Theoretical models for the regulation of DNA replication in fast-growing bacteria

Martin Creutziger¹, Mischa Schmidt¹ and Peter Lenz¹,²,³

¹ Fachbereich Physik, Philipps-Universität Marburg, D-35032 Marburg, Germany
² Zentrum für Synthetische Mikrobiologie, Philipps-Universität Marburg, D-35032 Marburg, Germany
E-mail: peter.lenz@Physik.Uni-Marburg.DE

New Journal of Physics 14 (2012) 095016 (23pp)
Received 9 May 2012
Published 18 September 2012
Online at http://www.njp.org/
doi:10.1088/1367-2630/14/9/095016

Abstract. Growing in always changing environments, Escherichia coli cells are challenged by the task to coordinate growth and division. In particular, adaption of their growth program to the surrounding medium has to guarantee that the daughter cells obtain fully replicated chromosomes. Replication is therefore to be initiated at the right time, which is particularly challenging in media that support fast growth. Here, the mother cell initiates replication not only for the daughter but also for the granddaughter cells. This is possible only if replication occurs from several replication forks that all need to be correctly initiated. Despite considerable efforts during the last 40 years, regulation of this process is still unknown. Part of the difficulty arises from the fact that many details of the relevant molecular processes are not known. Here, we develop a novel theoretical strategy for dealing with this general problem: instead of analyzing a single model, we introduce a wide variety of 128 different models that make different assumptions about the unknown processes. By comparing the predictions of these models we are able to identify the key quantities that allow the experimental discrimination of the different models. Analysis of these quantities yields that out of the 128 models 94 are not consistent with available experimental data. From the remaining 34 models we are able to conclude that

³ Author to whom any correspondence should be addressed.

Content from this work may be used under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
mass growth and DNA replication need either to be truly coupled, by coupling DNA replication initiation to the event of cell division, or to the amount of accumulated mass. Finally, we make suggestions for experiments to further reduce the number of possible regulation scenarios.

Online supplementary data available from stacks.iop.org/NJP/14/095016/mmedia

Contents

1. Introduction
2. Results
   2.1. Coupling schemes
   2.2. Growth models with noise
   2.3. Simulation results with DNA
3. Discussion
Acknowledgments
Appendix. Materials and methods
References

1. Introduction

*Escherichia coli* cells have astonishing abilities to grow rapidly in a variety of different environmental conditions. In a rich medium (where all amino acids are provided) *E. coli* can duplicate itself every 20 min. At first sight this is quite astonishing since replication of the whole chromosome takes 40 min and cell division occurs only another 20 min after termination of replication [1]. To circumvent these difficulties *E. coli* uses multiple replication forks by initiating replication at all available oriC sites before the preceding round of replication has finished [2].

These observations give rise to the fundamental question of how bacterial cells are able to adjust the rate of DNA replication to the growth rate supported by the growth medium. The regulatory scheme that guarantees maintenance of well-balanced growth under different conditions was revealed by the seminal work of Cooper and Helmstetter 40 years ago [2, 3]. According to their findings, the time at which DNA replication is initiated at all available oriC sites depends only on the doubling time (set by the surrounding growth medium). Cell division of the mother, daughter or grand-daughter cells occurs 60 min after the initiation of DNA replication at the corresponding oriC sites. In particular, this implies that if the doubling time is less than 40 min, replication occurs from several forks. This scheme predicts a constant cell mass per oriC site at the event of initiation (the so-called initiation mass), as observed and described in [4].

However, until now it has not been understood what molecular regulation gives rise to this rather simple macroscopic relation between cell mass and initiation of DNA replication. For fast-growing bacteria (i.e. with a doubling time shorter than 60 min), the DNA synthesis rate of DNA polymerase is essentially constant [1], suggesting that regulation of DNA replication has to take place at the level of initiation [5]. This regulation is not necessarily coupled to cell mass (or any other quantity proportional to cell mass). The observed constant initiation mass [4] could simply be the byproduct of a mass-independent regulatory mechanism, which itself regulates
mass growth and replication initiation [6]. Moreover, the observation of a constant initiation mass has been challenged [7] and discussed controversially [8].

An additional complication arises from the fact that cells are subject to a variety of different sources of molecular noise. Even under constant growth conditions [9–11], the main events of the cell cycle are affected by the presence of noise leading to stochastic division site placement [12] (with a coefficient of variation (CV) of 3–5% [12–14]), stochastic timing of the division event [15] (with a CV of 20–33% [16, 17]), yielding thus stochastic distributions of division mass [16] (a CV of about 8–12% [13, 18]), and individual growth rates [19, 20] (a CV of about 13% [20, 21]). However, only the presence of noise makes one regulatory strategy favorable over the others. In fact, if all cell-division-related processes were not subject to any kind of noise, all cells would always need the exact same time to reach the division mass, start one round of DNA replication once per cell cycle at the same time, so that they divide after the interdivision time set by the growth medium. In this case, different regulatory schemes give rise to the same global behavior, leading to no differences in observable quantities. In analyzing the advantages of regulatory schemes, we therefore have to take into account that cells are subject to noise.

This makes difficult the theoretical discussion of different possible coupling schemes of the DNA replication triggering mechanism to cellular growth and division. Although many molecular details of the replication initiation process have been investigated and are well established (for a review, see [22]), the understanding of cell cycle regulation is still incomplete [6] even though this question has been theoretically discussed for at least 40 years [23].

In the literature, several different models for the regulation of replication-initiation have been discussed. Typically, in these models initiation is coupled either to (i) cell mass [4, 24, 25] or (ii) to a quantity proportional to DNA content (such as DnaA) [26–28]. In other models, cells are able to keep track of time, which implies that some type of internal clock needs to be reset [23]. Like other cell cycle events, DNA replication is observed to be triggered only once in the cell cycle [6], possibly, however, at several replication forks simultaneously. Thus, the resetting of the internal clocks needs to occur once per cell cycle. Two obvious candidates for checkpoints at which the internal clock is reset are cell division or initiation of replication itself. Correspondingly, in our theoretical models the resetting event is assumed to occur at (iii) the time of birth or (iv) the initiation time of the previous round of replication. Variants of these time-induced models are discussed as two independently controlled cycles of mass growth and DNA replication, where an internal set of parameters plays the role of the cells clock which is reset at certain checkpoints during the cell cycle [6, 29]. These four simple schemes are the basis of most of the models discussed in the literature; see, e.g., [4, 19, 24, 26, 29–32].

In the above models, one also has to take into account molecular noise that smears out the DNA synthesis patterns [3]. However, this noise does not originate from the DNA replication process itself, since the DNA synthesis rate of DNA polymerase is essentially constant [1] and initiation at multiple origins is highly synchronized [33]. These findings suggest that the observed noise must originate from statistical fluctuations affecting the initiation process.

Finally, the theoretical models have to take into account that the cell cycle consists of two independent cycles (the cycle of mass duplication and the cycle of chromosome replication [6]) that both have to be finished before cell division takes place [5]. The actually achieved time between birth and subsequent division (interdivision time) of a single cell is therefore typically limited by either the time needed until two completely replicated DNA strands have segregated.
or the time needed to reach division mass. To coordinate these two tasks the regulatory system needs to sense both their environment and their achieved size or mass [24].

This coupling between cellular growth and DNA replication leads to a sensitive dependence of certain population observables (such as age distributions, DNA content or mean mass per cell) on molecular noise. This noise dependence provides essential information about the regulatory system and (as we demonstrate here) makes it possible to distinguish the different coupling schemes mentioned above.

In this study, we theoretically analyze the complex molecular regulation of DNA replication in *E. coli*. In doing so, we are challenged by the fact that many molecular details of the involved cellular processes (such as cell division, regulation of the cell cycle, etc) are unknown. To circumvent this difficulty, we develop here a novel approach. Similar to the idea of ensemble modeling, as introduced for network analysis of cellular pathways by the authors of [34], we introduce a series of different models that differ in basic assumptions about the unknown processes. In particular, we introduce various coupling schemes for the DNA replication initiation to cellular growth variables such as mass or DNA content. In doing so, we do not take into account the molecular details that give rise to the implemented coupling scheme; rather we describe these mechanisms at a coarse-grained level. Comparison with available experimental data then allows us to significantly reduce the number of models.

To be more specific, in the following we introduce a set of 128 different models (differing in at least one regulatory mechanism or in one source of molecular noise). The free parameters of these models are then fixed by comparison with available experimental data. Next, the models are classified whether they are in agreement with the remaining experimental data. Unfortunately, this procedure does not allow us to determine the precise regulatory mechanism since several models can be brought into agreement with all available experimental data. However, we conclude by making several suggestions for experiments that could further reduce the compatible models.

2. Results

2.1. Coupling schemes

To analyze the coupling between mass growth and DNA replication, we have developed a series of theoretical models. All models have in common that the growth of a bacterial population is simulated by a sequence of cell division events. We start from a single newborn cell and simulate DNA replication, mass growth and division of this cell and all its daughter cells; for details see the appendix. The individual doubling time $T_m$ (the time a cell needs to double its mass) is set by the growth medium [1]. In the following, we focus on fast-growing *E. coli* cells (with doubling times faster than 60 min).

Our models differ in the regulation of initiation of DNA replication. More specifically, we implemented the coupling schemes introduced above and analyzed four different models where replication is initiated:

- at a certain cell mass per ori (mass-induced (MI) model);
- at a certain time after birth (birth time-induced (TI) model);
- at a certain time after the preceding replication fork was started (replication time-induced (RTI) model);
- when the DNA amount reaches a critical threshold (DNA-induced (DI) model).
Figure 1. Schematic representation of coupling between mass growth and DNA replication. In our models the key quantities for cellular growth are birth mass, division mass and growth rate (slope of the line). In the example given the interdivision time is around 40 min, while the DNA replication time is 60 min. For these values the mother cell has to initiate a second replication fork in the middle of the cell cycle to guarantee that the daughter cells are able to keep DNA synthesis in pace with mass growth. The time point of initiation of the second replication fork can be characterized by the specific birth time, by a critical mass or DNA content reached at that time or by the time that has passed since initiation of the previous replication fork.

New Journal of Physics 14 (2012) 095016 (http://www.njp.org/)
sets in were calculated in such a way that in the absence of noise the simulations reproduce the observed replication fork pattern [2, 3]. For details, see appendix A.1.

2.2. Growth models with noise

As mentioned, growing bacteria are subject to various kinds of noise. In our models, we take the following sources of noise into account:

- noise on the accuracy of division site placement (that gives rise to a distribution of birth masses) (NB).
- noise on the timing of division (that gives rise to a distribution of division masses) (ND). We analyzed two variants of ND noise: (i) Po where the division noise has an exponential distribution \( p(x) = \lambda_D e^{-\lambda_D x} \) (with noise parameter \( \lambda_D \)), and (ii) Ga where the division noise is characterized by a Gaussian distribution (with standard deviation \( \sigma_D \)) of division masses.
- noise on the growth rate (slope of the red line in figure 1) (NGr) in two different variants: (i) Tr where growth rate changes occur randomly. Timing of the switching event is drawn from an exponential distribution (with parameter \( \lambda_{time} \)). (ii) Tf where each cell chooses a random growth rate at a fixed cell age \( a_{time} \) measured as a fraction of the cell’s total lifetime (relative cell age). In both models the randomly chosen growth rate is drawn from a normal distribution with standard deviation \( \sigma_{Gr} \).

First we analyzed the influence of the different sources of noise on the growth of a bacterial population. More specifically, we analyzed the dependences of the different sources of noise and
fixed the values of the free parameters by comparison with available experimental data on cell growth (for details, see appendix A.2). Our findings are as follows:

(i) **NB** noise is independent of **ND** and **NGr** noise. It is characterized by Gaussian distribution of mass partitioning at division with standard deviation $\sigma_B = 3$–5%.

(ii) Growth of the population only weakly depends on $\lambda_{\text{time}}$ and $a_{\text{time}}$ (see figures 2 and 3). To analyze the dependence of other observables on these parameters it is sufficient to only consider two different parameter values in each model, i.e. there are two model variants of $\textbf{Tf}$ with high (H) or low (L) $\lambda_{\text{time}}$ and two model variants of $\textbf{Tr}$ where switching occurs at birth (B, $a_{\text{time}} = 0$) or in the middle of the cell cycle (M, $a_{\text{time}} = 0.5$).

(iii) **ND** and **NGr** are independent as can be seen from systematic scans of the parameter values of $\sigma_D$, $\lambda_D$ and $\sigma_{\text{Gr}}$. Observed mass distributions depend only on the **ND** noise parameter and are independent of **NGr** noise, which in turn influences the observed growth rate distribution only. Figure 4 shows one example of the eight models; the remaining seven models are shown in the online supplementary data available from stacks.iop.org/NJP/14/095016/mmedia.

(iv) The remaining free parameters $\sigma_D$, $\lambda_D$ and $\sigma_{\text{Gr}}$ were adjusted to reproduce the experimentally observed CVs of division mass and growth rate.

With all parameters fixed we then tested which of the models correctly predict the remaining experimentally observed quantities (the CV of the distribution of individual interdivision times, $T_i$). Furthermore, we compared the predictions of the models with an

**Figure 2.** Influence of the noise parameter $\lambda_{\text{time}}$ on population observables for the $\textbf{Tr}$ models. CV = coefficient of variation, sisi = sister–sister correlation, damo = daughter–mother correlation.
Figure 3. Influence of $a_{\text{time}}$ on population observables for the TF models. Here, $a_{\text{time}}$ represents the relative cell age at which new growth rates are randomly chosen. CV = coefficient of variation, sisi = sister–sister correlation, damo = daughter–mother correlation.

An additional set of available experimental data on the correlation of interdivision times between sister cells (CCsisi) with a reported positive correlation coefficient larger than 0.5 [17, 35] and between mother and daughter cells (CCdamo) with reported negative correlation coefficient in the range −0.5 to 0 [16, 17]. All models produce positive sister–sister correlations and slightly negative mother–daughter correlations. Furthermore, they show a CV of $T_i$ distributions between 20% and 33% in good agreement with experimental observations [16, 17]. For details see table 2.

These parameter scans also demonstrate that all models robustly reproduce the expected positive CCsisi [17] and zero or negative CCdamo [16, 17]. The models are also robust with respect to variations in $\sigma_B$; see figure 5.

We also analyzed whether it is sufficient to introduce only one effective type of noise that is able to reproduce the main experimental observations mentioned above. To do so, we introduced a series of (hypothetical) mutants that all had only one source of noise. However, as explained in detail in section 2 of the online supplementary data available from stacks.iop.org/NJP/14/095016/mmedia, each of the noise types has its own characteristic influence on at least part of the experimental observations. It is thus inevitable to implement cell growth with all three types of noise.

2.3. Simulation results with DNA

In a next step, we included DNA replication in our simulations. As described above, in all four coupling schemes triggering of DNA replication is described by a critical variable that reaches
Figure 4. Parameter space for the GaTrH model. Different values of the varied quantities are shown in different colors (for example, in the left plot in the first line the CV of birth mass takes the values 14% (green), 12% (blue), 10% (magenta) and 8% (cyan)). All quantities are measured in units of the growth rate $\mu = \log 2 / T_i$. As can be seen, birth mass (BM), division mass (DM) and growth rate (Gr) are independently controlled by the respective noise parameter, whereas the CV of $T_i$ distribution and the observed correlation coefficients between sisters (CCsisi) and between daughters and mothers (CCdamo) depend on both parameters. Plots for the other seven models can be found in supplementary figures S1–S5 (online supplementary data available from stacks.iop.org/NJP/14/095016/mmedia).

a threshold value. The free parameters of the four coupling schemes were chosen in the way mentioned in appendix A.2, i.e. we aimed at finding a set of parameters that reproduces the experimental observations of growth rate noise and division mass noise simultaneously, while we fixed the noise parameter of NB noise at $\sigma_B = 3.3\%$ as above.

However, although DNA replication runs independently of mass growth both processes together determine the interdivision time. More specifically, cell division takes place only when both processes are finished, implying

$$T_i = \max(T_{DNA}, T_{mass}).$$

Here $T_{mass}$ denotes the time after birth when the cell reaches division mass and $T_{DNA}$ the time when DNA replication is completed. Note that therefore $T_{DNA}$ is not the time needed for replication of a whole DNA strand; instead, if replication starts already in the mother cell, $T_{DNA}$
Table 2. Overview of models with noise on cell growth. Data shown are for $T_i = 60$ min and $\sigma_B = 3.3\%$. The last two lines summarize the experimental data used for comparison.

| Model    | $\lambda_{time}$ | $\lambda_D$ | $1/\sigma_D$ | $\sigma_{Gr}$ (%) | CV DM (%) | CV $T_i$ (%) | CCsisi | CCdamo |
|----------|------------------|-------------|--------------|-------------------|-----------|-------------|--------|--------|
| PoTfB    | 0                | 0.05        |              | 13.2              | 9.9       | 26.6        | 0.25   | -0.25  |
| PoTfM    | 0.5              | 0.05        | 13.2         | 9.9               | 25.3      | 0.44        | -0.17  |
| PoTrL    | 0.005            | 0.05        | 12.9         | 9.9               | 25.8      | 0.40        | -0.20  |
| PoTrH    | 0.03             | 0.05        | 13.4         | 9.9               | 26.7      | 0.49        | -0.05  |
| GaTfB    | 0                | 0.08        | 13.2         | 10.1              | 26.4      | 0.31        | -0.34  |
| GaTfM    | 0.5              | 0.08        | 13.2         | 10.2              | 25.3      | 0.51        | -0.27  |
| GaTrL    | 0.005            | 0.08        | 13.2         | 10.2              | 26.5      | 0.54        | -0.14  |
| GaTrH    | 0.03             | 0.08        | 13.0         | 10.2              | 25.8      | 0.46        | -0.29  |
| Exp. values | 13              | 8–12        | 20–33        | ≈0.5              | -0.5 – 0  |

References [20, 21] [13, 18] [16, 17] [17, 35] [16, 17]

Figure 5. Influence of $\sigma_B$ on the observed CV of the $T_i$ distribution. All other parameters are kept fixed ($1/\sigma_D = 0.12$, $\lambda_D = 0.07$, $\sigma_{Gr} = 0.10$).

can be shorter than the sum of C- and D-time (where the C-time is the time required to replicate the chromosome and the D-time is the obligatory period between termination of replication and cell division). Again, no molecular details of this coupling are known and in our simulations we have to implement different variants:

- If DNA replication is the reason for delayed division (i.e. $T_{DNA} > T_{mass}$), then (i) the growth rate can be adjusted such that both cellular growth and DNA replication finish
simultaneously (DNA influences the growth rate \((D_g)\)), or (ii) the cells could continue growing after having reached division mass, until DNA replication allows for division (DNA influences the division mass \((D_m)\)).

- Because of the coupling, division mass noise now can originate from either noise concerning the length of the D-period of DNA replication \((N_D D)\) or as before only from mass growth \((N_D M)\).

Together with the four DNA replication initiation schemes and the eight variants of implementing mass growth noise, this leads to a total of 128 different model variants.

In analogy to the procedure described above for the case without DNA, we first performed two-parameter scans of the \(N_d\) and \(N_G r\) noise parameters for all of the 128 models. As it turned out, the \(N_d\) and \(N_G r\) parameters cannot be fixed independently. Instead, the observed mass distributions and growth rate distributions generally depend both on \(\sigma_D\), respectively \(\lambda_D\), and on \(\sigma_f\). Moreover, some of the model variants cannot be brought in agreement with the experimentally measured growth rate CV or the division mass CV. In some cases (denoted by CVlow in the sequel) the CV of the division mass distribution is below 4% for all parameter values. In these cases, we fixed parameters so as to best reproduce the expected noise on growth rate as well as on the \(T_i\) distribution. The latter was used instead of the division mass distribution only in this scenario. Furthermore, some models (denoted by CVhigh in the following) did not reach a steady-state division mass distribution (within 200 generations) for any set of reasonable values for \(N_d\) and \(N_G r\) noise parameters, see section 6 of the online supplementary data available from stacks.iop.org/NJP/14/095016/mmedia. We therefore could not assign these an equilibrated value for the CV of the mass distributions. No difficulties occurred in fixing the \(N_B\) noise to \(\sigma_B = 3.3\%\) as above.

As a next test for the models, we calculated the mean population content of DNA as a function of the observed mean interdivision time \(\bar{T}_i\). To do so we varied the doubling time set by the medium in the range \(21 \text{ min} \leq T_m \leq 80 \text{ min}\) and compared the numerically calculated mean population DNA content with experimental data [1] and the theoretical prediction by Cooper [2]. See figure 6 for examples. The results for all models are summarized in table 3.

Furthermore, we measured in the simulations (at a fixed \(T_m = 40 \text{ min}\)) the CV of the distribution, \(C_{C_sisi}\) and \(C_{C_damo}\) and (as above) compared their values with experimental data. See figure 7 for examples. The results for all models are also shown in table 3.

With these numerical results we were then able to classify the models. More specifically, we classified a model as being consistent with all experimental observations when the following constraints were fulfilled:

\[
5\% \leq CVT_i \leq 40\%,
\]
\[
0.4 \leq CC_{sisi},
\]
\[
CC_{damo} \leq 0.1.
\]

If one of these criteria was not met, or DNA content was measured too high, or the division mass distribution never reached steady state (CVhigh), models were marked red in table 3 and not taken into further consideration. In some cases we could not definitively exclude a model, namely in the case the observed DNA content was only a little too high or in the case of a too sharply regulated division mass accuracy (CVlow). In these cases, the models were marked yellow but not excluded.
Table 3. Results for simulations of all 128 model variants. Model names are compositions of abbreviations of their defining parameters, as explained in the text and summarized in table 1. Listed values include the coefficients of variation (CV) of the distributions of growth rates (gr), division mass (DM) and single-cell interdivision times \( T_i \), the coefficient of correlation of the interdivision times of sisters (CCsisi) and the mean DNA content per cell of the population as a function of the mean interdivision time \( \bar{T}_i \) (DNA). The latter quantity was obtained by varying the doubling time \( 21 \leq T_m \leq 80 \text{ min} \) set by the medium. If the calculated values agree with the experimental observations, then they are reported as OK and otherwise as fails. The last column shows whether the numerical results succeed (marked green, 23 variants) or fail in reproducing all experimental observations (marked red, 94 variants). Intermediate cases were marked yellow (11 variants); see text for details.

| Model         | CV gr (%) | CV DM (%) | CV \( T_i \) (%) | CCsisi | CCdamo | DNA | Fails due / OK |
|---------------|-----------|-----------|-----------------|--------|--------|-----|----------------|
| GaTfB_MiDgNdD | 11.9      | 3.5       | 11.8            | 0.21   | -0.36  | OK  | CCsisi, CVlow  |
| GaTfB_MiDgNdM | 12.5      | 9.8       | 17.4            | 0.50   | -0.29  | OK  |                 |
| GaTfB_MiDmNdD | 13.3      | 10.4      | 22.9            | 0.50   | -0.42  | OK  |                 |
| GaTfB_MiDmNdM | 13.3      | 10.0      | 21.3            | 0.56   | -0.45  | OK  |                 |
| GaTfB_TiDgNdD | 12.8      | 3.5       | 12.6            | 0.10   | -0.31  | OK  | CCsisi, CVlow  |
| GaTfB_TiDgNdM | 12.4      | 9.8       | 16.6            | 0.44   | -0.49  | OK  |                 |
| GaTfB_TiDmNdD | 13.3      | n/a       | 39.2            | 0.67   | -0.85  | OK  | CVhigh         |
| GaTfB_TiDmNdM | 13.3      | n/a       | 35.4            | 0.85   | -0.93  | OK  | CVhigh         |
| GaTfB_RtiDgNdD| 13.2      | 3.5       | 13.3            | -0.01  | -0.03  | Fails| DNA, CCsisi, CVlow|
| GaTfB_RtiDgNdM| 13.0      | 9.8       | 20.1            | 0.28   | -0.28  | Fails| DNA, CCsisi    |
| GaTfB_RtiDmNdD| 13.3      | 6.0       | 20.4            | 0.39   | -0.40  | Fails| DNA, CCsisi    |
| GaTfB_RtiDmNdM| 13.2      | 15.0      | 17.6            | 0.45   | -0.39  | Fails| DNA            |
| GaTfB_DiDgNdD | 12.9      | 3.5       | 13.2            | -0.05  | 0.00   | Fails| DNA, CCsisi, CVlow|
| GaTfB_DiDgNdM | 13.1      | 9.8       | 21.8            | 0.27   | -0.27  | Fails| DNA, CCsisi    |
| GaTfB_DiDmNdD | 13.2      | 0.4       | 14.4            | 0.06   | -0.14  | Fails| DNA, CCsisi, CVlow|
| GaTfB_DiDmNdM | 13.2      | 10.0      | 22.1            | 0.28   | -0.30  | Fails| DNA, CCsisi    |
| GaTfM_MiDgNdD | 12.7      | 3.5       | 11.2            | 0.22   | 0.32   | OK  | CCsisi, CVlow  |
| GaTfM_MiDgNdM | 12.9      | 9.8       | 18.4            | 0.61   | -0.16  | (OK) | DNA            |
| GaTfM_MiDmNdD | 13.2      | 10.5      | 22.0            | 0.47   | -0.32  | OK  |                 |
| GaTfM_MiDmNdM | 13.2      | 9.5       | 18.1            | 0.56   | -0.20  | OK  |                 |
| GaTfM_TiDgNdD | 12.9      | 3.5       | 11.5            | 0.30   | 0.19   | OK  | CCsisi, CVlow  |
| GaTfM_TiDgNdM | 12.9      | 9.8       | 17.8            | 0.63   | -0.38  | OK  |                 |
| GaTfM_TiDmNdD | 12.6      | n/a       | 40.7            | 0.68   | -0.83  | OK  | CVhigh         |
| GaTfM_TiDmNdM | 13.0      | n/a       | 30.3            | 0.79   | -0.57  | OK  | CVhigh         |
| GaTfM_RtiDgNdD| 13.2      | 3.5       | 12.1            | 0.51   | 0.40   | Fails| DNA, CCdamo, CVlow|
| GaTfM_RtiDgNdM| 13.0      | 9.8       | 20.2            | 0.52   | -0.20  | Fails| DNA            |
| GaTfM_RtiDmNdD| 13.2      | 5.6       | 19.6            | 0.48   | -0.34  | (Fails)| DNA        |
| GaTfM_RtiDmNdM| 13.2      | 16.0      | 9.3             | 0.57   | -0.33  | (Fails)| DNA        |
| GaTfM_DiDgNdD | 12.8      | 3.5       | 11.8            | 0.53   | 0.45   | Fails| DNA, CCdamo, CVlow|
| GaTfM_DiDgNdM | 13.3      | 9.8       | 21.2            | 0.50   | -0.17  | Fails| DNA            |
| GaTfM_DiDmNdD | 13.2      | 1.6       | 12.3            | 0.54   | 0.35   | Fails| DNA, CCdamo, CVlow|
| GaTfM_DiDmNdM | 13.2      | 10.1      | 21.6            | 0.50   | -0.20  | Fails| DNA            |
| GaTrL_MiDgNdD | 10.9      | 3.5       | 11.4            | 0.43   | 0.51   | OK  | CCdamo, CVlow  |
| GaTrL_MiDgNdM | 10.5      | 9.8       | 21.7            | 0.61   | -0.06  | OK  |                 |
### Table 3. Continued.

| Model               | CV gr (%) | CV DM (%) | CV Tisi (%) | CCSI | CCdamo | DNA   | Fails due / OK |
|---------------------|-----------|-----------|-------------|------|--------|-------|----------------|
| GaTrH_MidmDnD      | 13.0      | 9.1       | 18.8        | 0.55 | −0.03  | DNA   |                |
| GaTrH_TiDnD       | 12.8      | 11.4      | 0.53        | 0.22 | OK     | CCdamo, CVlow |                |
| GaTrH_TiDnD       | 12.9      | 9.8       | 21.8        | 0.63 | −0.07  | DNA   |                |
| GaTrH_TiDnD       | 12.8      | n/a       | 33.9        | 0.69 | −0.86  | CVhigh |                |
| GaTrH_TiDnD       | 12.6      | n/a       | 26.4        | 0.97 | −0.96  | CVhigh |                |
| GaTrH_RtiDnD      | 13.0      | 3.5       | 13.3        | 0.59 | 0.61   | DNA   | DNA, CCdamo, CVlow |
| GaTrH_RtiDnD      | 12.8      | 9.8       | 21.5        | 0.61 | −0.01  | DNA   |                |
| GaTrH_RtiDnD      | 12.9      | n/a       | 21.2        | 0.39 | −0.34  | OK    | CCdamo, CVhigh  |
| GaTrH_RtiDnD      | 12.6      | n/a       | 0.6         | 0.64 | −0.30  | DNA   | CV Tisi, CVhigh |
| GaTrH_DiDnD       | 12.2      | 3.5       | 12.5        | 0.59 | 0.65   | OK    | DNA, CCdamo, CVlow |
| GaTrH_DiDnD       | 12.3      | 9.8       | 22.0        | 0.59 | −0.01  | DNA   |                |
| GaTrH_DiDnD       | 13.0      | 2.6       | 15.7        | 0.53 | −0.03  | DNA   | DNA, CVlow    |
| GaTrH_DiDnD       | 13.1      | 9.7       | 21.6        | 0.59 | −0.01  | DNA   |                |
| GaTrH_MidmDnD      | 13.1      | 3.5       | 12.3        | 0.29 | 0.17   | OK    | CC CSI, CVlow  |
| GaTrH_MidmDnD      | 13.1      | 9.8       | 19.2        | 0.62 | −0.20  | OK    |                |
| GaTrH_MidmDnD      | 12.8      | 10.9      | 21.7        | 0.46 | −0.29  | DNA   |                |
| GaTrH_MidmDnD      | 12.8      | 9.3       | 18.2        | 0.59 | −0.18  | OK    |                |
| GaTrH_TiDnD       | 13.0      | 3.5       | 12.2        | 0.31 | 0.06   | OK    | CC CSI, CVlow  |
| GaTrH_TiDnD       | 14.8      | 9.8       | 19.0        | 0.63 | −0.35  | OK    |                |
| GaTrH_TiDnD       | 18.2      | n/a       | 37.3        | 0.68 | −0.83  | OK    | CV high       |
| GaTrH_TiDnD       | 12.9      | n/a       | 5.6         | 0.90 | −0.87  | OK    | CV high       |
| GaTrH_RtiDnD      | 12.8      | 3.5       | 12.4        | 0.45 | 0.28   | Fails  | DNA, CCdamo, CVlow |
| GaTrH_RtiDnD      | 13.1      | 12.3      | 19.6        | 0.55 | −0.19  | DNA   |                |
| GaTrH_RtiDnD      | 12.9      | n/a       | 16.9        | 0.50 | −0.36  | OK    | CV high       |
| GaTrH_RtiDnD      | 12.8      | 1.5       | 2.1         | 0.66 | −0.21  | OK    | CV Tisi, CVlow |
| GaTrH_DiDnD       | 13.0      | 3.5       | 12.6        | 0.48 | 0.34   | Fails  | DNA, CCdamo, CVlow |
| GaTrH_DiDnD       | 13.4      | 9.8       | 21.8        | 0.53 | −0.15  | Fails  | DNA           |
| GaTrH_DiDnD       | 12.9      | 1.8       | 12.7        | 0.48 | 0.24   | Fails  | DNA, CCdamo, CVlow |
| GaTrH_DiDnD       | 12.9      | 10.0      | 21.7        | 0.49 | −0.19  | Fails  | DNA           |
| PoTrB_MidmDnD      | 15.7      | 3.5       | 15.2        | 0.47 | −0.58  | OK    | CV low        |
| PoTrB_MidmDnD      | 13.3      | 8.0       | 17.3        | 0.28 | −0.26  | OK    | CC CSI        |
| PoTrB_MidmDnD      | 13.3      | 13.4      | 22.5        | 0.62 | −0.47  | OK    |                |
| PoTrB_MidmDnD      | 13.2      | 8.5       | 19.1        | 0.46 | −0.40  | OK    |                |
| PoTrB_TiDnD       | 13.1      | 3.5       | 12.6        | 0.19 | −0.40  | OK    | CC CSI, CVlow  |
| PoTrB_TiDnD       | 12.9      | 8.0       | 14.9        | 0.37 | −0.44  | OK    | CC CSI        |
| PoTrB_TiDnD       | 13.3      | n/a       | 36.5        | 0.78 | −0.90  | OK    | CV high       |
| PoTrB_TiDnD       | 13.3      | n/a       | 35.3        | 0.85 | −0.93  | OK    | CV high       |
| PoTrB_RtiDnD      | 13.0      | 3.5       | 13.1        | −0.01 | −0.05  | Fails  | CC CSI, CVlow, DNA |
| PoTrB_RtiDnD      | 12.8      | 8.0       | 17.5        | 0.23 | −0.23  | Fails  | CC CSI, DNA    |
| PoTrB_RtiDnD      | 13.3      | n/a       | 19.3        | 0.45 | −0.44  | Fails  | DNA, CV high   |
| PoTrB_RtiDnD      | 13.3      | 10.1      | 15.7        | 0.40 | −0.35  | Fails  | DNA           |
| PoTrB_DiDnD       | 12.8      | 3.5       | 13.0        | −0.05 | −0.01  | Fails  | CC CSI, DNA, CVlow |
| PoTrB_DiDnD       | 13.2      | 8.0       | 19.1        | 0.22 | −0.22  | Fails  | CC CSI, DNA    |
| PoTrB_DiDnD       | 13.3      | 2.1       | 16.7        | 0.28 | −0.42  | Fails  | CC CSI, DNA, CVlow |
| PoTrB_DiDnD       | 13.3      | 8.0       | 19.1        | 0.22 | −0.23  | Fails  | CC CSI, DNA    |
| PoTrB_MidmDnD      | 13.0      | 3.5       | 11.4        | 0.22 | 0.25   | OK    | CC CSI, CVlow  |

*New Journal of Physics 14 (2012) 095016 ([http://www.njp.org/](http://www.njp.org/))*. 
| Model           | CV gr (%) | CV DM (%) | CV Ti (%) | CCsisi | CCDamo | DNA    | Fails due / OK |
|----------------|-----------|-----------|-----------|--------|--------|--------|----------------|
| PoTfM_MiDgNdM  | 17.2      | 8.0       | 19.2      | 0.62   | 0.09   | OK     | OK             |
| PoTfM_MiDmNdD  | 13.2      | 10.0      | 16.4      | 0.60   | −0.23  | OK     | OK             |
| PoTfM_MiDmNdM  | 13.1      | 7.7       | 15.6      | 0.57   | −0.07  | OK     | OK             |
| PoTfM_TiDgNdD  | 13.1      | 3.5       | 11.6      | 0.30   | 0.14   | OK     | CCsisi, CVlow  |
| PoTfM_TiDgNdM  | 17.2      | 8.0       | 17.3      | 0.64   | −0.09  | OK     | OK             |
| PoTfM_TiDmNdD  | 13.1      | n/a       | 16.7      | 0.75   | −0.80  | OK     | CVhigh         |
| PoTfM_TiDmNdM  | 13.0      | 32.1      | 2.3       | 0.79   | −0.58  | OK     | CV T;         |
| PoTfM_RiDiDgNdD| 12.9      | 3.5       | 11.8      | 0.50   | 0.38   | Fails  | DNA, CCDamo, CVlow |
| PoTfM_RiDiDgNdN|M 12.7     | 8.0       | 17.3      | 0.54   | −0.11  | Fails  | DNA            |
| PoTfM_RiDiDmNdD| 13.2      | 10.0      | 15.5      | 0.50   | −0.30  | Fails  | DNA            |
| PoTfM_RiDiDmNdN|M 13.2      | 10.4      | 9.8       | 0.61   | −0.19  | Fails  | DNA            |
| PoTfM_DiDiNdD  | 12.7      | 3.5       | 11.7      | 0.52   | 0.43   | Fails  | DNA, CCDamo, CVlow |
| PoTfM_DiDiNdN|M 13.3      | 8.0       | 18.5      | 0.54   | −0.08  | Fails  | DNA            |
| PoTfM_DiDiNdD  | 13.2      | 1.2       | 12.9      | 0.54   | 0.01   | Fails  | DNA, CVlow     |
| PoTfM_DiDiNdN|M 13.1      | 8.1       | 18.4      | 0.53   | −0.10  | Fails  | DNA            |
| PoTrL_MiDgNdD  | 10.4      | 3.5       | 10.9      | 0.45   | 0.49   | OK     | CCDamo, CVlow  |
| PoTrL_MiDgNdN|M 8.6       | 8.0       | 17.9      | 0.60   | 0.01   | OK     | OK             |
| PoTrL_MiDmNdD  | 12.9      | 13.5      | 17.7      | 0.68   | −0.03  | (OK)   | DNA            |
| PoTrL_MiDmNdN|M 13.3      | 8.6       | 15.1      | 0.77   | 0.36   | (OK)   | CCDamo        |
| PoTrL_TiDgNdD  | 10.4      | 3.5       | 10.9      | 0.46   | 0.46   | OK     | CCDamo, CVlow  |
| PoTrL_TiDgNdN|M 10.6      | 8.0       | 17.8      | 0.62   | −0.04  | OK     | OK             |
| PoTrL_TiDmNdD  | 12.8      | n/a       | 32.1      | 0.82   | −0.93  | OK     | CVhigh         |
| PoTrL_TiDmNdN|M 12.5      | 8.1       | 28.0      | 0.97   | −0.96  | OK     | CVhigh         |
| PoTrL_RiDiNdD  | 12.0      | 3.5       | 12.4      | 0.58   | 0.60   | (Fails) | CCDamo, CVlow |
| PoTrL_RiDiNdN|M 11.9      | 8.0       | 18.2      | 0.63   | 0.10   | Fails  | DNA, CCDamo    |
| PoTrL_RiDiNdD  | 12.6      | n/a       | 17.9      | 0.40   | −0.33  | OK     | CVhigh         |
| PoTrL_RiDiNdN|M 12.7      | n/a       | 0.6       | 0.67   | −0.23  | OK     | CV T; CVhigh   |
| PoTrL_DiDiNdD  | 12.2      | 3.5       | 12.4      | 0.60   | 0.66   | fails  | DNA, CCDamo, CVlow |
| PoTrL_DiDiNdN|M 12.2      | 8.0       | 19.0      | 0.63   | 0.13   | Fails  | DNA, CCDamo    |
| PoTrL_DiDiNdD  | 12.4      | 3.5       | 17.9      | 0.62   | −0.41  | Fails  | DNA, CVlow     |
| PoTrL_DiDiNdN|M 12.7      | 10.2      | 28.3      | 0.87   | −0.21  | Fails  | DNA            |
| PoTrH_MiDiNdD  | 13.3      | 3.5       | 12.3      | 0.36   | 0.06   | OK     | CCsisi, CVlow  |
| PoTrH_MiDiNdN|M 13.2      | 8.0       | 18.1      | 0.58   | −0.08  | OK     | OK             |
| PoTrH_MiDiNdD  | 12.9      | 10.2      | 16.4      | 0.61   | −0.18  | OK     | OK             |
| PoTrH_MiDiNdN|M 12.8      | 7.8       | 16.2      | 0.58   | −0.07  | OK     | OK             |
| PoTrH_TiDiNdD  | 13.2      | 3.5       | 12.3      | 0.33   | −0.02  | OK     | CCsisi, CVlow  |
| PoTrH_TiDiNdN|M 13.0      | 8.0       | 15.9      | 0.62   | −0.31  | OK     | OK             |
| PoTrH_TiDiNdD  | 12.9      | n/a       | 32.3      | 0.77   | −0.89  | OK     | CVhigh         |
| PoTrH_TiDiNdN|M 12.9      | 5.3       | 8.9       | 0.89   | −0.87  | OK     | CVhigh         |
| PoTrH_RiDiNdD  | 13.1      | 3.5       | 12.5      | 0.45   | 0.26   | (Fails) | CCDamo, CVlow |
| PoTrH_RiDiNdN|M 12.5      | 8.0       | 16.6      | 0.56   | −0.12  | (Fails) | DNA            |
| PoTrH_RiDiNdD  | 12.9      | 1.0       | 13.4      | 0.53   | −0.36  | OK     | CVlow          |
| PoTrH_RiDiNdN|M 9.9       | 19.5      | 4.0       | 0.58   | −0.26  | OK     | CV T;         |
| PoTrH_DiDiNdD  | 12.7      | 3.5       | 12.3      | 0.47   | 0.32   | Fails  | DNA, CCDamo, CVlow |
| PoTrH_DiDiNdN|M 13.6      | 8.0       | 19.0      | 0.55   | −0.05  | Fails  | DNA            |
| PoTrH_DiDiNdD  | 12.9      | 1.7       | 13.9      | 0.51   | −0.16  | Fails  | DNA, CVlow     |
| PoTrH_DiDiNdD  | 13.0      | 8.0       | 18.6      | 0.51   | −0.08  | Fails  | DNA            |
Figure 6. DNA content as a function of doubling rate $\mu = 60 \text{ min}/T_i$. The MI model reproduces the experimental data nearly perfectly, while DI generally produces too high DNA content. Some models, such as the RTI model, yield an intermediate DNA content.

As can be seen from table 3, more than two thirds of the models (94 out of the 128 models) cannot be brought into agreement with all experimental observations. With the available experimental data the number of consistent models cannot be further reduced.

As the remaining number of 34 surviving models is still quite high, we searched for further possible variables to distinguish between those models. There are two quantities that are in principle experimentally accessible, but have so far not been systematically measured: (i) the correlation of interdivision times between sister cells at low growth rates (i.e. for $T_i > 80$ min), referred to as Sisi80 in the following, and (ii) the correlation between mass doubling time $T_m$ and interdivision time $T_i$ of individual cells, referred to as $CC_{TD}T_i$. For both quantities we expect significant positive correlation coefficients. In a next step, we calculated these quantities for the 34 consistent models. The results are shown in table 4.

As can be seen, a number of models show a small or even negative $CC_{sis80}$ (marked in red), despite a strongly positive correlation in the case of fast growth (i.e. for $T_m = 40$ min). In all models the correlation between $T_m$ and $T_i$ is positive. However, the models cluster into three distinct groups: ten models have a value close to zero (marked in red), 18 models a value around 0.3 (marked in yellow) and 11 models show a strong positive correlation with correlation coefficients greater than 0.5 (marked in green).

With the moderate assumption that the values of these two correlation coefficients are greater than 0.4, only ten of the original 128 models survive (marked in green).
Figure 7. Coefficients of correlation between sisters (CCsisi) and between daughters and mothers (CCdamo). Typical examples for very high correlation (left column), no correlation (right column) and a case in between (middle). All data shown are for $T_m = 40$ min.

3. Discussion

In this study, we theoretically analyze different regulatory models for mass growth and DNA replication of fast-growing E. coli cells. In these cells chromosome replication is initiated in the mother cell but completed only in the daughter cells. Thus, the individual cells cannot control the time point of DNA replication initiation. Rather, they have to rely on regulation of this process in the mother cell. This phenotypic inheritance effectively implements a sort of memory in the system, which is one of the main reasons why robust (i.e. stable with respect to stochastic variations in the composition of the individual cells) regulation of DNA replication is so complex. We tried to identify those regulatory schemes that are not able to show such a robust regulation of DNA replication and cell division under the influence of single-cell noise.

The constraints we used in order to classify a model as being consistent were therefore not very severe, as we aimed at excluding only those models that fundamentally fail. Given the quality of some of the experimental data (on, e.g., DNA content measures and sister–sister correlations of interdivision times) setting these limits is anyhow somewhat arbitrary. Nevertheless, we found that, surprisingly, almost three quarters (94 of 128) of the models were not able to reproduce the available experimental results, and can therefore be excluded from further consideration. The main advantage of our approach is thus that we are able to identify key variables that allow the comparison of the theoretical models with the available experimental data.
Table 4. Measures for CCsisi80 and CCTDti for the 34 surviving models. The coefficient of correlation (CC) of sister interdivision times at Tm = 80 min (CCsisi80) was marked in red in the case of CCsisi80 ≤ 0.2. The CC between mass doubling time Tm and interdivision time Ti (CCTDti) was marked in red for very low values (CCTDti ≤ 0.2), green for high values (CCTDti > 0.5) and yellow for those in between. For green marked models, both quantities are larger than 0.4.

| Model              | CCsisi80 | CCTDti |
|--------------------|----------|--------|
| GaTfB_MiDgNdM     | 0.34     | 0.30   |
| GaTfB_MiDmNdD     | 0.33     | 0.34   |
| GaTfB_MiDmNdM     | 0.20     | 0.43   |
| GaTfB_TiDgNdM     | 0.20     | 0.29   |
| GaTM_MiDgNdM      | 0.57     | 0.31   |
| GaTM_MiDmNdD      | 0.47     | 0.11   |
| GaTM_MiDmNdM      | 0.50     | 0.31   |
| GaTM_TiDgNdM      | 0.55     | 0.36   |
| GaTM_RtiDmNdD     | 0.39     | 0.14   |
| GaTM_RtiDmNdM     | 0.64     | 0.13   |
| GaTrL_MiDgNdM     | 0.62     | 0.51   |
| GaTrL_MiDmNdD     | 0.52     | 0.38   |
| GaTrL_TiDgNdM     | 0.64     | 0.55   |
| GaTrH_MiDgNdM     | 0.53     | 0.38   |
| GaTrH_MiDmNdD     | 0.45     | 0.17   |
| GaTrH_MiDmNdM     | 0.44     | 0.42   |
| GaTrH_TiDgNdM     | 0.47     | 0.45   |
| GaTrH_RtiDgNdM    | 0.49     | 0.44   |
| PoTfB_MiDgNdD     | -0.31    | 0.88   |
| PoTfB_MiDmNdD     | 0.11     | 0.45   |
| PoTfB_MiDmNdM     | 0.11     | 0.45   |
| PoTM_MiDgNdM      | 0.39     | 0.83   |
| PoTM_MiDmNdD      | 0.51     | 0.32   |
| PoTM_MiDmNdM      | 0.51     | 0.32   |
| PoTM_TiDgNdM      | 0.53     | 0.86   |
| PoTrL_MiDgNdM     | 0.27     | 0.78   |
| PoTrL_MiDmNdD     | 0.67     | 0.74   |
| PoTrL_TiDgNdM     | 0.43     | 0.79   |
| PoTrH_MiDgNdM     | 0.19     | 0.82   |
| PoTrH_MiDmNdD     | 0.44     | 0.48   |
| PoTrH_MiDmNdM     | 0.44     | 0.48   |
| PoTrH_TiDgNdM     | 0.24     | 0.79   |
| PoTrH_RtiDgNdM    | 0.33     | 0.84   |
| PoTrH_RtiDmNdD    | 0.68     | 0.16   |

In deriving these results we made the following general, well-established assumptions (see, for instance, [5] for a review): (i) exponential mass increase of individual cells; (ii) the regulation of DNA replication takes place at the level of initiation; (iii) only viable cells are produced (i.e. cell division is prevented in the case of incomplete DNA replication); (iv) every cell initiates
DNA replication exactly once per cell cycle, whereby replication is triggered from all oriC sites simultaneously.

In analyzing the properties of cellular populations, we took into account the effects arising from noise affecting the molecular composition of individual cells. External sources of noise, such as unequal accessibility of nutrients, temperature fluctuations or unsynchronized adaption phases, typically do not play a role in the relevant experimental situations and were therefore not considered here.

For given external conditions, mass growth of a single cell is characterized by three quantities: mass at birth, mass at division and growth rate at a given time. Our growth model includes noise coupled to all three quantities allowing us to explore the full parameter space for noisy single-cell mass growth. In doing so, we were able to show that the experimental observations cannot be explained by a single source of noise; rather all three quantities have to be noisy. In this case the free parameters of our models can be adjusted to yield good agreement with the experimental data (see table 2). In particular, our models then capture the CV of the distributions of division mass, growth rate and division site placement accuracy and yield the correct CV of $T_i$ distributions of about 25%. After having fixed all parameters for the various noise sources it turned out that the mass growth of bacteria is dominantly affected by the growth rate noise.

There is also a number of additional conclusions that we can draw from our findings summarized in table 3. The results essentially do not depend on the type of Nd noise. For a Gaussian and a Poisson distribution almost exactly the same model variants are in agreement with the experimental data. Furthermore, out of the 23 ‘green’ models in table 3, 11 are Ga and 12 are Po variants.

Furthermore, we found that for fast growth the details of growth rate changes do not have a significant effect on the observed sister–sister correlations, i.e. it hardly makes a difference whether growth rate changes occur randomly or at fixed points during the cell cycle. This is surprising, since the TfB models lead to completely uncorrelated growth rates of sister cells. In this case, growth rate noise does not contribute to positive sister–sister correlations. This is different from all other model variants, where sister cells on average grow with the same growth rate (for at least some time). We thus conclude that for fast-growing bacteria, an observed positive sister–sister correlation of interdivision times can be a feature of the DNA replication process. This finding is further supported by the fact that all of the TfB models show remarkably weaker sister–sister correlations when the doubling time is raised to $T_i = 80$ min (see tables 3 and 4). Note that in this case of slow growth, daughter cells do not receive DNA that is already being replicated and therefore can no longer provide a correlation between sisters. Except for one model (PoTFB_MiDgNdD, which turns to negative values), they all continue to show slightly positive sister–sister correlations which arise from the contribution of Nd noise.

This emphasizes the relevance of sister–sister correlations as a key variable. These findings can be tested experimentally by systematically analyzing sister–sister correlations under varying growth conditions that in particular include slow growth.

It is worth noting that almost all DgNdD models fail, while neither Dg nor NdD alone is the reason for this, emphasizing the importance of our approach, to combine and systematically test all different model assumptions as opposed to isolating and individually testing different variants.

The most important finding is that all DI and almost all RTI models fail, while those that are completely in agreement with all experimental data are either MI or TI models. From this
we conclude that the mass growth and the DNA replication cycle need to be coupled, either by coupling DNA replication initiation to the event of cell division (\(TI\)) or to the amount of accumulated mass (\(MI\)). It thus seems not sufficient if each of the cycles operates independently with only check points ensuring the avoidance of non-viable cells. However, complicated the molecular machinery of replication control works, as long as it just couples to itself (amount of DNA or time since last replication), it will fail to work robustly under noisy conditions.

In this study, we have demonstrated the importance of combining and systematically testing different model variants, especially with respect to noise on the single-cell level. With our novel approach of introducing a series of models for the unknown processes, we are able to significantly reduce the number of coupling schemes that are compatible with all available experimental data. To further reduce the number of compatible models additional experimental investigations are necessary. We propose here to measure two quantities, namely the sister–sister correlations of interdivision times of slow-growing *E. coli* cells and the correlation coefficient between mass doubling time and interdivision time of individual cells. We expect these quantities to both show positive correlations. If true, then only ten model variants will be compatible with all experimental data. It is notable that eight of these ten are NdM models, suggesting that cellular division timing inaccuracy is mostly due to mass growth alone and not due to DNA replication delays.

To summarize, in this study we have applied an ensemble approach to analyze the regulation of initiation of DNA replication. In this approach, we introduce a variety of different models that make different assumptions about unknown variables and processes. A similar approach has already successfully been applied to the analysis of the TOR pathway of *Saccharomyces cerevisiae* [34]. For our system this approach allows us to identify key variables that can be used to experimentally discriminate the different models. In this way, a refined qualitative understanding of this highly complex regulatory process is obtained with quantitative methods, despite the many unknowns.

**Acknowledgments**

We thank an anonymous referee for drawing our attention to [34]. This work was partially supported by the LOEWE program of the State Hessen.

**Appendix. Materials and methods**

**A.1. Implementation**

In our simulations, each cell is identified by an individual cell number and characterized by a set of 16 variables, such as current mass, growth rate, lifetime of the mother cell or the cell number of the sister cell; see section 3 of the online supplementary data available from stacks.iop.org/NJP/14/095016/mmedia for a complete list. The interdivision time \(T_{m}\) set by the medium is used to set the initial birth mass, DNA content and replication fork pattern of the first ideal cell; it thus defines a lower boundary for the achievable mean interdivision time \(\bar{T}_i\) of the population.

In all our simulations we set the time from initiation to completion of one round of DNA replication (C period) to 40 min, while the time from termination to subsequent cell division...
(D period) is set to 20 min. The latter is kept fixed in some cases and subjected to noise in others, as explained in the main text.

The amount of DNA already replicated at birth depends on the growth regime. To characterize this dependence it is useful to introduce the number of generations \( n \) back in time when the DNA replication to be completed in a given cell needs to have been initiated. With \( C + D = 60 \text{ min} \) and a lower boundary for \( T_i \approx D = 20 \text{ min} \), there are only three possible values for \( n = 0, 1, 2 \). Accordingly, the transition between the different regimes occurs at \((C + D)/(n + 1) = 60, 30, 20 \text{ min}\). In order to minimize the influence of these regime boundaries, we run simulations preferably for a medium with \( T_m = 40 \text{ min} \). In the cases where \( T_m \) was varied, we used the range \( 21 \text{ min} \leq T_m \leq 80 \text{ min} \).

In the absence of noise a cell needs to initialize DNA replication at age \( a_{\text{crit}} [2, 3] \):\[ a_{\text{crit}} = T_m \left( 1 + \left[ \frac{C + D}{T_m} \right] - \frac{C + D}{T_m} \right). \] (A.1)

Here, \( C + D \) is the total time for the DNA replication process and \([x]\) denotes the integer part of \( x \).

In the TI models, replication starts at this age on average (due to the effects of noise). The MI models start replication also on average at this age triggered, however, by the corresponding critical mass given as \[ m_{\text{crit}} = 2^{(a_{\text{crit}}/T_m)}. \] (A.2)

where we have assumed that cell mass increases exponentially with time [36, 37]. The RTI coupling scheme sets the initial age at which to start DNA replication for the first cell to \( a_{\text{crit}} \), and continues to initiate the following rounds of DNA replication at \( T_m \) time intervals.

In the case of the DI models, DNA replication is supposed to start when a critical variable (Dna\(^*\) per oriC) that is produced proportional to a certain section of DNA reaches a threshold value. We performed simulations of this model and measured the content of different sections of DNA at the time of initiation. Figure A.1 shows the results for the full range of interdivision times for fast-growing \( E. \) coli cells. As can be seen, this mechanism only works if Dna\(^*\) reaches the threshold value within the first 40% of the whole DNA replication time. Otherwise replication initiation occurs too late at faster growth rates. For our simulations, we therefore chose the first 40% of DNA measured from oriC to serve as the threshold value needed.

All simulations start with a single cell initialized with a non-noisy history. With evolving time, cell variables such as age, mass and DNA content are updated. Eventually, cells divide when both the division cycle and the DNA replication cycle are completed. The properties of newborn cells are calculated at birth, but, depending on the model, they might be re-adjusted at certain events throughout the lifetime of a cell. While DNA strands are always distributed equally among the two daughter cells, mass partitioning is subject to noise. Other noise sources affect growth rates or division timing. The mean values of all cell variables are recorded every minute, while population observables such as correlation coefficients are calculated each time the population has grown tenfold. Because the values of all relevant variables (i.e. lifetime of the mother cell and the amount of already replicated DNA by the mother and grandmother cells) are recorded for all cells, these calculations can be easily performed at any point in time.

A simulation typically ran for at least 5000 min. All correlation coefficients were calculated using the Spearman rank correlation method from Numerical Recipes [38]. All programs were written in C, compiled using gcc and run of PC architecture.
Figure A.1. The amount of a substance Dna* per cell per oriC at the time of initiation $a_{\text{crit}}$.

For further details on the simulation, see online supplementary data available from stacks.iop.org/NJP/14/095016/mmedia.

A.2. Noise parameters

To determine the parameter values for the various noise sources, we performed a series of simulations of bacterial growth. We did this in the absence of DNA replication to see whether implementation of the various sources of noise is sufficient to explain the available experimental data. Furthermore, in this way we analyze the robustness of our growth model with respect to the presence of noise.

We first implemented only noise on the accuracy of division site placement (NB). From experimental data it is evident that noise has a symmetric distribution. We therefore assumed a Gaussian distribution of mass partitioning at division with standard deviation $\sigma_B$. Fitting to the available experimental data [12–14] suggests a value of $\sigma_B = 3$–5%.

Inclusion of ND and NGr noise had no visible influence on these findings. In the following we set $\sigma_B = 3.3\%$.

In a next step, we checked if ND and NGr depend on each other. As a first test, we analyzed the influence of the timing parameters from NGr, $\lambda_{\text{time}}$ and $a_{\text{time}}$, on population growth; see figures 2 and 3. As it turned out, the influence of $\lambda_{\text{time}}$ on distribution of individual interdivision times $T_i$ is only weak; however, the observed mean population interdivision time, $\bar{T}_i$, decreases as $\lambda_{\text{time}}$ decreases. In the TF models, $a_{\text{time}}$ only slightly affects the distribution of $T_i$, yielding a minimum value of $\bar{T}_i$ for values $a_{\text{time}} \approx 0.5$. 

New Journal of Physics 14 (2012) 095016 (http://www.njp.org/)
Given this weak dependence of the other observables on $\lambda_{\text{time}}$ and $a_{\text{time}}$, we performed the parameter scans by varying only the two remaining parameters ($\sigma_{\text{Gr}}, \lambda_D$), respectively ($\sigma_{\text{Gr}}, \sigma_D$), for the four model variants: (H) with $\lambda_{\text{time}} = 0.03$, (L) with $\lambda_{\text{time}} = 0.005$, (B) with $a_{\text{time}} = 0$, and (M) with $a_{\text{time}} = 0.5$. As each was simulated with both types of ND noise, this yields a total of eight model variants. Figure 4 shows an example of the noise parameter space for the model GaTrH (which is the model Ga in variant Tr at high $\lambda_{\text{time}}$); similar plots for the other seven model variants can be found in online supplementary data available from stacks.iop.org/NJP/14/095016/mmedia. Analysis of all these simulations then showed that in the absence of DNA the observed CV of the division mass (DM) distribution and that of the growth rate distribution can be independently chosen in all model variants. Thus, ND and NGr do not affect each other.

To reproduce the observed CV of division mass of 10% [13, 18] we adjusted the noise parameters of ND to $\lambda_D = 0.05$ and $1/\sigma_D = 0.08$. See supplementry figures S1–S4 for details. Finally, to reproduce the observed CV of growth rate noise of 13% [20], we have adjusted the only remaining free parameter $\sigma_{\text{Gr}}$ yielding $\sigma_{\text{Gr}} = 0.13$.

References

[1] Bremer H, Dennis P and Neidhart F 1996 Modulation of chemical composition and other parameters of the cell by growth rate Escherichia coli and Salmonella: Cellular and Molecular Biology (Washington, DC: ASM) pp 1553–69
[2] Cooper S and Helmstetter C E 1968 Chromosome replication and the division cycle of Escherichia coli B/r J. Mol. Biol. 31 519–40
[3] Helmstetter C E and Cooper S 1968 DNA synthesis during the division cycle of rapidly growing Escherichia coli B/r J. Mol. Biol. 31 507–18
[4] Donachie W D 1968 Relationship between cell size and time of initiation of DNA replication Nature 219 1077–9
[5] Vinella D and D’Ari R 1995 Overview of controls in the Escherichia coli cell cycle Bio Essays 17 527–36
[6] Boye E and Nordström K 2003 Coupling the cell cycle to cell growth EMBO Rep. 4 757–60
[7] Wold S, Skarstad K, Steen H B, Stokke T and Boye E 1994 The initiation mass for DNA replication in Escherichia coli K-12 is dependent on growth rate EMBO J. 13 2097–102
[8] Cooper S 1997 Does the initiation mass for DNA replication in Escherichia coli vary with growth rate? Mol. Microbiol. 26 1137–60
[9] Spudich J and Koshland D 1976 Non-genetic individuality: chance in the single cell Nature 221 467–71
[10] Elowitz M B, Levine A J, Siggia E D and Swain P S 2002 Stochastic gene expression in a single cell Science 297 1183–6
[11] Smits W, Kuipers O and Veening J 2006 Phenotypic variation in bacteria: the role of feedback regulation Nature Rev. Microbiol. 4 259–71
[12] Trueba F J 1982 On the precision and accuracy achieved by Escherichia coli cells at fission about their middle Arch. Microbiol. 131 55–9
[13] Koppes L H, Woldringh C L and Nanninga N 1978 Size variations and correlation of different cell cycle events in slow-growing Escherichia coli J. Bacteriol. 134 423–33
[14] Guberman H M, Fay A, Dworkin J, Wingreen N S and Gitai Z 2008 PSICIC: noise and asymmetry in bacterial division revealed by computational image analysis at sub-pixel resolution PLoS Comput. Biol. 4 1000233
[15] Reshes G, Vanounou S, Fishov I and Feingold M 2008 Timing the start of division in E. coli: a single-cell study Phys. Biol. 5 046001
[16] Wakamoto Y, Ramsden J and Yasuda K 2005 Single-cell growth and division dynamics showing epigenetic correlations Analyst 130 311–7

New Journal of Physics 14 (2012) 095016 (http://www.njp.org/)
[17] Schaechter M, Williamson J P, Hood J R Jr and Koch A L 1962 Growth, cell and nuclear divisions in some bacteria J. Gen. Microbiol. 29 421–34
[18] Bremer H 1986 A stochastic process determines the time at which cell division begins in Escherichia coli J. Theor. Biol. 118 351–65
[19] Browning S T, Castellanos M and Shuler M L 2004 Robust control of initiation of prokaryotic chromosome replication: essential considerations for a minimal cell Biotechnol. Bioeng. 88 575–84
[20] Tsuru S, Ichinose J, Kashiwagi A, Ying B-W, Kaneko K and Yomo T 2009 Noisy cell growth rate leads to fluctuating protein concentration in bacteria Phys. Biol. 6 036015
[21] Wang P, Robert L, Pelletier J, Dang W L, Taddei F, Wright A and Jun S 2010 Robust growth of Escherichia coli Curr. Biol. 20 1099–103
[22] Kaguni J M 2006 DnaA: Controlling the initiation of bacterial DNA replication and more Annu. Rev. Microbiol. 60 351–71
[23] Bremer H and Churchward G 1991 Control of cyclic chromosome replication in Escherichia coli Microbiol. Rev. 55 459–75
[24] Koch A L 2002 Control of the bacterial cell cycle by cytoplasmic growth. Crit. Rev. Microbiol. 28 61–77
[25] Cooper S 2006 Bacterial Growth and Division (Weinheim: Wiley-VCH)
[26] Mahaffy J M and Zyskind J W 1989 A model for the initiation of replication in Escherichia coli J. Theor. Biol. 140 453–77
[27] Skarstad K, Løbner-Olesen A, Atlung T, von Meyenburg K and Boye E 1989 Initiation of DNA replication in Escherichia coli after overproduction of the DnaA protein Mol. Gen. Genet. 218 50–6
[28] Riber L and Løbner-Olesen A 2005 Coordinated replication and sequestration of oriC and DnaA are required for maintaining controlled once-per-cell-cycle initiation in Escherichia coli J. Bacteriol. 187 5605–13
[29] Nordström K, Bernander R and Dasgupta S 1991 The Escherichia coli cell cycle: one cycle or multiple independent processes that are co-ordinated? Mol. Microbiol. 5 769–74
[30] Løbner-Olesen A, Skarstad K, Hansen F G, Meyenburg K v and Boye E 1989 The DnaA protein determines the initiation mass of Escherichia coli K-12 Cell 57 881–9
[31] Donachie W D and Blakely G W 2003 Coupling the initiation of chromosome replication to cell size in Escherichia coli Curr. Opin. Microbiol. 6 146–50
[32] Haeusser D P and Levin P A 2008 The great divide: coordinating cell cycle events during bacterial growth and division Curr. Opin. Microbiol. 11 94–9
[33] Skarstad K, Boye E and Steen H B 1986 Timing of initiation of chromosome replication in individual Escherichia coli cells EMBO J. 5 1711–7
[34] Kuepfer L, Peter M, Sauer U and Stelling J 2007 Ensemble modeling for analysis of cell signaling dynamics Nature Biotechnol. 25 1001–6
[35] Koch A L and Schaechter M 1962 A model for statistics of the cell division process J. Gen. Microbiol. 29 435–54
[36] Trueba F J and Koppes L J H 1998 Exponential growth of Escherichia coli B/r during its division cycle is demonstrated by the size distribution in liquid culture Arch. Microbiol 169 491–6
[37] Cooper S 1988 What is the bacterial growth law during the division cycle? J. Bacteriol. 170 5001–5
[38] Press W, Flannery B, Teukolsky S and Vetterling W 1992 Numerical Recipes in C: The Art of Scientific Computing (Cambridge: Cambridge University Press)