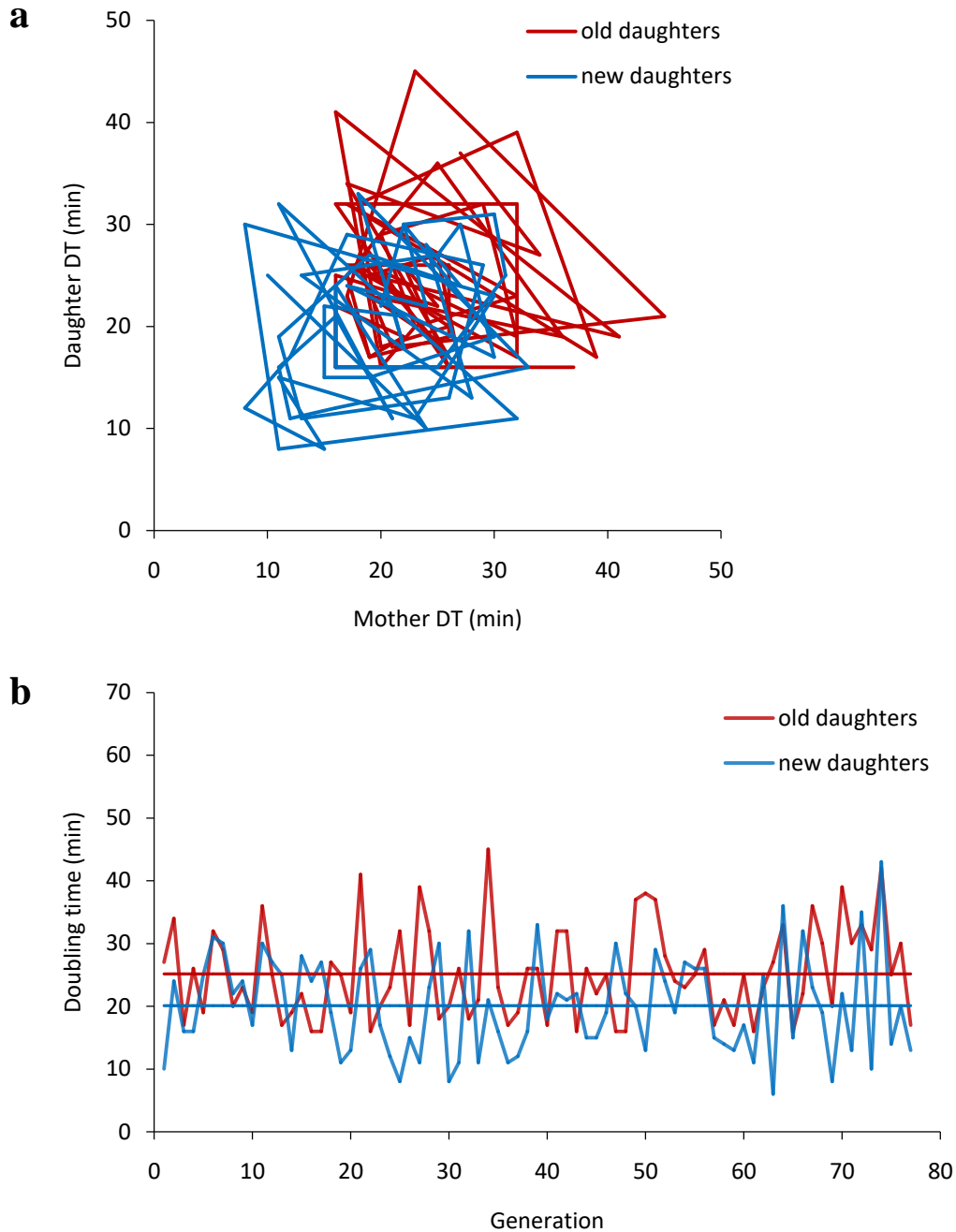
Figure 2.3 | Simulated equilibrium of doubling time of old and new lineages. a) phase trajectories of doubling time of two randomly chosen consecutive old and new generations show stability under fluctuations over 40 generations. b) same dataset plotted over 80 generations shows mean doubling times of old and new daughter generations equilibrate at 20.0779 min and 25.169 min.

44
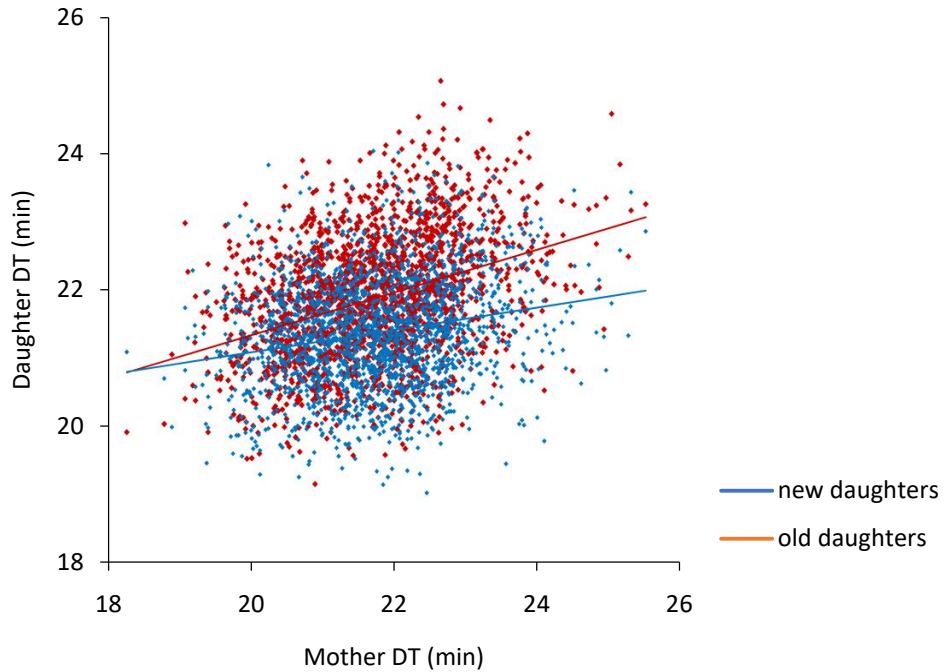
Figure 2.4 | Simulated phase graph of doubling time of an aging E.coli population in the mother machine. Old daughters (24.232±1.263 min, n=1000) showed longer doubling times than new daughters (20.939±0.927 min, n=1000). The separation of old and new sub-populations is significant (two-tailed paired t-test, n=1000, p<1e-10).
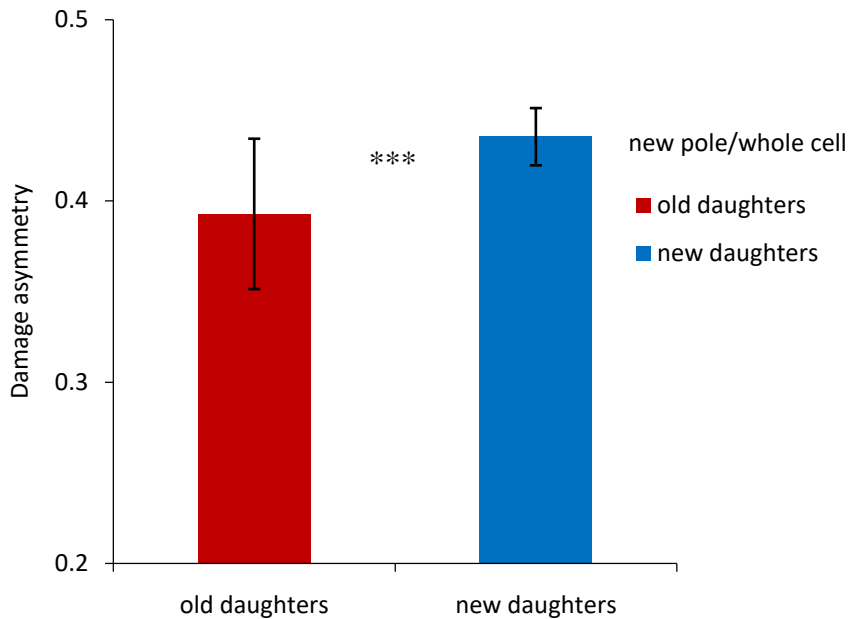
Figure 2.5 | Damage asymmetry in old and new daughters. Damage asymmetry is calculated as fraction of damage in new pole. 0.5 represents symmetry. Old daughters (0.393±0.0415, n=1000) show higher damage asymmetry than new daughters (0.435±0.0158, n=1000) (two-tailed paired t-test, n=1000, p<1e-10). Note that old daughters also show higher variance of asymmetry, which is in consistent with their more variant age difference between poles.
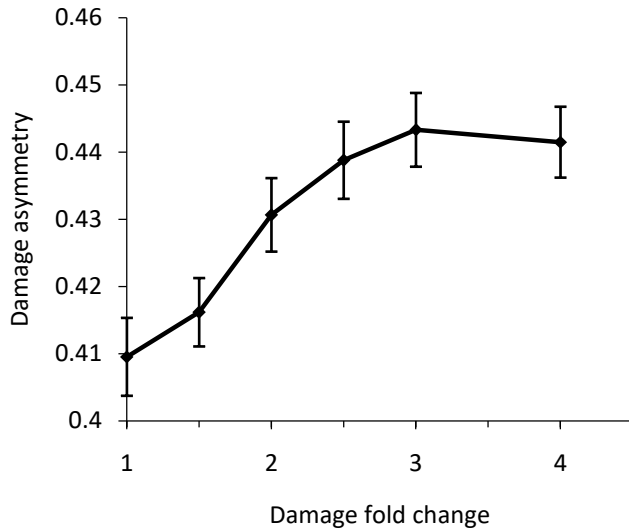
Figure 2.6 | Damage partitioning disrupted by increasing rate of external damage. Damage asymmetry is calculated as fraction of damage in new pole. 0.5 represents symmetry. External damage is imposed uniformly across the cell. As external damage rate elevating, the damage distribution in every cell becomes more uniform. Therefore damage partitioning between poles and following daughter cells will not be as effective.

2.6 Methods

WGA cell wall staining

To monitor the cell wall dynamics, we labeled cell wall by wheat germ agglutinin (WGA)- Alexa flour 488 conjugate from Life Technologies Corporation. Overnight culture is prepared from single colony by plating permanent E.coli K-12 MG1655 strain stock. Cells are then transferred to fresh media and grow in 37 degree to exponential phase and OD ~0.5. Then the culture is diluted 10 times into 1mL and incubated with 40 ul 1mg/mL fluorescent WGA. The sample was vortexed and kept shaking in dark incubator for 20 minutes. When the incubation is finished, the cells were spinned down and washed twice in fresh M9 media to clean the unbinded WGA conjugates. Then cells are placed on top of pre-sterilized Lb agar pad for imaging.

Single cell time lapse imaging

A Nikon Eclipse Ti-S microscope is used for imaging. And NIS-Element AR software is used for time lapse image acquisition. One phase contrast and one fluorescent image were taken at same time every 30 mins with 150 ms exposure. After imaging, cells are identified and fluorescent signals are measured with software ImageJ (NIH, https://imagej.nih.gov/ij)

Single cell modeling

The model of individual based E.coli aging population is performed by R 3.3.0.

Each individual cell is described by one dimension array with each array element representing a growth unit. Damage particles are localized in each growth unit and taking one dimension random walk to neighbor growth units depending on its size. We choose the length of growth unit in the following way. A very small growth unit will achieve highest resolution of damage distribution, but the simulation time for a whole cell will be very long since diffusion in each time step can across several neighboring growth units under reflective boundary condition. A smaller time step will help to hold a smaller distance of diffusion, but it takes more time step to simulate the doubling of single cell. On the other hand, a very big growth unit guarantees fast simulation of diffusion, but the resolution of damage distribution will be low and asymmetry will be weak. Taken these into consideration, growth unit size is designed to be 0.5um, with each cell born with 7 units and divide with 14 units. And time step is set to be one minute.

As cell elongates, the checkpoint of cell division is set to be at least twice as many of the initial number of growth unit (cell length doubled from birth), and the number of growth unit is checked in each time step. Different regime of emerging and positioning of the division ring in *E.coli* is reported, but majority of data supports the position of division ring being in the middle of dividing cell. Therefore in the simulation, when cell length doubled, cell divided from the middle.

Finite population modeling

To start as close to an equilibrated population, our starting population is built from simulated old and new equilibrated lineages. The population size 1000 is kept by

49

random culling at each time any individual divides. The population is simulated until the average doubling time of the population stabilizes.

Model parameterization

Parameterization is done by fminsearch function from R package pracma, with Hooke-Jeeves optimization algorithm. From the simulation of a finite bacteria population, we were able to reproduce four key observations we got from culturing bacteria under microscope in LB media. i.e. the equilibrate doubling time of old and new daughter, the relation of doubling time between two daughters and their mother. These observations are therefore used to calibrate three unknown parameters in the model. i.e. the size of free-diffusing damage particle (that determines diffusion rate under the relation of cytosolic mass-diffusion relation [23]), the disaggregation rate of aggregated damage (only applied in both poles), and new damage rate per time step.

2.7 Acknowledgments

2.8 References

1. Ackermann M, Stearns SC and Jenal U (2003) Senescence in a bacterium with asymmetric division. Science.300, 1920.

2. Stewart EJ, Madden R, Paul G and Taddei F (2005) Aging and death in an organism that reproduces by morphologically symmetric division. PLoS Biol. 3, 295–300.

3. Søren V, Harry N, Andrej K, Szabolcs S and Ala T (2017) Asymmetric Damage Segregation Constitutes an Emergent Population-Level Stress Response. Cell Systems , Volume 3 , Issue 2 , 187 – 198

4. Shapiro L, McAdams HH and Losick R (2002) Generating and exploiting polarity in bacteria. Science. 298, 1942–1946.

5. Stephens C (2005) Senescence: even bacteria get old. Curr.Biol. 15, R308–R310.

6. Winkler J., Seybert A., Ko¨ nig L., Pruggnaller S., Haselmann U., Sourjik V., Weiss M., Frangakis A.S., Mogk A and Bukau B (2010) Quantitative and spatio-temporal features of protein aggregation in Escherichia coli and consequences on protein quality control and cellular ageing. The EMBO Journal 29, 910–923

7. Rokney A., Shagan M., Kessel M., Smith Y., Rosenshine I and Oppenheim AB (2009) E. coli Transports Aggregated Proteins to the Poles by a Specific and Energy-Dependent Process. J. Mol. Biol. 392, 589–601. doi:10.1016/j.jmb.2009.07.009

8. Lindner AB., Madden R., Demarez A., Stewart EJ and Taddei F. (2007) Asymmetric segregation of protein aggregates is associated with cellular aging and rejuvenation. PNAS February 26, 2008. 105 (8) 3076-3081

9. Coquel AS., Jacob JP., Primet M., Demarez A., Dimiccoli M., Julou T., Moisan L., Lindner AB and Berry H (2013) Localization of Protein Aggregation in Escherichia coli Is Governed by Diffusion and Nucleoid Macromolecular

10. Ackermann M., Chao L., Bergstrom CT and Doebeli M (2007) On the evolutionary origin of aging. Aging Cell, 6: 235–244.Crowding Effect.PLoSComputBiol 9(4): e1003038. doi:10.1371/journal.pcbi.100303

11. Chao L (2010) A Model for Damage Load and Its Implications for the Evolution of Bacterial Aging. PLoS Genet 6(8): e1001076. doi:10.1371/journal.pgen.1001076

12. López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G.The hallmarks of aging.Cell. 2013;153: 1194–1217. pmid:23746838

13. Erjavec N, Cvijovic M, Klipp E, Nyström T. Selective benefits of damage partitioning in unicellular systems and its effects on aging. Proc Natl AcadSci U S A. 2008;105: 18764–18769. pmid:19020097

14. Laney SR, Olson RJ, Sosik HM. Diatoms favor their younger daughters. LimnolOceanogr. 2012;57: 1572–1578.

15. Rang CU, Peng AY, Chao L. Temporal dynamics of bacterial aging and rejuvenation. Curr Biol. 2011;21: 1813–6. pmid:22036179

16 . Proenca AM, Rang CU, Qiu A, Shi C, Chao L (2019) Cell aging preserves cellular immortalityin the presence of lethal levels of damage. PLoSBiol 17(5): e3000266.

17. Schlissel G, Krzyzanowski MK, Caudron F, Barral Y, Rine J. Aggregation of the Whi3 protein, not loss of heterochromatin, causes sterility in old yeast cells. Science. 2017;355(6330):1184-1187. doi:10.1126/science.aaj2103

18. Ursell TS, Trepagnier EH, Huang KC and Theriot JA (2012) Analysis of Surface Protein Expression Reveals the Growth Pattern of the Gram-Negative Outer Membrane. PLoSComputBiol 8(9): e1002680. doi:10.1371/journal.pcbi.1002680

19. Turner RD., Hurd AF., Cadby A., Hobbs JK and Foster SJ (2012) Cell wall elongation mode in Gram-negative bacteria is determined by peptidoglycan architecture. Nat Commun 4:1496 doi: 10.1038/ncomms2503

20. Huang KC., Mukhopadhyay R, Wena B., Gitai Z and Wingreen NS (2008) Cell shape and cell-wall organization in Gram-negative bacteria. PNAS 105 (49) 19282–19287

21. Ursell TS., Nguyen J., Monds RD., Colavin A., Billings G., Ouzounov N., Gitai Z., Shaevitz JW and Huang KC (2013) Rod-like bacterial shape is maintained by feedback between cell curvature and cytoskeletal localization. PNAS 111 (11) E1025-E1034

22. Proenca, A. M., Rang, C. U., Buetz, C., Shi, C. & Chao, L. (2018). Age structure landscapes emerge from the equilibrium between aging and rejuvenation in bacterial populations. Nature Communications 9

23. Kumar M, Mommer MS, Sourjik V. Mobility of cytoplasmic, membrane, and DNA-binding proteins in Escherichia coli. Biophys J. 2010 Feb 17;98(4):552-9. doi: 10.1016/j.bpj.2009.11.002. PMID: 20159151; PMCID: PMC2820653.

24. Einstein A (1905) Über die von der molekularkinetischenTheorie der WärmegeforderteBewegung von in ruhendenFlüssigkeitensuspendiertenTeilchen. Annalen der Physik. 322 (8): 549–560

25. von Smoluchowski M (1906). Zurkinetischen Theorie der BrownschenMolekularbewegung und der Suspensionen.Annalen der Physik. 326 (14): 756–780

26. Shi Chao, Chao Lin, Proenca Audrey Menegaz, Qiu Andrew, Chao Jasper and Rang Camilla U. (2020) Allocation of gene products to daughter cells is determined by

the age of the mother in single Escherichia coli cellsProc. R. Soc. B.2872020056920200569

27. Fisher, R.A. (1930) TheGenetical Theory of Natural Selection, Clarendon Press, Oxford

Chapter 3: Optimal bacterial size results from multi-linear elongation

3.1 Abstract

Optimal phenotypes can evolve by stabilizing selection (1,2). However, the characterization of the phenomenon in one example, rather than from several disparate ones, is rare (3,4). Here we show that an optimal birth length, previously undocumented, maximizes growth to division in E.coli. From an old study, we proved that E.coli length grew exponentially by a succession of linear phases with increasing but decelerating rates. Short cells are not optimal for their lengths span fewer phases than long cells, therefore elongate more linear. Very long cells are not optimal for the decelerating rates of elongation. Thus, the optimum is a birth length that maximizes the sum of growth at all lengths from birth to division. The discovery of optimal E.coli size provides a key answer to the origin of cell division and biological population.

3.2 Background

Reproduction is a hallmark of biological organism. By reproduction, population size grows exponentially. In single cell organism, population growth is achieved by continuous cell division. Since the invention of microscope, the elongation dynamics of E.coli has been widely reported as exponential (5-8), and cell width is unchanged (7,9). However this widely accepted view negates the necessity of cell division and the formation of an E.coli population because the population dynamics can be realized by a single cell, besides all the extra machinery and energy cost of cell division (16). This fact therefore calls for new evolution theory of cell division and more careful examine of single cell elongation dynamics.

With the development of single cell tracking and imaging, researchers have observed the elongation of E.coli is composed of two or more of linear phases with increasing rates (13), which has been regarded as exponential in low resolution observations. This brought up about new examining of fundamental growth unit of E.coli. The exponential view of E.coli elongation requires each unit biomass of cell contributes equally to the addition of new biomass. Therefore an E.coli cell is a population of fundamental growth units. This is supported by the observation that E.coli growth rate correlates clearly with count of ribosomes (10,11), and ribosome here, with its self-replicating nature, can be regarded as fundamental growth unit. On the contrary, the multiple linear elongation regime indicates larger scale and more integrated growth units that achieve self-replication for only one or few times during a doubling, revealed by the number of linear phases. One hypothesis of growth units is peptidoglycan factories for cell wall growth (13). They function as a unit and exist at integer numbers. The number

of functional factories increases step wise, and number of factories correlates with length.

The observation of multiple linear elongation gives us a new chance to revisit the evolution of cell division. Starting from massive microfluidic experiment, we firstly proved the precision of multiple linear elongation dynamics over exponential dynamics. Furthermore, our sample size allowed us to map out the length landscape that outside the equilibrated range of cell length, where our data identified an optimal cell size that emerges from multiple linear phases of elongation. The observation of optimal cell size for the first time provides support of cell division and origin of population.

3.3 Results

Observation of E.coli optimal birth length

We first verified that E. coli cell width did not change with length under our experimental conditions. We found no change and thereafter quantified cell growth by length. We examined next the relationship between elongation rate and birth length of single E. coli cells. As previous study has shown, we found a local fitness peak connected by higher fitness with long birth length, and no global optimal birth length is observed (Figure 3.1a). However, the distribution of birth length shows good consistency with abundant representation of local fitness peak but under-representation of long birth length (Figure 3.2b). It drew our interest to exploring cell elongation regime.

Characterizing of E.coli cell elongation

To re-examine the dynamics of E.coli elongation, we cultured E.coli strain MG1655 in LB media in microfluidic device, and time-lapse elongations were recorded every 1 min from 540 randomly chosen E.coli in the growth chamber (Figure 3.2a). All cells were then pooled by length. The mean length steps (δ, dμm per min) show a sigmoid relation with current length, featuring two phases of constant steps connected by a transition (Figure 3.2c). This is consistent with the observation of two linear elongation phases connected by an inflection in single cell data (Figure 3.2b), and also cell that divides early with no inflection (Figure 3.2b). The pooling assumed that cell elongation at a later length was independent of birth length. We tested the correlation between lengths of birth and δ within every bin of current length. No significance is

observed. Thus δ is independent of birth length and elongation is a Markovian process in E.coli.

When we further examine the shape of mean δ, the expectation of mean δ of bi-phasic growth is a step function, but the location of the step can be stochastic for different cells and the outcome is a sigmoid. And the expectation of a continuous exponential growth is linear. The sigmoid fit of mean δ being significantly better than linear fit (Akaike test, p=0.00366) proved that cell elongation is bi-phasic.To minimize the noise of step location in δ, we also picked the single cells that show significant bi-phasic growth and aligned their inflection point (Figure 3.2d). The average location of step is estimated as 41.74dμm. Therefore on average, cell that divides earlier than 41.74dμm will not cover next linear phase and will keep linear, whereas cell that exceeds 41.74dμm will cover next linear phase and will be bi-phasic. Both observation and simulation (by integrating δ) of single cell elongation trajectories proves this expectation (Figure 3.2e).

Landscape of multi-phasic growth

The shape of mean δ increases as a step indicates an underlyinggrowth unit production process as units increase with cell length, as discussed by Reshes et al (13). Assuming the equivalence of all growth units, the number of growth unit should be in equilibrium with the length of a cell, i.e. the growth unit number at division should be twice as the number at birth. Given the observation of single increase event of this number, the growth unit number can only be born at 2 units, increases to 3 and divides

at 4 units (Figure 3.3). And given 41.47dμm is the length at inflection, the length threshold of a new growth unit addition is 41.47/3 = 13.82dμm. From here, we predicted all the inflections on graph of mean δ, which has been obscured by limited sample size, by taking a integer multiple of this threshold.

Our data showed good consistency with predicted inflections (Figure 3.4), and the magnitude of each δ step is the midpoint of a linear fit on all binned δ in this step. From δ landscape, the growth trajectory of an arbitrary cell can be reconstructed by integrating δ from length of birth to death, and its division time and growth rate can be predicted.

Prediction of E.coli optimal birth length

All E.coli reaches length equilibrium when its birth length doubles before division at the middle. But when cells are away from equilibrium, different regimes (most prominently adder and sizer) are proposed to explain their returning to equilibrium. The relation of birth length and length change we found in our dataset (Figure 3.5a) shows intermediate pattern of adder and sizer. To simulate the equilibrium growth rates of cell of different birth lengths, we assumed all cells in our dataset had its birth length doubles before division. Therefore a fitness landscape of E.coli birth length is predicted and there we observed an optimal birth length (Figure 3.5b). This optimum was also the most common length (Figure 3.1b), which is most important because it indicates that the lengths in our sample were representative of a population under selection. The optima were 20% higher than the lowest rates, and bacterial populations are readily large enough to support natural selection of this magnitude (14,15).

3.4 Discussion

E.coli elongation has been viewed as exponential, which assume independence of self-replicating units, and therefore negate the significance of division and optimal birth length. Our discovery of growth unit number transits from 2 to 3 indicates units will have cooperated and mixed their products to assemble the third unit. Acquiring a third unit from two is equivalent to accruing early benefits by compound interest. Therefore it is also evolutionarily beneficial over model of non-cooperating units (Figure 3.3b,c). Cells with short birth lengths are slow because they divide before their length reaches the inflection. Short cells are born with one unit and divide at two. Since their units are isolated in different cells, there is no cooperation and a population of short cells increases exponentially only by division. Cells with long birth lengths benefit from more occurrence of compounding, but are hurt by decreased magnitude of δ steps. This could be well expected considering the distance limitation of cooperation between growth units. A constant δ step requires products of all growth units being concentrated at one position regardless of their localization in the cell, which will be difficult when cells are long. The fact that E.coli ribosomes and gene products are concentrated at the poles (12,17) could explain the decline if the polar units in long cells are under-supplied. Thus the decreasing δ steps indicating longer distance, local cooperation and limited product mixing. In short, short cells suffer from lack of compounding and linear growth, and long cells have limited cooperation offsetting their benefit of having more chance of compounding. Thus, medium cells are optimum and this is the evolutionary significance of cell division and potentially the primordial motivation of ecological population.

3.5 Availability of data and materials

The authors declare that all data supporting the findings of this study are available within the paper.
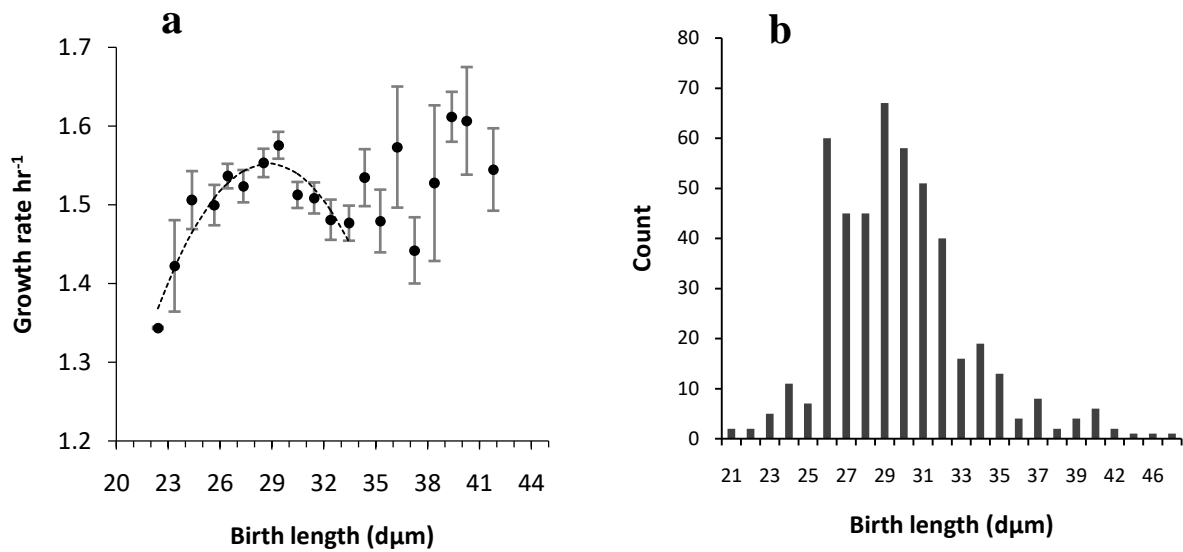
Figure 3.1 | Cell birth length affects growth rate in E.coli.(a) Growth rate (±SEM) relative to birth length in microfluidic device. Dashed line shows a local optimal which peaks at the highest enrichment of birth length in b). (b) Cell frequency relative to birth length in microfluidic device.
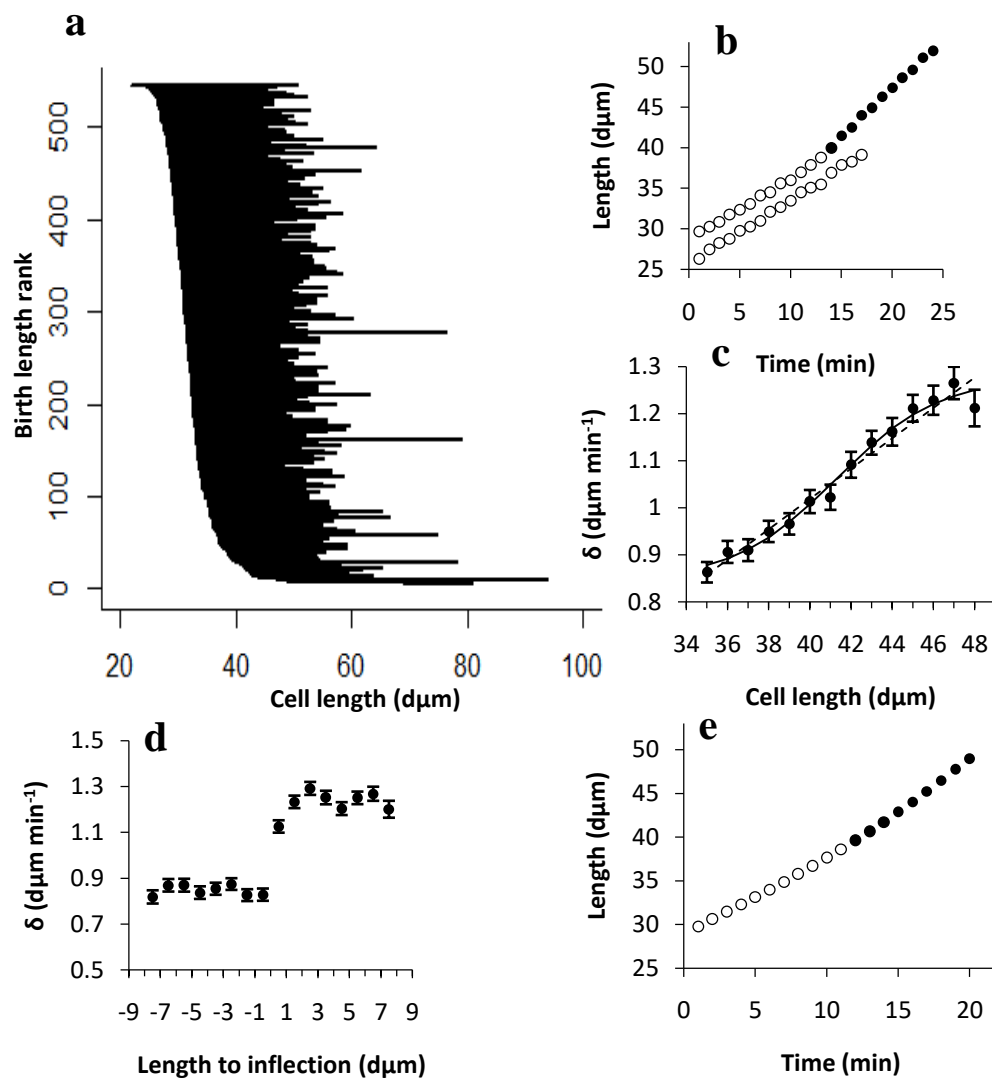
Figure 3.2 | E.coli growth relative to cell length. (a) Cell length from birth to division ranked by birth length. (b) Representative cell showing mono (open circle) and bi-phasic (open and close circle) growth from birth to division. (c) δ relative to cell length (±SEM) with tanh fit (solid line) and linear fit (dashed line). tanh fits significantly better (p=0.00366) than linear. The expectation of δ of exponential growth is linear. (d) δ relative to inflection point (±SEM) shows a clean step across two linear phases. The expectation of δ of bi-linear growth is step. (e) Empirical standard cell showing bi-phasic growth with no division

64
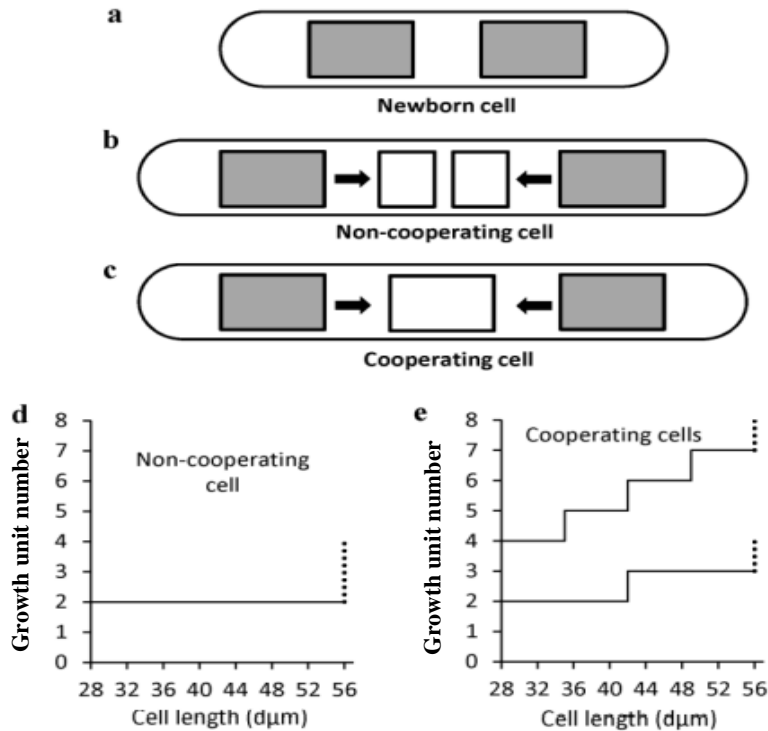
Figure 3.3 | Growth unit cooperation. (a) Cell born with 2 units. (b) New units built slowly without cooperation. (c) New functional unit built quickly with cooperation. (d) Growth unit number relative to length in Figure 2c cell. 3 units coexisting is not realized. Cell divides when growth unit number doubles to 4 but the step (dashed) coincides with division and is not observed. (e) Growth unit number relative to length in cooperating cell born with 2 units (lower line) or 4 units (upper line). 3 and 4, 5, 6, 7 units coexisting are realized and observed. 4 and 8 (dashed) units coexisting are realized but not observed.
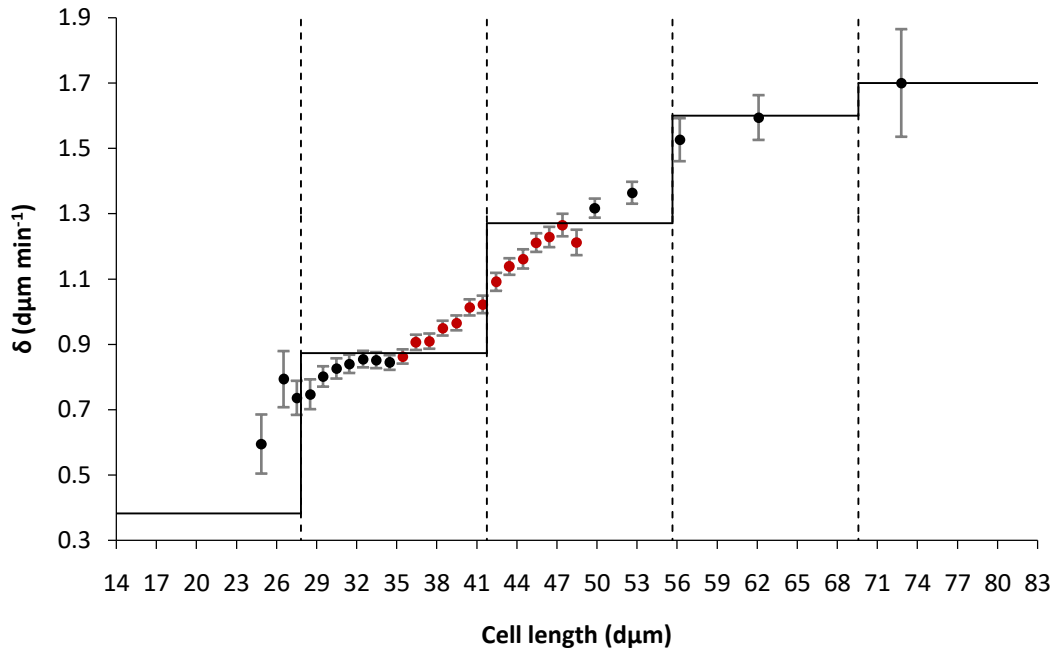
Figure 3.4 | Observed and fitted δ over current length of E.coli. δ is predicted on shorter and longer length of cell where sample size is limited. Red dots show area of Figure 3.2c where sample size exceeds 200 and inflection point is revealed by fitting tanh. Then all inflection points are predicted subsequently (dashed lines). Within each step between two inflections, expected δ (solid line) is obtained by fitting a polynomial line and predict midpoint. The expected δ is then used to construct empirical standard cell (Figure3.2e) and predict division time and growth rate.
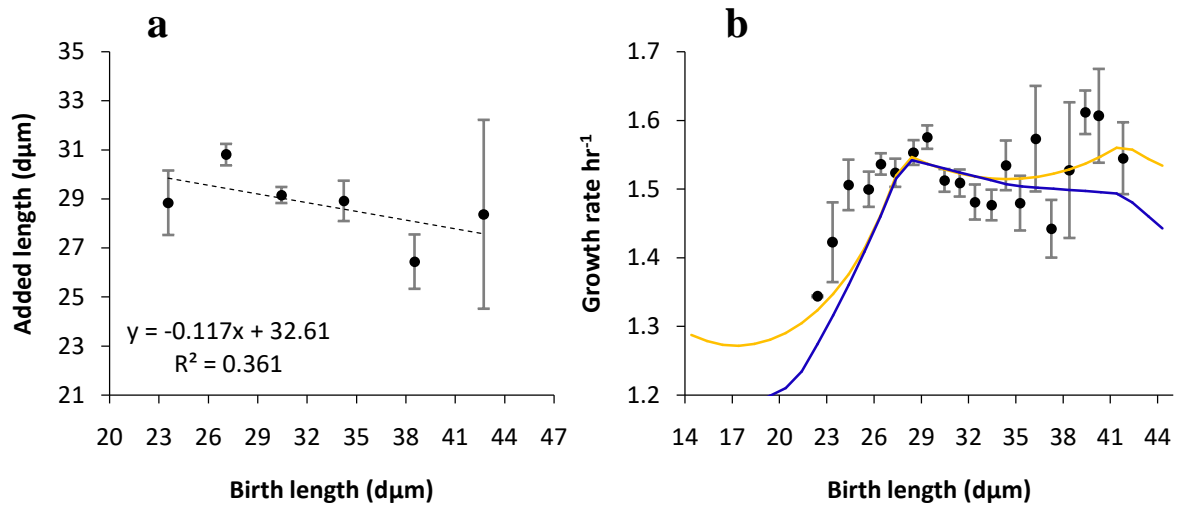
Figure 3.5 | E.coli elongation regime and optimal cell size.(a) The amount of E.coli growth before division (added length) relative to its birth length. A flat line is the expectation of "adder" regime. A straight line with slope -1 is the expectation of "sizer" regime. Our data shows an intermediate of both features. (b) Optimal cell size landscape realized by assuming doubler (division at twice of birth length, blue line) but not adder (yellow line). In one E.coli population, adder regime stabilizes the length of population, but is insufficient to give fitness landscape of populations of different birth lengths.

3.7 Methods

Cell culture and experiment conditions

K-12 E. coli wild-type strain MG1655 are grown in Luria-Bertani (LB) media overnight in 37 degree Celsius incubator before loaded with LB media to microfluidic device. The device is designed to have multiple channels in parallel with each row of culture chambers opening on both ends between two channels. The depth of the chamber confines E.coli colony to be mono-layer once cells are loaded by centrifugatopn. Microfluidic fabrication, cell loading, time-lapse imaging, and image collection and analysis have been described (5, 12). As cells start elongation and division, the chamber will be filled and cells will be pushed out from both side of chamber to the channels and being washed out to the outlet of device. Therefore the number of cells in a chamber will reach equilibrium. Study has shown (12) no significant enrichment of age or physiology being observed in the equilibrated chamber population caused by configuration or limitation of this device. The imaging is done on the bottom of selected culture chambers. After cell entering the chamber, the device is left running for 8.6 hours (i.e 23.37 generations with average generation time 22.11 minutes) before data acquisition to allow stabilization of cell physiology.

Data acquisition

The imaging process is directed by NIS-Elements AR software under Nikon Eclipse Ti-S microscope. Time-lapse phase images are taken every 20s throughout the whole experiment. The frequency of cell length sampling is one frame out of every 3

frames from birth to division. Images were analyzed with software ImageJ (NIH, https://imagej.nih.gov/ij). Statistical analysis was done by R version 3.3.0.

Data sets

The dataset for this study consists of 540 cells with 9853 length measurements. The cells are ranked by length of birth ranges from 21 to 50 dμm and every cell is represented by a trace of length span from birth to death. The data of following study starts from every cell at the length span 29 to 49 dμm where the collective sample for each specific length exceeds 200.

Length bins and measuring δ

To get the correct inflection point estimation, we starts from length span 29 to 49 dμm where the collective sample of δ for each 1 dμm bin exceeds 200. A tanh function is fit and inflection point is predicted. The inflection point is the transition of 2 to 3 growth unit. Therefore the length threshold of growth unit transition is inflection position/3. Due to the noise of inflection of single cell, the transition between linear phases is not a sharp step but a slope connecting neighbor linear phases. And similar slopes should be observed before and after these two linear phases because the existence of other linear phases. To eliminate the effect of slopes due to neighboring linear phases on the estimation of this inflection point, from the predicted inflection point and predicted length threshold, we picked the area half threshold before and after the inflection point to be least impacted by neighbor slopes, and fit a tanh function again. A new inflection point is

predicted and then its half threshold area is being fit again. This process iterates until the position of inflection point and predicted length threshold does not change.

As length threshold of growth unit transition is fixed, we can predict the length boundary of each linear phase by taking multiple of the threshold. δ of the first and second linear phases are averaged with bin size 1 dμm. As talked above, the beginning of third phase has 1 dμm bins that sample size of δ exceeds 200. For the rest of the phase, all δ are considered collectively and being divided to two bins that has sample size as close as possible. Same two bins are implemented in the fourth linear phase. And for the fifth linear phase, due to very limited sample size, all points are pooled in one bin.

To eliminate the noise of inflection on estimation of δ in each linear phase, a linear fit of the mean δ of all bins within a linear phase is performed, and its interpolation on the midpoint of this phase is used to represent the mean δ of the entire phase. The landscape of mean δ therefore appears step-like nature with each step representing constant δ within each linear phase of growth.

AIC (Akaike Information Criterion) test

The fit of models (e.g., sigmoid versus linear) to data was evaluated by likelihood ratio test and the AIC test was used to evaluate the likelihood probability using packages in R.

Detection of inflection point of single cell

Single cell growth curve is highly noisy. And the length trajectory of each single cell is measured by every 1 min. To find the inflection point of each cell during transition of 2 to 3 growth units, we firstly picked all the cells whose growth spans the length range of second and third linear phases. Then the growth trajectory of each cell is fit to one linear line. The same trajectory is then fit to bi-linear lines assuming inflection point being each length point from the first 3 to the last 3 time point. Each bi-linear fit is compared with linear fit by Bartlett's test of homoscedasticity. Then all the p-values from bi-linear fit are collected and the minimum is found by interpolation. If the minimum is less than threshold of 0.05, the minimum position is a significant inflection, and the trajectory of this cell is showing significant bi-linear behavior than linear.

Standard cells, estimating expected division time DiT and growth rate

The empirical standard cell growing over time was created to illustrate bi-phasic growth over time (Figure 3.2e). The cell, born with a length of L0 = 29 dµm, was constructed by taking the δ values in Figure 3.2c and using them to recreate the growing lengths. A cell of length L increases by a unit length to L+1 over a time interval of 1/δ. Thus, the time needed for a cell to elongate from L0 to Lt is $t = \sum 1/\delta L$ from L0 to Lt. The curve represents a cell that elongates without dividing. However, if Lt is the division length, then t becomes the division time DiT. Thus, although this standard cell was based on empirically determined δ values, an expected standard cell can also be generated from hypothetical δ values. Such an expected standard cell can then be used to predict expected DiT values given any pair of L0 and Lt. And $r = \log(Lt/L0)/DiT$.

3.8 Acknowledgements

Chapter 3, in full, is currently being prepared for submission for publication of the material: Chao Shi, Lin Chao, Audrey MenegazProenca, and Camilla U. Rang. "Optimal bacterial size results from multi-linear elongation." In preparation. The dissertation author was the primary investigator and author of this material.

3.9 References

1. Charlesworth, B., Lande, R. &Slatkin, M. A Neo-Darwinian Commentary on Macroevolution.Evolution 36, 474-498 (1982).

2. Uyeda, J. C., Hansen, T. F., Arnold, S. J. &Pienaar, J. The million-year wait for macroevolutionary bursts. Proceedings of the National Academy of Sciences 108, 15908-15913 (2011).

3. Kingsolver, J. G., Diamond, S. E., Siepielski, A. M. & Carlson, S. M. Synthetic analyses of phenotypic selection in natural populations: lessons, limitations and future directions. Evolutionary Ecology 26, 1101-1118 (2012).

4. Holand, H. k., Kvalnes. Thomas.,Roed. Knut H., Holand. Oystein.,Saether. Bernt-Erik and Kumpula.Jouko.Stabilizing selection and adaptive evolution in a combination of two traits in an arctic ungulate. Evolution 74, 103-115.

5. Proenca, A. M., Rang, C. U., Qiu, A., Shi, C. & Chao, L. Cell aging preserves cellular immortality in the presence of lethal levels of damage. Plos Biology 17 (2019).

6. Schaechter, M., Williamson, J. P., Hood, J. R., Jr. & Koch, A. L. Growth, cell and nuclear divisions in some bacteria. J Gen Microbiol 29, 421-34 (1962).

7. Stewart, E. J., Madden, R., Paul, G. &Taddei, F. Aging and death in an organism that reproduces by morphologically symmetric division. Plos Biology 3, 295-300 (2005).

8. Kennard, A. S., Osella, M., Javer. A., Grilli. J., Nghe. P., Tans. S., Cicuta.P. and Lagomarsino. Marco Cosentino. Individuality and universality in the growth-division laws of single E. coli cells.Phys Rev E 93, 012408 (2016).

9. Taheri-Araghi, S.,Bradde. Serena., Sauls. John T., Hill. Norbert S., Levin. Petra Anne., Paulsson. Johan.,Vergassola., Massimo and Jun. Suckjoon. Cell-Size Control and Homeostasis in Bacteria.Current Biology 25, 385-391 (2015).

10. Ecker, R. E. &Schaechter, M. Ribosome Content and the Rate of Growth of Salmonella Typhimurium. BiochimBiophysActa 76, 275-9 (1963).

11. Kjeldgaard, N. O. & Kurland, C. G. The distribution of soluble and ribosomal RNA as a function of growth rate.Journal of Molecular Biology 7, 341-348 (1963).

12. Proenca, A. M., Rang, C. U., Buetz, C., Shi, C. & Chao, L. Age structure landscapes emerge from the equilibrium between aging and rejuvenation in bacterial populations. Nature Communications 9 (2018).

13. Reshes, G., Vanounou, S., Fishov, I. & Feingold, M. Cell shape dynamics in Escherichia coli. Biophys J 94, 251-64 (2008).

14. Ohta, T. & Gillespie, J. H. Development of Neutral and Nearly Neutral Theories. TheorPopulBiol 49, 128-42 (1996).

15. Chao, L. Evolution of sex and the molecular clock in RNA viruses. Gene 205, 301-8 (1997).

16. Sánchez-Gorostiaga A, Palacios P, Martínez-Arteaga R, Sánchez M, Casanova M, Vicente M. Life without Division: Physiology of Escherichia coli FtsZ-Deprived Filaments. mBio. 7(5):e01620-16. doi:10.1128/mBio.01620-16 (2016)

17. Chao Shi, Lin Chao, Audrey MenegazProenca, Andrew Qiu, Jasper Chao and Camilla U. Rang. Allocation of gene products to daughter cells is determined by the age of the mother in single Escherichia coli cells. Proc BiolSci 287, 20200569 (2020).