Self-Consistent Examination of Donachie's Constant Initiation Size at the Single-Cell Level

Sattar Taheri-Araghi [1,2*]

[1] Department of Physics, University of California, San Diego, La Jolla, CA, USA, [2] Department of Physics and Astronomy, California State University, Northridge, Northridge, CA, USA

How growth, the cell cycle, and cell size are coordinated is a fundamental question in biology. Recently, we and others have shown that bacterial cells grow by a constant added size per generation, irrespective of the birth size, to maintain size homeostasis. This "adder" principle raises a question as to when during the cell cycle size control is imposed. Inspired by this question, we examined our single-cell data for initiation size by employing a self-consistency approach originally used by Donachie. Specifically, we assumed that individual cells divide after constant C + D minutes have elapsed since initiation, independent of the growth rate. By applying this assumption to the cell length vs. time trajectories from individual cells, we were able to extract theoretical probability distribution functions for initiation size for all growth conditions. We found that the probability of replication initiation shows peaks whenever the cell size is a multiple of a constant unit size, consistent with the Donachie's original analysis at the population level. Our self-consistent examination of the single-cell data made experimentally testable predictions, e.g., two consecutive replication cycles can be initiated during a single cell-division cycle.

Keywords: cell size, adder principle, cell cycle, chromosome replication, critical initiation size, single-cell analysis

1. INTRODUCTION

The coordination between growth and the cell cycle is a fundamental aspect of cellular physiology. The classic work of Schaechter, Maaløe and Kjelgaard established the "growth law," which states that the average size of bacterial cells in steady-state growth condition scales exponentially with the respective average growth rate (Schaechter et al., 1958). This is one of the first quantitative principles in bacterial physiology. Another important quantitative principle is the bacterial cell cycle model, whose two cornerstone assumptions are (i) in balanced growth the duration of replication (C period) of *Escherichia coli* chromosome is constant independent of the growth condition and (ii) cell divides after a constant time (C + D period) has elapsed since replication initiation (Cooper and Helmstetter, 1968; Helmstetter, 1968; Cooper, 1969).

In an important work, Donachie studied the consequences of the growth law and the cell cycle model together (Donachie, 1968). He concluded that, if both models are correct, the size of the cell per origin at the moment replication is initiated should be constant for all growth conditions. Furthermore, if the two models are correct, then the growth law can be expressed using the measured C + D as

$$m(T) = m_0 2^{(C+D)/T},$$ (1)

where m is the average cell size and T is average cell doubling time. In other words, Donachie was able to make experimentally testable predictions by self-consistently examining the relationship between two different assumptions. Furthermore, conversely, the predicted relationship can be used to estimate $C + D$ using size m and the average doubling time T, which can be measured and tested independently. In Appendix A, we present another example of self-consistency check, i.e., by self-consistently combining the initiator model (Cooper, 1969; Helmstetter, 1969) and the cell cycle model, we can show that the growth law emerges.

In recent years, single-cell experiments have significantly improved our understanding of growth and cell-size control in bacteria [For a review see Taheri-Araghi et al. (2015b) and discussions therein]. Single-cell data reveal information about fluctuations, heterogeneity and correlations between measurable parameters, which are masked in population measurements. In particular, we and others have shown that bacteria employ an "adder" principle to maintain size homeostasis during steady-state growth (Campos et al., 2014; Taheri-Araghi et al., 2015a). That is, cells grow by a constant size from birth to division, irrespective of the birth size. This automatically ensures that deviations in cell-size are corrected within a few generations. The adder principle however raises an important issue of when during the cell cycle size control is imposed.

This work presents a single-cell version of Donachie's analysis to our data in Taheri-Araghi et al. (2015a). We assume that $C + D$ is constant for all cells. Using this assumption, we retrace $C + D$ minutes backward in time from each cell division to extract a hypothetical initiation size of individual cells. We then ask if these assumptions lead to constant initiation size at the single-cell level. We found that, if the $C + D$ period is indeed constant for all cells, the constant initiation size is consistent with the adder principle at the single-cell level. Another prediction of our self-consistent analysis is that a cell can initiate two rounds of replication between birth and division. These predictions can be tested experimentally to verify the validity of the assumptions.

2. MATERIALS AND METHODS

2.1. Experimental Data on Growth and Division of *E. coli*

We used experimental data of cell length vs. time for seven different growth conditions for *E. coli* reported in Taheri-Araghi et al. (2015a). The media, average generation time, and average newborn size of cells are listed in **Table 1**. For the details of the experiments and growth media see Taheri-Araghi et al. (2015a) and its Supplementary Material. For the details of the single-cell growth experiment see Wang et al. (2010).

2.2. Retracing Length vs. Time Data to Infer Initiation Size

We apply the cell cycle model by Helmstetter and Cooper (Cooper and Helmstetter, 1968; Helmstetter, 1968; Cooper, 1969) to infer the initiation size. That is, we assume that individual cells initiate replication $C + D$ minutes prior to cell

TABLE 1 | Name of the growth conditions, average generation time, and average cell size at birth.

Name of growth media	Generation time (minutes)	Size at birth (μm^3)
TSB	17.1	2.73
Synthetic Rich	22.5	1.64
Glucose+12 a.a.	26.7	1.04
Glucose+6 a.a.	30.2	0.80
Glucose	37.7	0.59
Sorbitol	50.8	0.46
Glycerol	51.3	0.42

division (**Figure 1A**). We estimate $C + D$ self-consistently by fitting the population average size vs. growth data from Taheri-Araghi et al. (2015a) to Equation (1). The fitting outcome is that $C + D = 69$ min.

Since we do not have direct experimental data on the actual fluctuations of C and D periods, we cannot quantify the error arising from the retracing method. However, we can add noise to $C + D$ extracted by fitting data to Equation (1), and use it to check robustness of our conclusions. In Appendix B, we present a detailed discussion on the effect of noise in $C + D$. We find that the predictions of our analysis are robust to random fluctuations in the C and D periods, unless the added noise is larger than $\gtrsim 20\%$ of the generation time (**Figure A2**).

We provide a final self-consistency check that our single-cell analysis agrees with the population level results in Appendix C.
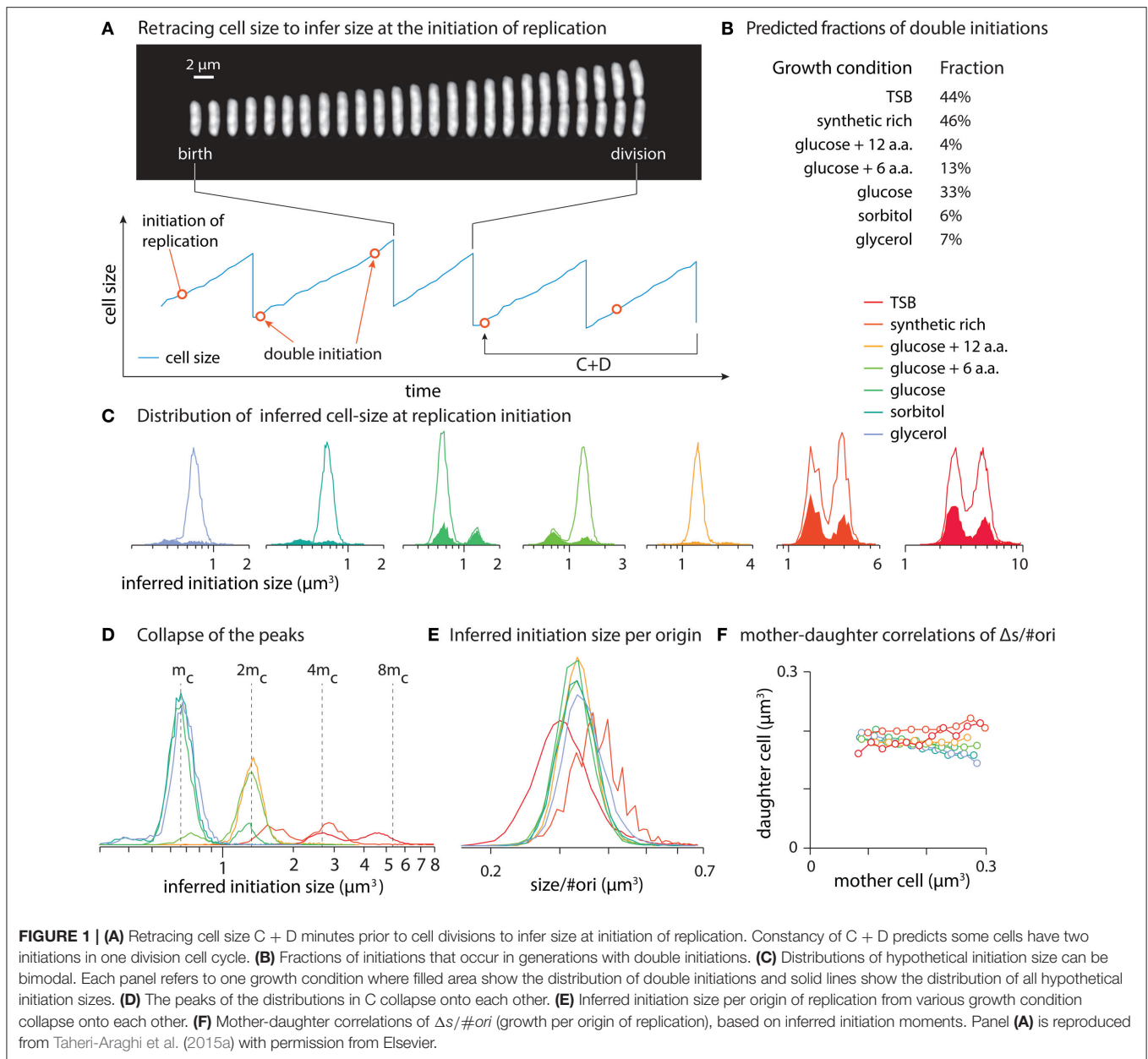
3. RESULTS AND DISCUSSION

3.1. Distribution of Inferred Initiation Size Shows Distinct Peaks, Consistent with Donachie's Constant Initiation Size Model

We computed distributions of hypothetical initiation size by retracing the single-cell length vs. time data for seven different growth conditions (**Figure 1B**). All distributions showed peaks. An obvious question is whether these peaks are multiples of constant initiation size as Donachie inferred from population data. To answer this question, we overlaid the distributions (**Figure 1D**).

Indeed, we found that the peaks of the inferred initiation size distributions collapse onto each other, with the peak positions increase in exponent of 2 from the position of the left-most peak. We then calculated inferred initiation size per replication origin (**Figure 1E**). Distributions from various growth conditions collapse on each other in the form of single-peak distributions. This is consistent with the model that replication initiates whenever the cell size per origin reaches a constant critical size, regardless of the growth condition (Donachie, 1968; Pritchard, 1968). (In Appendix D we show how the number of replication origins is calculated.)

A prediction of our self-consistent analysis is the possibility of double initiations. For significant fractions of subpopulations of cells, retracing by constant $C + D$ predicted two initiations

FIGURE 1 | (A) Retracing cell size C + D minutes prior to cell divisions to infer size at initiation of replication. Constancy of C + D predicts some cells have two initiations in one division cell cycle. **(B)** Fractions of initiations that occur in generations with double initiations. **(C)** Distributions of hypothetical initiation size can be bimodal. Each panel refers to one growth condition where filled area show the distribution of double initiations and solid lines show the distribution of all hypothetical initiation sizes. **(D)** The peaks of the distributions in C collapse onto each other. **(E)** Inferred initiation size per origin of replication from various growth condition collapse onto each other. **(F)** Mother-daughter correlations of Δs/#ori (growth per origin of replication), based on inferred initiation moments. Panel **(A)** is reproduced from Taheri-Araghi et al. (2015a) with permission from Elsevier.

separated by growth of a constant size per origin between them within a single generation (**Figures 1B–E**). This is not what is expected from the basic assumptions of the cell cycle control, which requires one-to-one correspondence between replication cycle and division cycle (Mitchison, 1971). Since this prediction seemingly violates a basic assumption, direct experimental test at the single-cell level will be important.

3.2. Conditions for Consistency of Constant (C + D) Model with Adder Principle

Another important question is whether the Helmstetter-Cooper model based on constant C + D is consistent with the adder principle. The organized pattern of inferred initiation size in **Figures 1D,E** can support such consistency. Unfortunately, with our current data we cannot answer whether replication starts at a critical size or after the cell grows for a constant size per origin from previous initiation. However, we can test if the constant C + D assumption and the adder principle are consistent by mother-daughter correlations. In **Figure 1F**, we show that there are no significant correlations between the mother and the daughter cells in terms of added size per origin ($\Delta s/\#ori$), as expected by the adder principle. That is, growth of the daughter cell by a constant $\Delta s/\#ori$ between initiation events is independent of the mother. Since $\Delta s/\#ori$ has been estimated by the constant C + D assumption, our analysis suggest that the two assumptions are mutually consistent.

Growth by a constant size per origin is consistent with the classic initiator model by Helmstetter and Cooper, stating that chromosome replication starts once the accumulation of initiators reach a critical threshold level (Cooper, 1969; Helmstetter, 1969). A feedback mechanism was proposed by Sompayrac and Maaloe (1973) to maintain initiator level proportional to cell size. We showed in Appendix A how the initiator and the cell cycle model by Helmstetter and Cooper can lead to the growth law.

While the initiator model seems plausible for the coordination of cell size and the replication cycle, there are experimental data that cannot be explained by the initiator model. For example, it has been shown that both an ectopic origin and the original wildtype origin initiate simultaneously without significant changes in growth kinetics (Wang et al., 2011). Another example is synchronous replication of minichromosomes that carry similar origin of replication in cells (Messer et al., 1978; Leonard and Helmstetter, 1986). In these examples, the relationship between size and number of origins do not follow the wild-type. At this point, we do not have sufficient experimental evidence to confirm the initiator model and the critical size for initiation and its link to the adder principle. Nevertheless, one way to reconcile a consistency between adder and constant C + D is to have an adder-like behavior for cell size at the initiation of chromosome replication.

3.3. Future Work

In this work, we applied Donachie's self-consistent analysis to the single cell data we reported recently. With the assumption that C + D is constant for individual cells, our analysis makes two predictions that can be directly tested experimentally in the future work: (i) double initiations of chromosome replication in one division cycle, and (ii) growth by a constant size between two consecutive replication initiations. Single-cell level test of these predictions will clarify whether our assumption of constancy of C + D is valid. Cell-size dependency or large fluctuations of C + D can change these predictions.

Several recent models discussed various size control routes in bacteria (Amir, 2014; Campos et al., 2014; Iyer-Biswas et al., 2014a,b; Kennard et al., 2014; Osella et al., 2014; Taheri-Araghi et al., 2015a). An interesting, unresolved question is how size control principles align with the cell cycle control. For a conclusive answer, we need direct experimental data on the progression of cell cycle in individual cells.

Finally, while adder principle appears general for all bacterial organisms tested so far, eukaryotes are not perfect adder (Jun and Taheri-Araghi, 2015). Further insights on the molecular mechanism of the adder principle can be gained through experimental tests in which we can perturb the perfect adder. Previously, perturbation of cell division machinery has been experimentally linked to variations of cell size (Weart et al., 2007; Chien et al., 2012; Hill et al., 2013). The timing of replication initiation was also linked to cell size, where *E. coli* mutants of smaller size delay initiation until they reach the appropriate initiation size (Hill et al., 2012). Interestingly, a modest over expression of DnaA-ATP can recover the replication initiation timing. We believe experiments on wild-type or size mutants in which the rate of accumulation of possible initiators can be temporarily decoupled from cell size (with overexpression or inhibition of their expression) will reveal valuable information on the regulation of cell size and the coordination of the cell cycle with cell size.

REFERENCES

Amir, A. (2014). Cell size regulation in bacteria. *Phys. Rev. Lett.* 112, 208102. doi: 10.1103/PhysRevLett.112.208102

Campos, M., Surovtsev, I. V., Kato, S., Paintdakhi, A., Beltran, B., Ebmeier, S. E., et al. (2014). A constant size extension drives bacterial cell size homeostasis. *Cell* 159, 1433–1446. doi: 10.1016/j.cell.2014.11.022

Chien, A. C., Hill, N. S., and Levin, P. A. (2012). Cell size control in bacteria. *Curr. Biol.* 22, R340–R349. doi: 10.1016/j.cub.2012.02.032

Cooper, S. (1969). Cell division and DNA replication following a shift to a richer medium. *J. Mol. Biol.* 43, 1–11. doi: 10.1016/0022-2836(69)90074-6

Cooper, S., and Helmstetter, C. E. (1968). Chromosome replication and the division cycle of *Escherichia coli* B/r. *J. Mol. Biol.* 31, 519–540. doi: 10.1016/0022-2836(68)90425-7

Donachie, W. D. (1968). Relationship between cell size and time of initiation of DNA replication. *Nature* 219, 1077. doi: 10.1038/2191077a0

Helmstetter, C. E. (1969). "Regulation of chromosome replication and cell division in *Escherichia coli*," in *The Cell Cycle,* eds G. M. Padilla, G. L. Whitson, and I. L. Cameron (New York, NY; London: Academic Press), 15–35.

Helmstetter, C. E. (1968). DNA synthesis during the division cycle of rapidly growing *Escherichia coli* B/r. *J. Mol. Biol.* 31, 507–518. doi: 10.1016/0022-2836(68)90424-5

Hill, N. S., Buske, P. J., Shi, Y., and Levin, P. A. (2013). A moonlighting enzyme links *Escherichia coli* cell size with central metabolism. *PLoS Genet.* 9:e1003663. doi: 10.1371/journal.pgen.1003663

Hill, N. S., Kadoya, R., Chattoraj, D. K., and Levin, P. A. (2012). Cell size and the initiation of DNA replication in bacteria. *PLoS Genet.* 8:e1002549. doi: 10.1371/journal.pgen.1002549

Iyer-Biswas, S., Crooks, G. E., Scherer, N. F., and Dinner, A. R. (2014a). Universality in stochastic exponential growth. *Phys. Rev. Lett.* 113, 028101. doi: 10.1103/PhysRevLett.113.028101

Iyer-Biswas, S., Wright, C. S., Henry, J. T., Lo, K., Burov, S., Lin, Y., et al. (2014b). Scaling laws governing stochastic growth and division of single bacterial

cells. *Proc. Natl. Acad. Sci. U.S.A.* 111, 15912–15917. doi: 10.1073/pnas.1403232111

Jun, S., and Taheri-Araghi, S. (2015). Cell-size maintenance: universal strategy revealed. *Trends Microbiol.* 23, 4–6. doi: 10.1016/j.tim.2014.12.001

Kennard, A. S., Osella, M., Javer, A., Grilli, J., Nghe, P., Tans, S., et al. (2014). Individuality and universality in the growth-division laws of single *E. coli* cells. ArXiv :1411.4321.

Leonard, A., and Helmstetter, C. E. (1986). Cell cycle-specific replication of *Escherichia coli* minichromosomes. *Proc. Natl. Acad. Sci. U.S.A.* 83, 5101–5105. doi: 10.1073/pnas.83.14.5101

Messer, W., Bergmans, H. E., Meijer, M., Womack, J. E., Hansen, F. G., and von Meyenburg, K. (1978). Mini-chromosomes: plasmids which carry the *E. coli* replication origin. *Mol. Gen. Genet.* 162, 269–275. doi: 10.1007/BF00268852

Mitchison, J. (1971). *The Biology of the Cell Cycle.* Cambridge, UK: Cambridge University Press.

Osella, M., Nugent, E., and Cosentino Lagomarsino, M. (2014). Concerted control of *Escherichia coli* cell division. *Proc. Natl. Acad. Sci. U.S.A.* 111, 3431–3435. doi: 10.1073/pnas.1313715111

Pritchard, R. H. (1968). Control of DNA synthesis in bacteria. *Heredity* 23, 467–476.

Schaechter, M., Maaløe, O., and Kjeldgaard, N. O. (1958). Dependency on medium and temperature of cell size and chemical composition during balanced growth of *Salmonella typhimurium*. *J. Gen. Microbiol.* 19, 592–606. doi: 10.1099/00221287-19-3-592

Sompayrac, L., and Maaloe, O. (1973). Autorepressor model for control of DNA replication. *Nat. New Biol.* 241, 133–135. doi: 10.1038/newbio241133a0

Taheri-Araghi, S., Bradde, S., Sauls, J. T., Hill, N. S., Levin, P. A., Paulsson, J., et al. (2015a). Cell-size control and homeostasis in bacteria. *Curr. Biol.* 25, 385–391. doi: 10.1016/j.cub.2014.12.009

Taheri-Araghi, S., Brown, S. D., Sauls, J. T., McIntosh, D. B., and Jun, S. (2015b). Single-cell physiology. *Annu. Rev. Biophys.* 44, 123–142. doi: 10.1146/annurev-biophys-060414-034236

Wang, P., Robert, L., Pelletier, J., Dang, W. L., Taddei, F., Wright, A., et al. (2010). Robust growth of *Escherichia coli*. *Curr. Biol.* 20, 1099–1103. doi: 10.1016/j.cub.2010.04.045

Wang, X., Lesterlin, C., Reyes-Lamothe, R., Ball, G., and Sherratt, D. J. (2011). Replication and segregation of an *Escherichia coli* chromosome with two replication origins. *Proc. Natl. Acad. Sci. U.S.A.* 108, E243–E250. doi: 10.1073/pnas.1100874108

Weart, R. B., Lee, A. H., Chien, A.-C., Haeusser, D. P., Hill, N. S., and Levin, P. A. (2007). A metabolic sensor governing cell size in bacteria. *Cell* 130, 335–347. doi: 10.1016/j.cell.2007.05.043

APPENDIX A

Consistency Between the Initiator Model, Growth Law and Critical Initiation Size

In this appendix we model a nutrient shift experiment where the timing of the initiation of chromosome replication and, thus, cell divisions are calculated based on the initiator model proposed and tested by Helmstetter and Cooper (Cooper, 1969; Helmstetter, 1969). We derive an analytical solution for cell size and the cell cycle of bacteria that experience a nutrient shift. From that, we show that the growth law and the critical size model emerge from the initiator model.

Below is the list of the assumptions of the model:

- Chromosome replication initiates once the initiators accumulate to a critical level. When a round of replication starts, initiators get destroyed.
- There is a constant time gap, $C + D$, between each initiation of replication and cell division.
- During steady-state and nutrient-shift the timing of initiations of replication are given by:

  (i) In steady-state: period of initiators accumulation up to the threshold is T, equivalent to cells' generation time.

  (ii) During nutrient shift from doubling times T_1 to T_2: If shift occurs at t_1 minutes after the last initiation, next initiation happens at $t_1 + t_2$, where $t_1/T_1 + t_2/T_2 = 1$ (Cooper, 1969).

- The cell grows exponentially and the size increase rate changes promptly upon a nutrient shift.

To begin with, consider a cell growing in steady-state condition with doubling time T_1. At time T_s after a division, the nutrient condition changes. The new condition imposes an eventual doubling time of T_2 in the final steady-state. Here we choose the reference of the time, $t = 0$, the birth of the cell in which nutrient shift happened (**Figure A1**). Thus, nutrient shift happens at $t = T_s$ during a cell cycle that the cell was expected to divide at $t = T_1$. The time of planned division at T_1 can change or remain unchanged depending on the timing of T_s with respect to rounds of chromosome replications. We take three steps to proceed with the calculation:

  (i) We find the timing of initiations of chromosome replication, both before nutrient shift and after nutrient shift.

  (ii) From the timing of replication initiations, we find timing of cell divisions, assuming every initiation results in a cell division after $C + D$ minutes.

  (iii) From the timing of cell division we calculate cell size considering that size increases exponentially with a rate instantaneously proportional to nutrient condition.

Since multiple cell cycles can overlap in bacteria, we assume at the time of nutrient shift, T_s, there are n cell cycles overlapping ($n = 0, 1, 2, ...$). The case $n = 0$ refers to the cells in slow growth condition if nutrient shift happens in the gap between the birth and initiation of chromosome replication (B period). There is a relationship between n, T_1, T_s, and $C + D$, the shorter the T_1

the larger the n can be. Also, n can vary depending on when the nutrient shift happens during the cell cycle. Without loss of generality, we choose not to elaborate further on this relationship as the final results can be expressed in terms of T_1, T_2, and $C + D$.

Timing of Chromosome Replications

Since there is a one-to-one correspondence between initiation of replication and cell division, the time gap between the n rounds of ongoing cell cycles at $t = T_s$ must be the generation time in the pre-shift condition, T_1. As the first cell division after $t = 0$ was scheduled at T_1, the "oldest" of the n overlapping cell cycles must have started at $T_1 - (C + D)$. The rest of them started every T_1 minutes thereafter. The initiation times of these n cell cycles are, thus, the following series:

$$T_1 - (C + D), 2T_1 - (C + D), ..., nT_1 - (C + D). \quad (A1)$$

Let's consider the next chromosome replication (the first after nutrient shift) starts at $t = T_r$, which depends on both time of nutrient shift, T_s, and generation time in the second growth condition, T_2. Next rounds of replication after T_r initiate every T_2 minutes. Thus, the timing of replication initiation after the nutrient shift is simply:

$$T_r, T_r + T_2, T_r + 2T_2, ... \quad (A2)$$

The time T_r can be calculated based on the third assumption of the model that we discussed earlier. Until the moment of nutrient shift, T_s, initiators have been accumulating since the start of the last round of replication, $t = nT_1 - (C + D)$, with the rate of $1/T_1$. [e.g., $3T_1 - (C + D)$ in **Figure A1**] From $t = T_s$ until $t = T_r$, initiators accumulate at rate $1/T_2$. The following equation can be solved for T_r,

$$\frac{1}{T_1}\left[T_s - nT_1 + (C + D)\right] + \frac{1}{T_2}(T_r - T_s) = 1, \quad (A3)$$

which yields

$$T_r = (n + 1)T_2 + T_s - \frac{T_2}{T_1}(T_s + C + D). \quad (A4)$$

Timing of Cell Divisions

Each initiation of chromosome replication leads to a cell division after a time gap of $(C + D)$. From Equations (A1) and (A2) and **Figure A1**, the cell division times from ongoing cell cycles and the ones starting after nutrient shift are:

$$\underbrace{T_1, 2T_1, ..., (n - 1)T_1, nT_1,}_{\text{divisions from ongoing cell cycles}}$$

$$\underbrace{T_r + (C + D), T_r + (C + D) + T_2, T_r + (C + D) + 2T_2, ...}_{\text{divisions from cell cycles that start after nutrient shift}}$$

$$(A5)$$

Cell-Size after the Nutrient Shift

Let's consider $m(T_1)$ denotes the newborn cell size during steady-state growth with generation time T_1. We aim to calculate $m(T_2)$,

FIGURE A1 | Cell cycle and nutrient shift. Here $n = 3$ cell cycles overlap at the moment of nutrient shift, $t = T_s$. In this example $T_1 > T_2$ and thus $m(T_1) < m(T_2)$. The three rounds of cell cycle started before nutrient shift lead to divisions at T_1, $2T_1$, and $3T_1$ (blue arrows). The first initiation of replication after the nutrient shift, sets up the start of the new steady-state (green arrows starting at T_r).

the newborn cell size at steady-sate with generation time T_2. Cells reach steady-state after the division at $T_r + (C + D)$, where the gap between cell divisions is T_2 and the size increase rate is $1/T_2$. Thus, the cell size after that division is the newborn cell-size in the new steady state, $m(T_2)$.

From $t = 0$ until $t = T_s$ cell size increases exponentially with rate $1/T_1$ (pre-shift rate). Upon nutrient shift, the rate of cell size increase changes instantaneously to $1/T_2$. The division at $t = T_r + (C + D)$ (corresponding to start of the new steady-state) is the $(n+1)th$ division after $t = 0$ (see Equation A5). Thus, the newborn cell size is given by

$$m(T_2) = \frac{1}{2^{n+1}} m(T_1) 2^{T_s/T_1} \times 2^{(T_r+C+D-T_s)/T_2}, \quad (A6)$$

where the first factor on the right-hand-side counts $(n + 1)$ divisions, the second term refers to the growth from birth at $t = 0$ until nutrient shift at $t = T_s$, and the third accounts for the size increase from the nutrient shift, $t = T_s$, until division at $t = T_r + C + D$.

Substituting T_r from Equation (A4) in Equation (A6), we get

$$m(T_2) = m(T_1) 2^{-(C+D)/T_1} \times 2^{(C+D)/T_2} \quad (A7)$$

Equation (A7) denotes that if cells grow in any steady-state condition with generation time T, newborn cell size is exponentially related to T,

$$m(T) = m_\circ 2^{(C+D)/T}. \quad (A8)$$

This is the growth law (Schaechter et al., 1958), with the exponent being $(C + D)$ if the relationship presented in the base of two. The exponent is consistent with Donachie's constant size at initiation of chromosome replication, as we elaborate below.

Cell-Size at Initiation of Replication
Consider that cells growing in a steady-state condition with generation time T and that up to N cell cycles overlap

($N = 1, 2, 3...$). That means from an initiation of chromosome replication until the cell division (this corresponds to a gap of $(C + D)$) we have N cell division events. For slowest growth conditions where cell cycles do not overlap we have $N = 1$. If $m_c(T)$ refers to cell size at which initiation occurs, the newborn cell size after the corresponding cell division is

$$\frac{1}{2^N} m_c(T) 2^{(C+D)/T} = m(T), \quad (A9)$$

where the first factor accounts for N divisions and the rest accounts for growth for $(C+D)$ minutes. Substituting $m(T)$ from Equation (A8), we obtain

$$m_c(T) = 2^N m_\circ. \quad (A10)$$

Since 2^N is the total number of replication forks when N cell cycle overlap, Equation (A10) shows that cell size per origin of replication at the initiation of replication is constant, independent of growth condition. This is the Donachie's observation in 1968 by combining the growth law and Helmstetter-Cooper model.

APPENDIX B

Noise and Uncertainty in Retracing Analysis

One may question the effect of noise in $C + D$ in retracing analysis and the extent it influences the distributions reported in **Figure 1**. The retracing analysis has an intrinsic "reading error." Since we do not have experimental data on the fluctuations of the durations of C and D periods for individual cells, we cannot completely avoid or quantify this reading error. However, in this appendix we present how we can choose a constant $C + D$ to minimize the error. Here, we show that adding any noise to retracing time, $C + D$, merely adds *extra* reading error and does not capture actual fluctuations of C and D periods.

FIGURE A2 | (A) To test the effect of noise in the analysis, a Gaussian noise δ_n with standard deviation of standard deviation of σ_n is added to the retracing time. **(B)** Various levels of noise is tested with standard deviations, σ_n, chosen at 0, 2.5, 5, 7.5, and 10% of C + D. **(C)** Solid lines refer to the distribution from whole cell population and filled 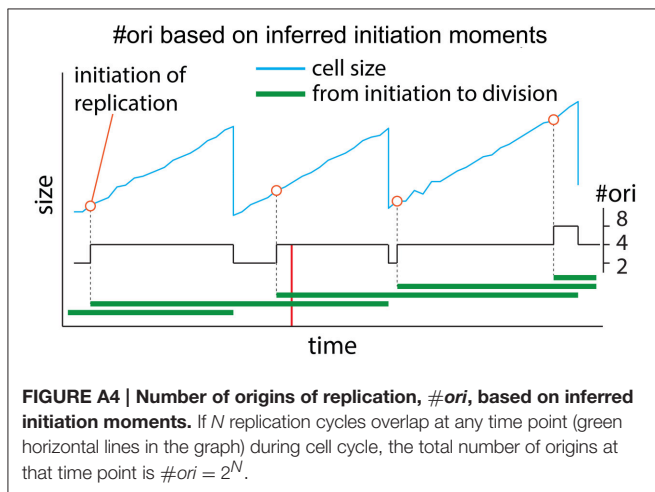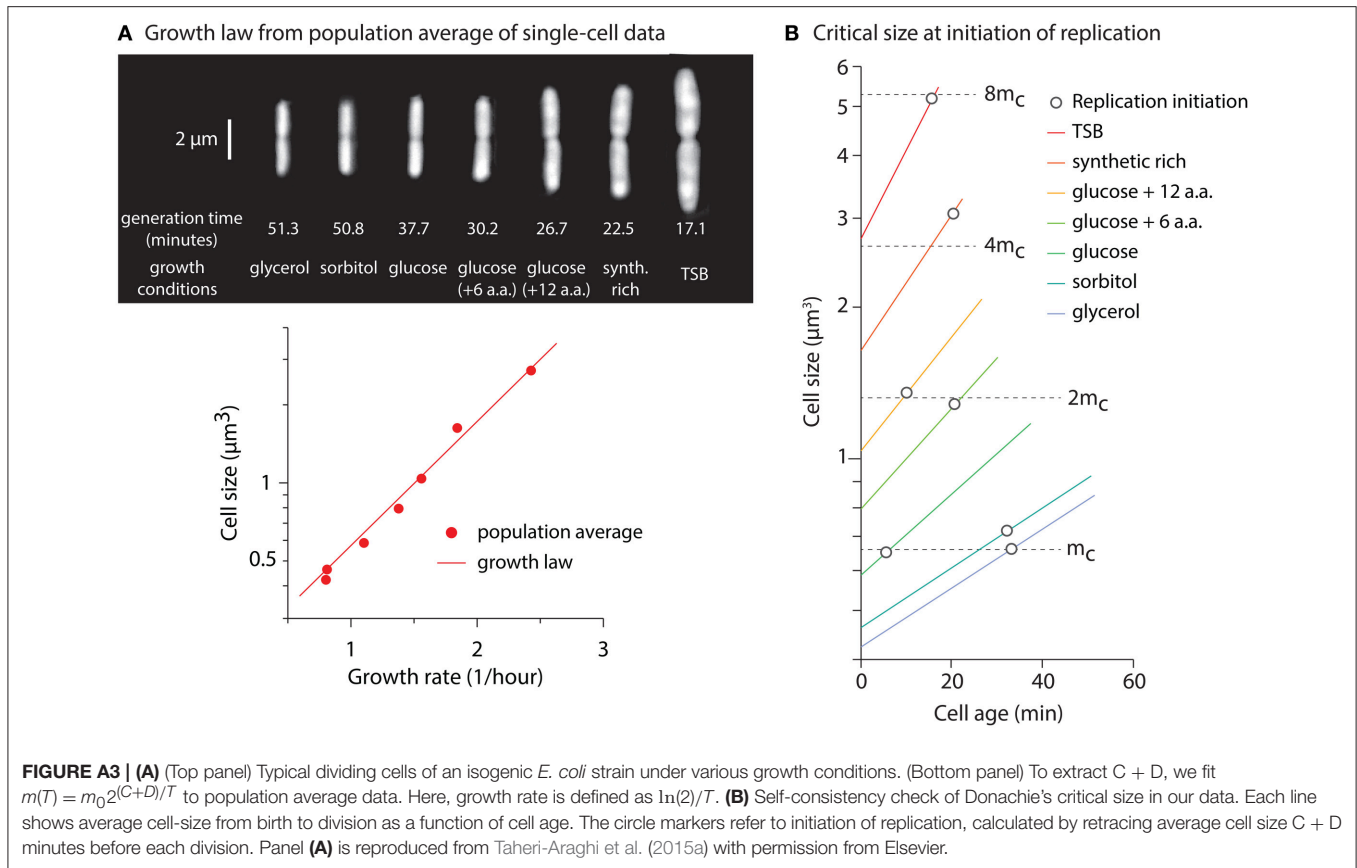area refer to cells with double initiations in a division cycle. Standard deviation, σ_n, in each row is the same as the one in the corresponding row in **(B)**. The ratio of the standard deviations, σ_n, to the average doubling time, T, is noted on the left side of each sub panel. **(D)** Mother-daughter correlations of $\Delta s/\#ori$ with noise added in retracing.

There are two different points that should not be mixed:

 (i) The actual fluctuations in C and D periods.
(ii) The noise that can be possibly added to retracing time, C + D.

The reading error, defined as the difference between the actual moment of initiation and the inferred moment, is a result of the combination of (i) and (ii). Regarding the point (i), we believe that there are fluctuations in C and D periods. However, to date, we do not have any experimental data measuring and addressing these fluctuations at a single-cell level. Regarding the point (ii), retracing is an indirect method of

estimating initiation time and size. Adding noise to retracing time can test the extent the outcome is robust with respect to noise.

Let's consider the actual fluctuations of C + D periods (combined) have the standard deviation of σ_a. In the case that we do not include noise in retracing time, the inferred readings miss these actual fluctuations. Thus, the readings error has the standard deviation of σ_a as well. Now consider that we add a noise in retracing time such that initiation times are estimated C + D + δ_n prior to cell divisions. Here δ_n is a stochastic variable. Let's assume the standard deviation of (C + D + δ_n) is σ_n. The noise δ_n is independent of fluctuations in C and D periods. Thus, this

A Growth law from population average of single-cell data

B Critical size at initiation of replication

FIGURE A3 | **(A)** (Top panel) Typical dividing cells of an isogenic *E. coli* strain under various growth conditions. (Bottom panel) To extract C + D, we fit $m(T) = m_0 2^{(C+D)/T}$ to population average data. Here, growth rate is defined as $\ln(2)/T$. **(B)** Self-consistency check of Donachie's critical size in our data. Each line shows average cell-size from birth to division as a function of cell age. The circle markers refer to initiation of replication, calculated by retracing average cell size C + D minutes before each division. Panel **(A)** is reproduced from Taheri-Araghi et al. (2015a) with permission from Elsevier.

FIGURE A4 | **Number of origins of replication, #ori, based on inferred initiation moments.** If N replication cycles overlap at any time point (green horizontal lines in the graph) during cell cycle, the total number of origins at that time point is $\#ori = 2^N$.

type of noise does not capture the actual fluctuations, but only adds more uncertainty to the reading of inferred initiation size and time.

If both actual fluctuations of C + D and added noise are Gaussian, we can calculate the standard deviation of the reading error. If the actual initiation points happen $(C + D + \delta_a)$ prior to cell divisions, the retracing analysis reads those initiations at

$(C + D + \delta_n)$. (δ_a and δ_n are independent stochastic variables). Thus, the reading error is given by:

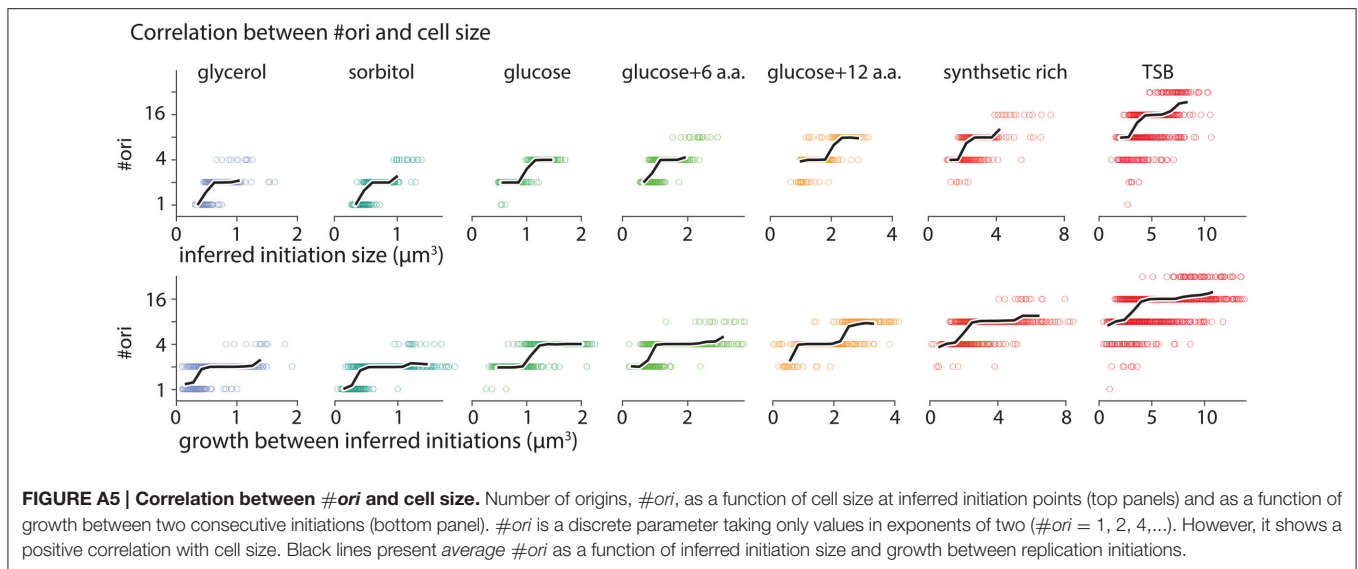$$error = (C + D + \delta_n) - (C + D + \delta_a). \quad (A11)$$

Since δ_n and δ_a are independent Gaussian variables, the standard deviation of the error is given by

$$\sigma_{error}^2 = \sigma_n^2 + \sigma_a^2. \quad (A12)$$

Therefore, to minimize the error in retracing analysis, we should minimize the noise in $C + D + \delta_n$ in the analysis. For this reason using a constant C + D is a better choice for this study.

To visualize the effect of noise in the analysis, we added Gaussian noise in retracing time and tracked the changes in the distributions of the inferred initiation sizes. As expected, the noise widens the distributions and beyond certain points, it influences the bimodal shape of the distributions. Various levels of noise is tested with standard deviations, σ_n, chosen at 0%, 2.5%, 5%, 7.5%, and 10% of C + D. **Figure A2B**, shows distributions of the initiation size for various level of noise, as noted on the left side of each panel.

Figure A2C depicts the distributions of inferred initiation size for individual growth conditions. The solid lines in each panel refer to the distributions from whole cell population and the filled area refer to cells with double initiations in

FIGURE A5 | Correlation between _#ori_ and cell size. Number of origins, _#ori_, as a function of cell size at inferred initiation points (top panels) and as a function of growth between two consecutive initiations (bottom panel). _#ori_ is a discrete parameter taking only values in exponents of two (_#ori_ = 1, 2, 4,...). However, it shows a positive correlation with cell size. Black lines present _average #ori_ as a function of inferred initiation size and growth between replication initiations.

one division cycle. The level of noise, σ_n, in each row of **Figure A2C** is the same as the one in the corresponding row in **Figure A2B**. The numbers on top of each distribution show the ratio of the standard deviation, σ_n, of the noise to the average doubling time in each growth condition, σ_n/T. For noises with $\sigma_n/T \gtrsim 20\%$, the widening of the distributions are such that peaks of the bimodal distributions start to overlap. However, the prediction of double initiations are robust with respect to the shape of the distributions (see filled area in panels of **Figure A2C**).

In **Figure A2D**, we show the effect of added noise in the mother-daughter correlations of $\Delta s/\#ori$. We do not see significant correlations between mother and daughter cells. However, at high level of added noise, the correlations tends to slightly tilt toward negative values since a random noise affects the value of $\Delta s/\#ori$ oppositely in two consecutive generations.

In conclusion, for the retracing analysis used in this work, a constant value for retracing time C + D minimizes the analysis error. Our test on the distributions of the inferred initiation cell size shows that the shape of distributions are more or less conserved if the noise in the retracing time is up to ∼ 20% of the generation time. Beyond this threshold the shape of the distributions start to change.

APPENDIX C

Consistency Check between Population Average and Single-Cell Analysis

In this appendix, we reproduce the Donachie's analysis on the _average_ cell-size data from our single-cell measurements. To this end, we use C + D extracted from fitting average data to Equation (1) (**Figure A3A**), where C + D = 69 min. As expected, applying retracing analysis reproduces Donachie's critical size model (**Figure A3B**). This analysis is solely for a secondary self-consistency check of critical size model on our single-cell data

(i.e., given the method used here, we do not expect any result otherwise).

Figure A3B shows cell size from birth to division as a function of cell age for seven different growth conditions. The circle markers refer to the moment of the replication initiation. Since C + D is longer than the generation times for the growth conditions in the study, each initiation size essentially refers to cell size in previous generations. The initiation size of different growth conditions can be clustered such that the initiation size per replication origin is constant and independent of growth condition. For the slowest condition, $m_c = 0.66\,\mu m^3$ and for the rest, critical initiation size increases by factors of two [1].

APPENDIX D

Calculating Number of Origin of Replication Based on Initiation Times

Based on inferred moments of initiation, one can mathematically calculate the number of origins during the cell cycle. A graphical example is presented in **Figure A4**. The graph shows cell size and the inferred moments of initiation as a function of time. The horizontal green lines show replication cycles from initiation of replications until the corresponding cell divisions. In slowest growth conditions, where we have either zero or one replication cycle going on in the cell ($N = 0$ or 1), we know that $\#ori = 1$ or 2 (not shown in the figure). In fast growth conditions, N (>1) replication cycles can overlap. At any initiation of replication, number of overlapping cell cycles, N, increases by one and number of origins of replication, $\#ori$, increases by a factor of two. At any cell division, number of overlapping cell cycles, N, decreases by one and $\#ori$ decreases by a factor of two. Thus, since

---

[1]As a consistency check we can also extract $m_\circ$ from **Figure A3A** using Equation (1) and compare it with m_c through Equation (A10). We find $m_\circ = 0.168\mu m^3$. In the slowest condition in the study, $N = 2$ cell cycles overlap. $m_c = 0.66\mu m^3 \sim 2^2 m_\circ$. This agrees with Equation (A10) that states $m_c = 2^N m_\circ$.

for $N = 0$, #$ori = 1$ and for $N = 1$, #$ori = 2$ and since increasing N by one increases #ori by a factor of two, we conclude if N replication cycles overlap at any moment, #ori at that moment is given by #$ori = 2^N$. For instance, at the moment shown with the red vertical line in **Figure A4**, $N = 2$ replication cycles overlap (two horizontal green lines cross the red line). Thus, the total number of #ori at that moment is $2^N = 4$. This can be applied to any moment throughout the cell cycle to calculate #ori as a function of time.

To examine correlation between #ori and cell size, we plotted #ori calculated at inferred initiation points as a function of initiation cell size (**Figure A5** top panels) and as a function of growth between initiations (**Figure A5** bottom panels). #ori is a discrete quantity taking only values in exponents of two (#$ori =1$, 2, 4,...). There is, however, a positive overall correlation between computed #ori and cell size as depicted by black lines in each panel of **Figure A5** (black lines show *average* #ori as a function of corresponding size or growth). The positive correlation arises from the fact that size and generation time are negatively correlated and that shorter generation time results in more overlapping replication cycles, thus larger #ori. The parameter $\Delta s/$#ori that is used in **Figure 1F** is calculated based on growth between inferred initiation points per #ori during that growth period.