CONTENTS

I. MODEL DEVELOPMENT

The following subsections provide additional details of the core assumptions, calculations, and methodologies used to formulate the heterogeneous multiphase growth model (HMG model) described in the main text.

Nomenclature

Term	Definition
V_i	Volume at initiation, also called critical volume ($0.9\ \mu m^3 \equiv 0.9\ fL$)[1-2]
V_i **Noise**	Gaussian noise standard deviation (10% for dt = 0.01 min)[3]
V_a	Cell volume at cell cycle time a
V_a **Noise**	Gaussian noise standard deviation (5% for dt = 0.01 min)[3]
μ	Growth rate (min^{-1})
τ	Doubling rate (min^{-1}), where $\tau = \frac{ln(2)}{\mu}$
C	Genome replication time (min)
C Noise	Gaussian noise standard deviation (5% for dt = 0.01 min)[3]
D	Binary fission time (min)
D Noise	Gaussian noise standard deviation (for dt = 0.01 min)
Segregation Asymmetry	Gaussian noise standard deviation (10% for dt = 0.01 min)[3] describing the asymmetry of binary fission
Chance of Initiation	Given critical volume reached, probability that the OriC opens (1.75 for dt = 0.01 min)[3]
Chance of DNA damage	Probability that the replicating chromosome experiences any DNA damage
Ratio of DNA damage	If the replicating chromosome experiences DNA damage, this ratio represents the chance that the DNA damage leads to degradation of only the replicating strand, divided by the chance that such damage leads to degradation of the whole chromosome

1. Relating experimental and simulated DNA distributions

Experimentally measured DNA distributions reflect accumulated experimental variation associated with both labeling (e.g., variable efficiencies of fixation, permeabilization, and binding of DAPI to DNA) and detection (e.g., variability in the signal measured by flow cytometry given a fixed quantity of DAPI in a single sample). Therefore, even if a population of cells included a discrete and uniform amount of DNA per cell, the measured DNA distribution would be "spread" by these other sources of variability,[4] and this spread must be quantified in order to relate our simulation results to measured DNA distributions[3, 5]. To quantify the spread associated with our experimental methods, we chose a sample in which cells were grown to stationary phase, since under these conditions, it is reasonable to assume that all chromosomes had completed replication and thus actual DNA content per cell represents a discrete number of full genome units. The DAPI-associated fluorescence was thus fit to a sum of Gaussians (Figure S5). If the cell contains a wild type (WT) version of the RecA gene, then each peak would be assigned a value of 1, 2, 4, or 8 genome equivalents of DNA, since all chromosomes are assumed to initiate replication at OriC at the same time, and thus the number of completed chromosomes should be a power of two [6]. However, because the TOP10 bacterial strain used in this study contains a mutation in the gene encoding RecA (*recA1*, see section 3 below)[7], the cells have aberrant chromosome copy numbers[8], and as a consequence, each peak was assigned an integer value (1, 2, 3, …) of genome equivalents of DNA.

From these fits, we determined two important parameters. From the means of the Gaussians, we generated a calibration curve for relating discrete numbers of chromosomes to flow cytometry channels (levels of DAPI-associated fluorescence per cell). Notably, the standard deviation of the peaks followed the same pattern observed in previous reports[4] (Figure S2); with an increasing amount of DNA, the standard deviation of the peaks increases linearly. This makes sense physically, since when more DAPI labeling takes place, the variability associated with both labeling and quantifying cell-associated DAPI would also increase. These parameters enabled us to convert simulated DNA distributions to simulated *measured* DNA distributions by using Gaussian blurring. Due to the potential for variation in measured fluorescence values between flow cytometry runs performed on different days, the calibration was repeated on each day.

2. Simulation initiation

2.1. Identification of exponential growth

In order to automatically predict DNA distributions given an experimental growth curve, we require a method for identifying the exponential (and post-exponential) phases of growth for calculating instantaneous growth rates across the different phases of growth to which our model applies. To this end, each experimental growth curve (OD vs. time) was fit to a B-spline function, including a user-defined smoothing parameter that may be tuned to minimize sensitivity to experimental noise while generating a fit that is suitable given the temporal spacing between subsequent OD measurements[9]. This B-spline function (OD = f(t)) was then analyzed computationally to define the exponential phase of growth (the first time point, or time interval,

over which $\left|\frac{d^2(\ln(f(t)))}{dt^2}\right|$ is minimized). Given the typical noise associated with OD measurements, particularly at low OD values, we manually excluded earlier time points (low OD values) from this consideration, and the temporal interval identified as representing exponential growth was also confirmed manually (Figure S3). This methodology leads to the definition of exponential growth as representing a temporal window over which the minimization condition is satisfied (thus, any early inflection point, which also satisfies the minimization condition, is not identified as the window of exponential growth).

2.2. Populating the simulation assuming Malthusian growth

To initiate the simulation, we assumed the Malthusian growth model of unrestricted growth. For the Malthusian growth law to be valid however, it requires the population to be in balanced growth (also called steady state). By growing cells in ideal conditions, i.e. with an abundance of space, nutrients, and aeration, it is believed that the cells are not experiencing any growth limiting factors, and thus each individual in the population is growing at its maximal rate[6, 10]. This translates to a population of cells that are in the same state, albeit asynchronous vis-à-vis the cell cycle. This theory has been used to mathematically describe bacterial growth through each stage of the cell cycle, such that given the following probability density function describing the age distributions of an asynchronous bacterial population growing exponentially:

$$n(a) = 2 \cdot \ln(2) e^{-a\ln(2)}$$

the number of cells within the population inhabiting each cell cycle state can be deduced[3, 11]. In this expression, $n(a)$ is the probability for a cell to be at age a, where $0 \leq a \leq 1$, such that $a = 0$ would correspond to a newly divided cell, and $a = 1$ is the age at which a cell divides[11].

If indeed the balanced growth theory holds, it follows that given an exponentially growing population, regardless of when a sample is measured within a window of exponential growth, the state of the population (distribution of cells inhabiting each cell cycle state) should be unchanging. Thus, for each of the four different growth conditions (cells grown in LB or M9 media and shaken at either 230RPM or 23RPM), we used the method discussed in section 2.1 to identify the exponential phase of growth (Figure S4). Figure S5 depicts the experimentally measured DNA distributions observed across the identified window of exponential growth. Although some variation across this period was observed, overall the DNA distributions were relatively unchanging, which supports our use of the Malthusian growth condition/assumption to initiate the simulator. Note that because the model does not include a mechanism for cell death, any periods in which decreasing OD values were experimentally observed were not considered in this analysis.

2.3. Cell volume and instantaneous volumetric calculations

To initiate the simulation, we start with a single cell containing a single chromosome that has not started replication, and with volume $V_a = \frac{V_i}{2}$, and then we advance the model until it reaches

5000 cells. Under the assumption of exponential growth, the first step of the HMG simulation involves growing a population assuming the Malthusian growth model for each cell, where growth rate (μ) is extracted from the linear section of growth (see section 2.1 above):

$$V_{a+dt} = V_a(1 + (\mu * dt)) \quad (1)$$

and dt is a discrete time interval.

For non-exponential growth, the HMG simulation uses the injection growth method to track true volumetric changes in a population (see main manuscript). Therefore, it is necessary to convert the OD growth curve to its total cell volume equivalent. To perform this task, we use a constant relationship reported by Volkmer *et al.*, where regardless of the growth rate, total cell volume of a population of bacterial cell is $3.6\ \mu L \cdot OD^{-1} \cdot mL^{-1}$ [12]. Because the model requires the change of volume in time, we use the following equation to deduce the change in volume from OD growth curves:

$$Population\ V_{a+dt} = 3.6 * OD(a + dt) * 10^3 - 3.6 * OD(a) * 10^3 \quad (2)$$

where $OD(a)$ is the optical density at time a, and $Population\ V_{a+dt}$ is expressed in $fL \cdot mL^{-1}$. The reason we convert from μL to fL is because V_i is expressed in μm^3 and $\mu m^3 \equiv fL$. Furthermore, due to computational limitations, it is only possible to reach a fraction of the measured total population volume by our simulation (V_0). Thus it is necessary to normalize the total cell volume calculated from the OD to the total volume of the simulation at the start of the simulation, where:

$$Normalization\ constant = \frac{Population\ V_0}{Simulation\ Population\ V_0} \quad (3)$$

Thus to grow an individual cell using the injection method, we use the following:

$$V_{a+dt} = \frac{(Population\ V_{a+dt}/Normalization\ constant)}{Cell\ Concentration(a)} \quad (4)$$

Where $Cell\ Concentration(a)$ is the number of cells in the simulation at time a.

3. Development of *recA* mutant cell cycle model

3.1. WT RecA function

Because chromosome replication is, by nature, error prone, and because cells are constantly exposed to stresses that cause DNA damage such as oxidative stress and UV radiation, bacterial cells possess different mechanisms to manage damage to their DNA, including excision repair, mismatch repair, and the SOS response system[13]. For repair involving templating by homologous DNA sequences, the RecA protein plays a central role in matching the damaged DNA to its complementary sequence [13-15]. RecA expression not only alleviates repression of DNA repair protein expression through a co-proteolytic function that cleaves the repressor LexA, but upon coating single stranded DNA, RecA catalyzes the pairing to a complementary strand[13, 15-16]. Since elucidating the detailed mechanisms of bacterial DNA repair remains an area of active investigation, we represented the process of DNA repair using a simplified mechanism that captures the essence of the current consensus understanding[14, 17]. In this consensus model, when a replicating fork encounters any type of DNA damage, it is arrested at the site of the damage[14-15, 17]. As a consequence of this stalling, a double stranded break is created that leads to one of the replicating strands containing an exposed double stranded break[14-15, 17]. Exposure of the double-stranded break leads to recruitment of the RecBCD enzyme, whose helicase and nuclease functions degrade the DNA past the source of the damage, until the nuclease reaches a Chi sequence[14-15, 17]. RecA is then recruited, binds to a single strand of DNA, and through homologous recombination, initiates a faithful repair of the error, resulting in a fully restored version of the chromosome[14-15, 17].

3.2. Effect of the *recA* null mutant on chromosome replication

The gene encoding RecA is often mutated (e.g., *recA1*) in laboratory strains used for the purpose of genetic engineering, since this mutation impairs the ability of the cell to perform recombination of inserted plasmid genetic material and thus increases the stability of inserted plasmids. Unsurprisingly, such bacterial strains are particularly sensitive to any type of DNA damage, and during normal laboratory growth, as much as half of the cells in a population are inviable[8], mostly due to the occurrence of anucleate cells[8, 17]. A few key observations indicate the consequences of *recA* mutation. Early studies showed that *recA1* cells contained aberrant chromosome copy numbers[8]. In a population of WT cells, most individuals would have 1, 2, 4, 8 etc… (2^n) chromosomes per cell, since a new round of replication initiates simultaneously at all OriC locations in a cell once the cell reaches critical mass. On the other hand, *recA1* populations contain cells with 1, 2, 3, 4, 5 etc… copies per cell[8].

Substantial work has been devoted to elucidating how the *recA1* mutation leads to aberrant chromosome replication. Timing and coordination of initiation was shown to not be affected by this mutation[8, 18]. However, *recA1* bacteria contain a higher number of free floating nucleotides than do their WT counterparts, and this phenotype is exacerbated by faster growth rates[19]. *In vitro* experiments show that RecA inhibits the nuclease activity of RecBCD, and knock out experiments for this protein are more lethal to the cell than are *recA* knockouts[20]. If, for any reason, the replication fork is arrested long enough, then the arrest leads to double-stranded breaks, and RecBCD is recruited and degrades the replicating strand[13-14, 17, 21-22]. Furthermore, there seems to be a RecA-independent damage avoidance mechanism which involves

suppression and removal of the damaged strand by RecBCD during replication,[23] which leads to the collapse of the replicating strand, and the chromosome returns to its original form[14, 23]. Other evidence suggests that *recA1* mutant bacteria also experience whole chromosome degradation[8]. Finally, the lack of functional RecA may affect the synchronous segregation of chromosomes at division, such that a given complement of chromosomes is divided less evenly between daughter cells in *recA* mutant strains[18].

Altogether, these observations led us to summarize these *recA*-associated mechanisms for inclusion in our model in the following way. If the replicating strand encounters DNA damage (which it is unable to repair due to the lack of RecA), then a double stranded break occurs. If the break occurs upstream of the replication fork, then RecBCD degrades the replicating strand until it reaches the end of the other replication fork[21], which leads to restoration of the replicating chromosome to its original state. If the double-stranded break occurs downstream of the replication fork, then RecBCD degrades the whole chromosome. Finally, in *recA mutant* cells, a parameter was included which quantified the degree of asymmetric segregation of chromosomes into daughter cells at cell division.

To investigate how we might best describe the effects of *recA* mutation in our model, given this diversity of mechanisms and a lack of quantitative measurements for these processes, we first formulated a series of models that did, or did not include a combination of these three mechanisms – whole chromosome degradation, replicating strand collapse/degradation, and asymmetric segregation. By attempting to fit each of these models to our experimental data, by parametric optimization (data not shown), we observed that the best fits came from a model that included whole chromosome degradation and replicating strand collapse/degradation, but did not include asymmetric segregation, which is a formulation that is consistent with some recent reports[8]. Therefore, we formulated a concise mechanism to capture the impact of *recA* mutation in our model, which is summarized in Figure S6, and used this formulation for all modeling work reported here. In further support of this choice, we observed via a parametric sensitivity analysis (Figure S7) that the asymmetric segregation parameter had the lowest impact on the extent to which our model could match the experimentally measured DNA distribution associated with an exponentially growing population[3].

4. Model implementation

For a summary of the model algorithm, please see Figure 2. Here we describe a few such steps in detail to clarify how this algorithm was implemented in our HMG simulator.

4.1. Noise parameters and cell cycle parameter stochasticity

Each noise parameter, which we will generically call NoiseParam, (and which includes V_i Noise, V_a Noise, C Noise and D Noise, expressed in %) is used in the following way: for Parameter X, with mean value M, noise is introduced by selecting a random number from a Gaussian distribution with mean M and standard deviation σ, where σ = M*NoiseParam/100.0. The selected random number is then used as a noisy version of Parameter X.

4.2. Stochastic opening of replication forks

The "Chance of Initiation" parameter describes the "degree of synchronous initiation at multiple origins"[3]. This can be described as a cumulative distribution function given a standard normal distribution (or in other terms the area under the bell curve). Following this previously developed method[3], we randomly select a number from a Gaussian distribution with mean 0.0 and standard deviation of 1.0, and we then check if this random number is larger or equal to 1.75. This means that at every time step, the probability of asynchronous initiation is equal to 0.0401 given a time step (dt) of 0.01 minutes, and using the following cumulative distribution of a standard normal distribution:

$$\phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2}\, dt \tag{5}$$

we calculate the asynchronicity amongst replication forks (approximately 4%)[3]:

$$Asynchronicity = (1.0 - \phi(Chance\ of\ Initiation)) * dt \tag{6}$$

4.3. DNA Damage

A subtlety to be noted is that in practice, to avoid the physically redundant scenario of inducing both whole chromosome degradation and replicating strand collapse, we implemented the two DNA degradation mechanisms in a mutually exclusive fashion. During chromosome replication, a random number is generated according to a Gaussian distribution with mean 0 and standard deviation of 1. This random number is compared against the parameter "Chance of DNA damage" to determine whether the currently replicating chromosome experiences either of the two types of DNA damage. If damage does occur, a second random number is generated from a Gaussian distribution with mean of 0 and standard deviation of 1, and this random number is compared against the parameter "Ratio of DNA damage", in order to choose whether the chromosome experiences replicating strand degradation or, instead, whole chromosome degradation (where if random number $<$ Ratio of DNA damage then a single replication forks collapses, and if random number $\geq$ Ratio of DNA damage then the whole chromosome is degraded). Note that our parameters were defined in such a fashion (as random Gaussian numbers) to enable comparison with parameters that were previously measured by experiment [3]. To convert such a parameter to the equivalent number defined as a random number sampled from a uniform distribution across an interval from 0 to 1, one simply calculates:

$$\begin{aligned} &Chance\ of\ Chromsome\ Degradation \\ &\quad = \left(1.0 - \phi(Chance\ of\ DNA\ Damage)\right) \\ &\quad * \left(1.0 - \phi(Ratio\ of\ DNA\ Damage)\right) * dt \end{aligned} \tag{7}$$

Similarly, to determine the chance that a replication fork collapses, one calculates:

$$Chance\ of\ Replication\ Fork\ Collapse = \left(1.0 - \phi(Chance\ of\ DNA\ Damage)\right) * \left(\phi(Ratio\ of\ DNA\ Damage)\right) * dt \quad (8)$$

Where ϕ is the cumulative distribution of a Gaussian distribution with mean of 0 and σ of 1 (as described in equation 5), and dt is the time step per minute.

4.4. C and D periods

Under exponential growth, simple functions have been derived from experiments to describe replication time as a function of growth rate ($C(\mu) = C_1 e^{-C_2\mu} + C_3$), and to describe segregation time as a function of growth rate ($D(\mu) = D_1 e^{-D_2\mu} + D_3$), as previously reported[3]. Here, we hypothesized that a similar functional form could be used to describe growth outside of exponential growth, and to that end, we subjected these expressions to parametric optimization (see section 5.2 below).

5 Optimization strategy

5.1. Objective Function and definition of Similarity Score

To compute the difference between measured DNA distributions and simulated DNA distributions, we used a previously reported functional form to define an objective function for our optimization[11], and we describe this value as a "similarity score":

$$s = \sqrt{\sum_{i=1}^{m} \frac{(\sqrt{y_i} - \sqrt{N_i})^2}{m-1}} \quad (9)$$

Note that because we use this definition as an objective function, better fits between experiments and simulations result in lower (smaller) scores). In this expression, for each bin (i) of a flow cytometry histogram, y_i is the normalized original number of cells (i.e., the frequency represented by this bin), N_i is the normalized calculated number of cells (again, a frequency), and m is the total number of bins. For a graphical illustration of the scoring, please refer to Figures S9, S10, S11, and S12.

5.2 Optimization using a genetic algorithm (GA)

To optimize parameters for the equations describing replication time as a function of growth rate ($C(\mu) = C_1 e^{-C_2\mu} + C_3$), and describing segregation time as a function of growth rate ($D(\mu) = D_1 e^{-D_2\mu} + D_3$), and the parameters 'Chance of DNA damage' and 'Ratio of DNA damage', we

used a genetic algorithm (GA) (see Materials and Methods in the main manuscript). To effectively search the large potential parameter space, we generated Sobol sequences[24] with the following boundary conditions:

$$C(\mu) > 30 \text{ minutes}$$
$$D(\mu) > 15 \text{ minutes}$$
$$C(\mu) > D(\mu)$$
$$Chance\ of\ DNA\ damage < 3.75$$

Using data collected from TOP10 cells grown in LB at 23RPM and 230RPM (as presented in Figure 3), optimization of the model yielded the following results (Figure S8):

$$C(\mu) = 250.02 \cdot e^{-304.41\mu} + 60.97$$
$$D(\mu) = 100.59 \cdot e^{-137.07\mu} + 19.93$$
$$Chance\ of\ DNA\ damage = 3.907$$
$$Ratio\ of\ DNA\ damage = -0.777$$

Where C and D are expressed in minutes and growth rate is expressed in min^{-1}. Given the definitions described above, the 'Chance of DNA damage' parameter obtained through this fitting means that any chromosome would have a 0.467% chance of experiencing DNA damage during replication (as described in section 4.2). Similarly, this 'Ratio of DNA damage' means that a replicating chromosome has 0.365% chance of experiencing replicating strand degradation and a 0.102% chance of experiencing whole chromosome degradation (as described in section 4.3) (these mutually exclusive possibilities sum to a 0.467% chance that either occurs).

II. SUPPLEMENTAL FIGURES

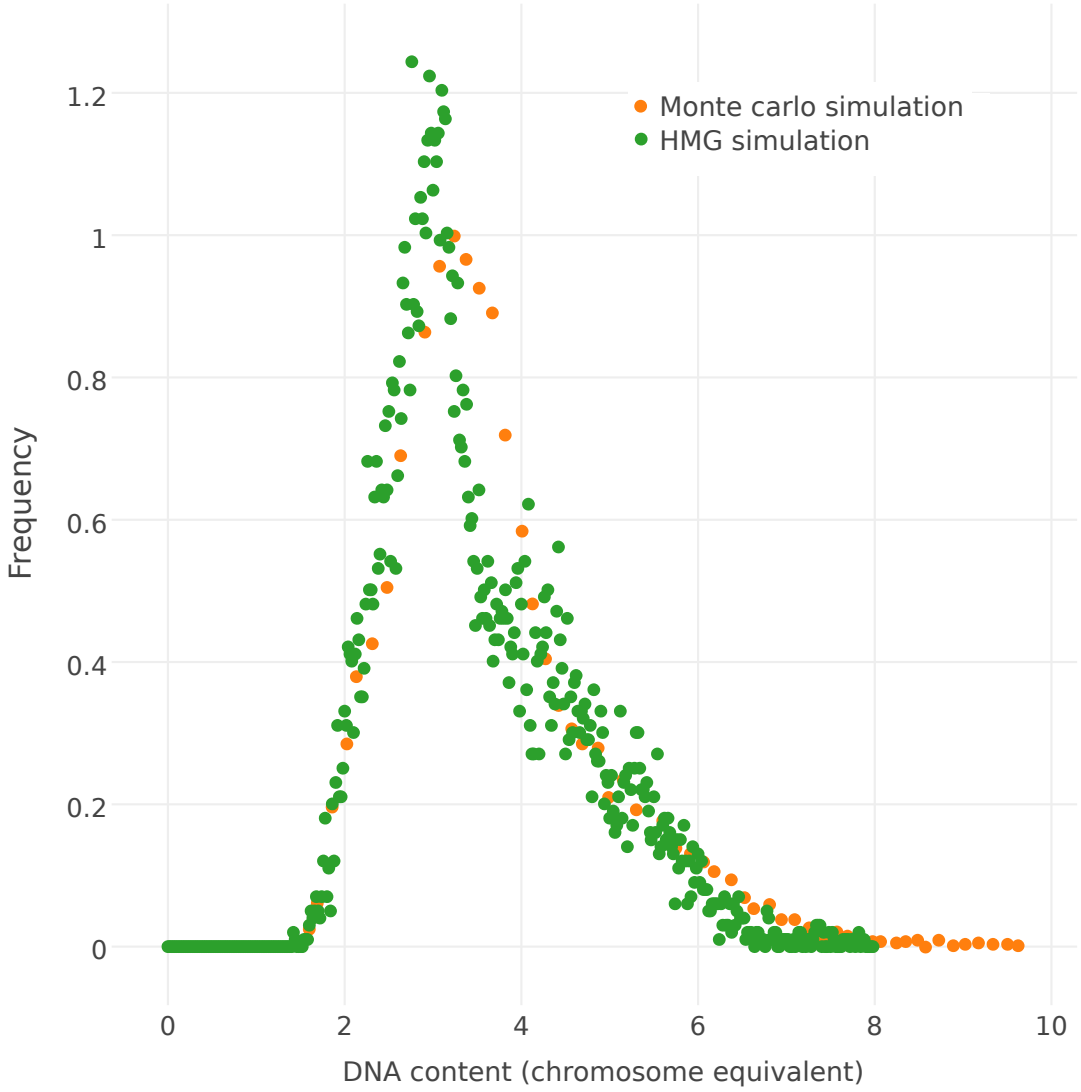

Figure S1. Comparison of the HMG simulator to a Monte Carlo simulation of the canonical CH model during exponential growth. The HMG simulator used a model parameterized using growth rate, parameter noise, and C & D parameters previously reported[3], and these results are plotted alongside the Monte Carlo Simulation results reproduced from that same study. The HMG simulator generates comparable results to this canonical model, when specifically examining the case of exponential growth, hence validating our injection growth formulation.

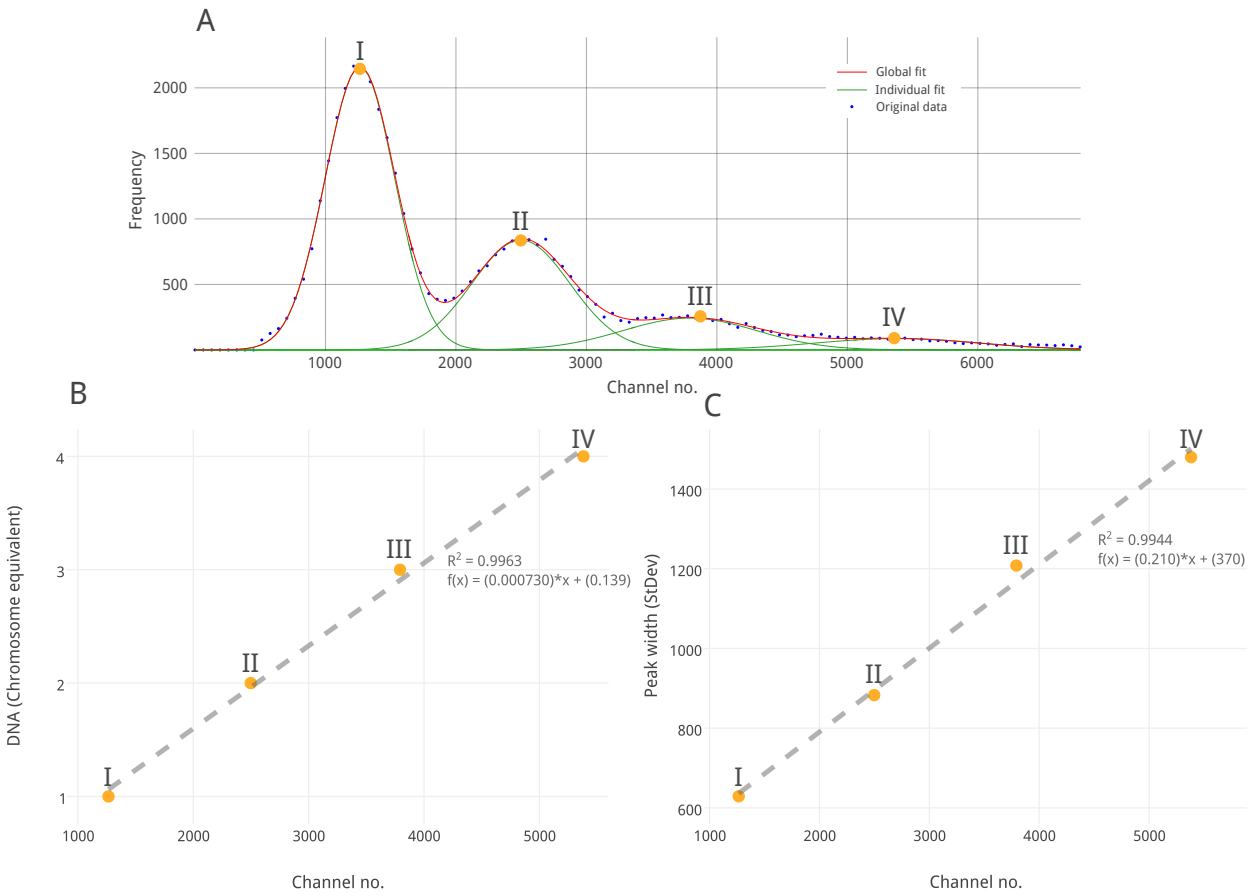

Figure S2. Computational deconvolution of DNA distributions. This figure illustrates the use of the MATLAB package ipf11 to identify a sum of Gaussian distributions that is maximally consistent with the DNA distributions measured by flow cytometry of DAPI-labeled cells. The upper panel (A) shows the original flow cytometry data (blue dotted line) plotted against a global fit (red) that is the sum of the fits of individual Gaussian distributions (green). Because *recA* mutants are known to have aberrant chromosome copy numbers[8], the orange circles associated with the peaks represent the following chromosome copy numbers (from left to right): 1, 2, 3, 4. The sample shown corresponds to cells grown in LB and shaken at 230 RPM, and samples were collected for DNA distribution analysis after 24 hours of growth post-inoculation. First order polynomials were fit to relate DNA content (B), and peak width (standard deviation) (C), to the channel number. Notably, peak width also increases linearly with channel number, which has been previously observed[5]. Peaks labels defined in panel A (I, II, III, IV) are indicated next to corresponding data points in panels B and C.

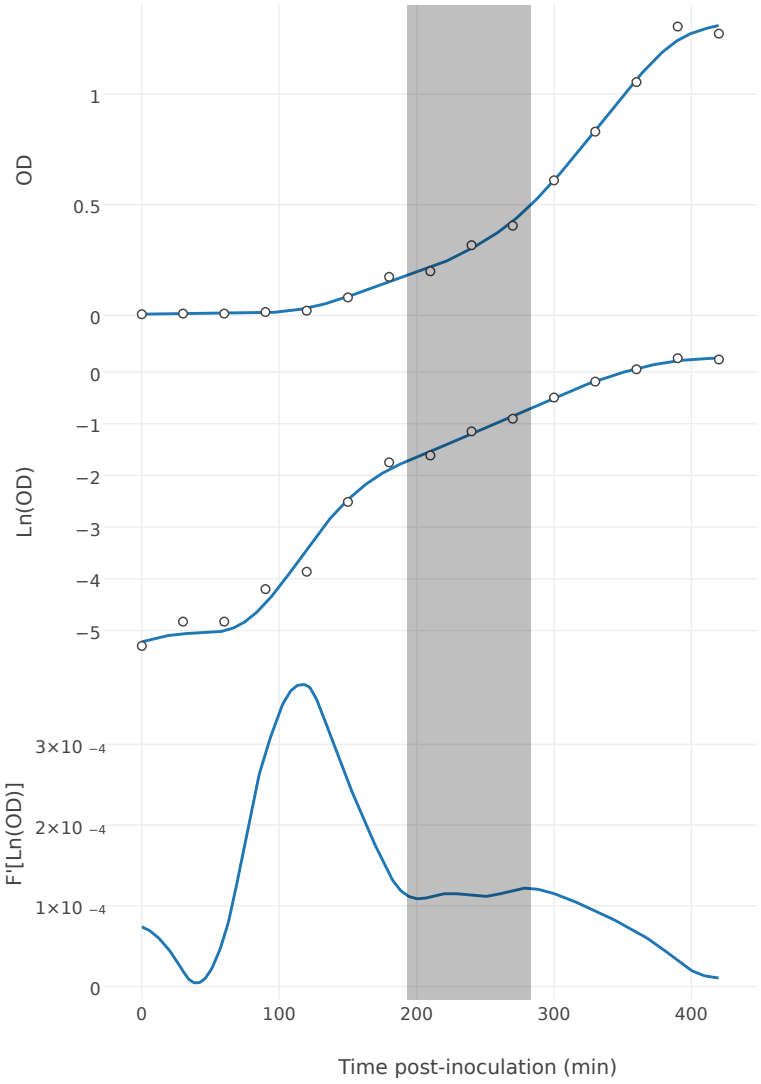

Figure S3. Identification of exponential growth regimes. This figures illustrates our method for identifying the exponential growth regime, in this case using cells grown in LB and shaken at 23RPM. The top panel shows the B-spline fit (blue line) to the original OD measurements (open circles). Given this fit, the middle panel shows the natural log of both the original OD measurements and the fit. The bottom panel shows the first order differential of the natural log of the fit OD growth curve, which is used to identify the period of exponential growth that appears as a roughly horizontal line on this bottom panel. Although a naïve interpretation of these data would indicate that maximal growth occurs at around 100 minutes post-inoculation, we can clearly see from the lower panel that growth is stable between ~200-300 minutes post-inoculation. Because OD measurements are relatively noisy at low cell densities, we would argue that the identified window (shaded in gray) represents true exponential growth, it remains possible that our method favors the identification of prolonged steady growth phases over early rapid phases of growth.

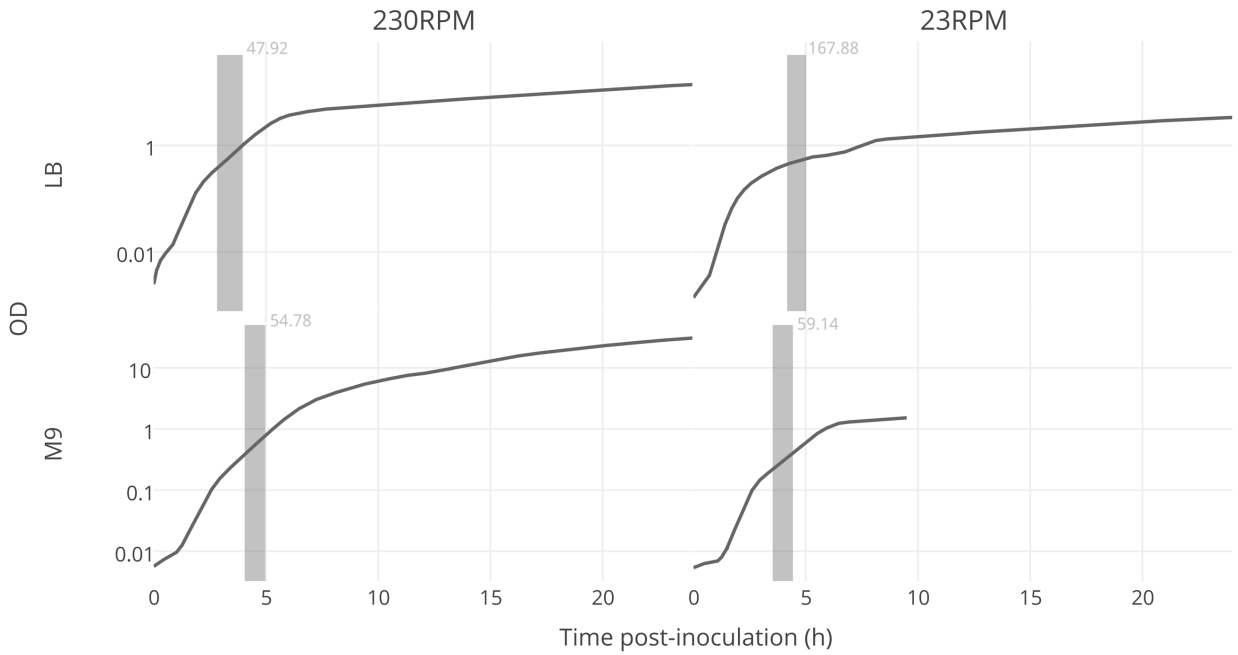

Figure S4. Growth curves. Shown here are raw OD measurements from the different growth conditions used throughout this study, using a combination of different carbon sources (LB, M9) and shaking rates (230RPM and 23RPM). The shaded areas are the identified exponential sections of the growth curves, as described in Figure S3, with the calculated doubling rates (in minutes) indicated next to each shaded region.

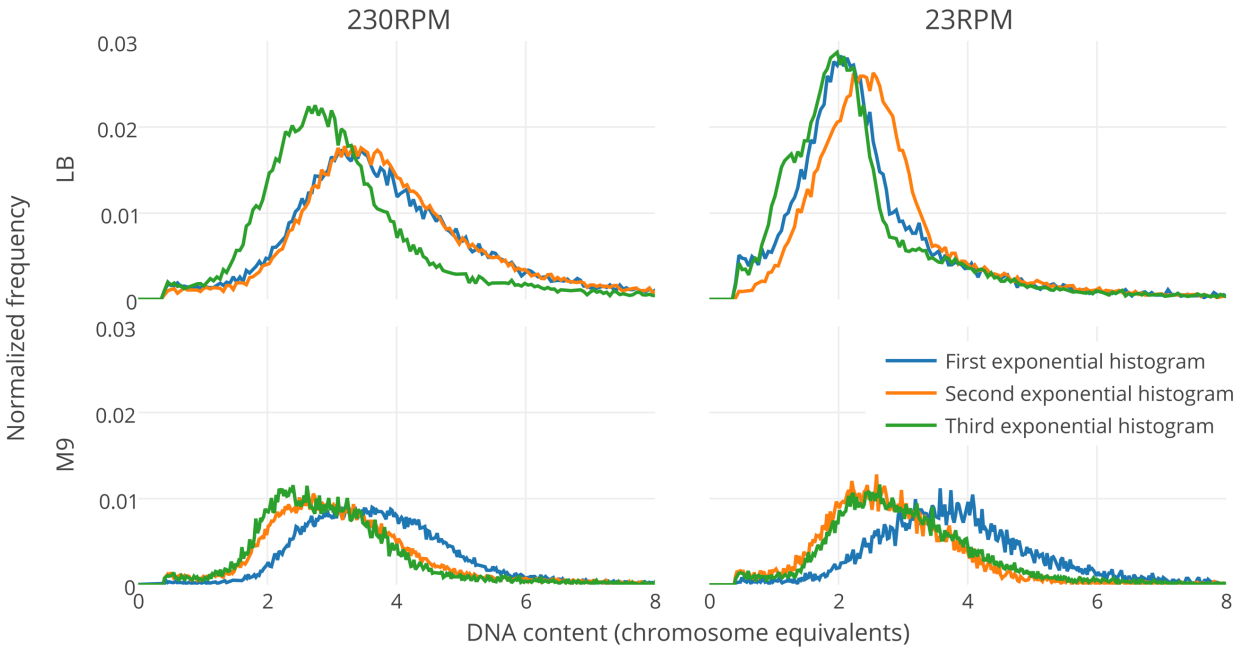

Figure S5. Measured DNA distributions for exponential growth regions identified in Figure S4. For each panel, the first, second and third DNA distributions (see key) correspond to the first, middle, and final time points within the identified window of exponential growth. Because cells are assumed to be in balanced growth during exponential growth, in theory the DNA distributions should be unchanging across these windows. These data indicate that the DNA distributions are indeed relatively unchanging, supporting the assumption of balanced growth within this window, and validating our strategy of initiating (inoculating) our simulations by assuming exponential balanced growth during these identified windows.

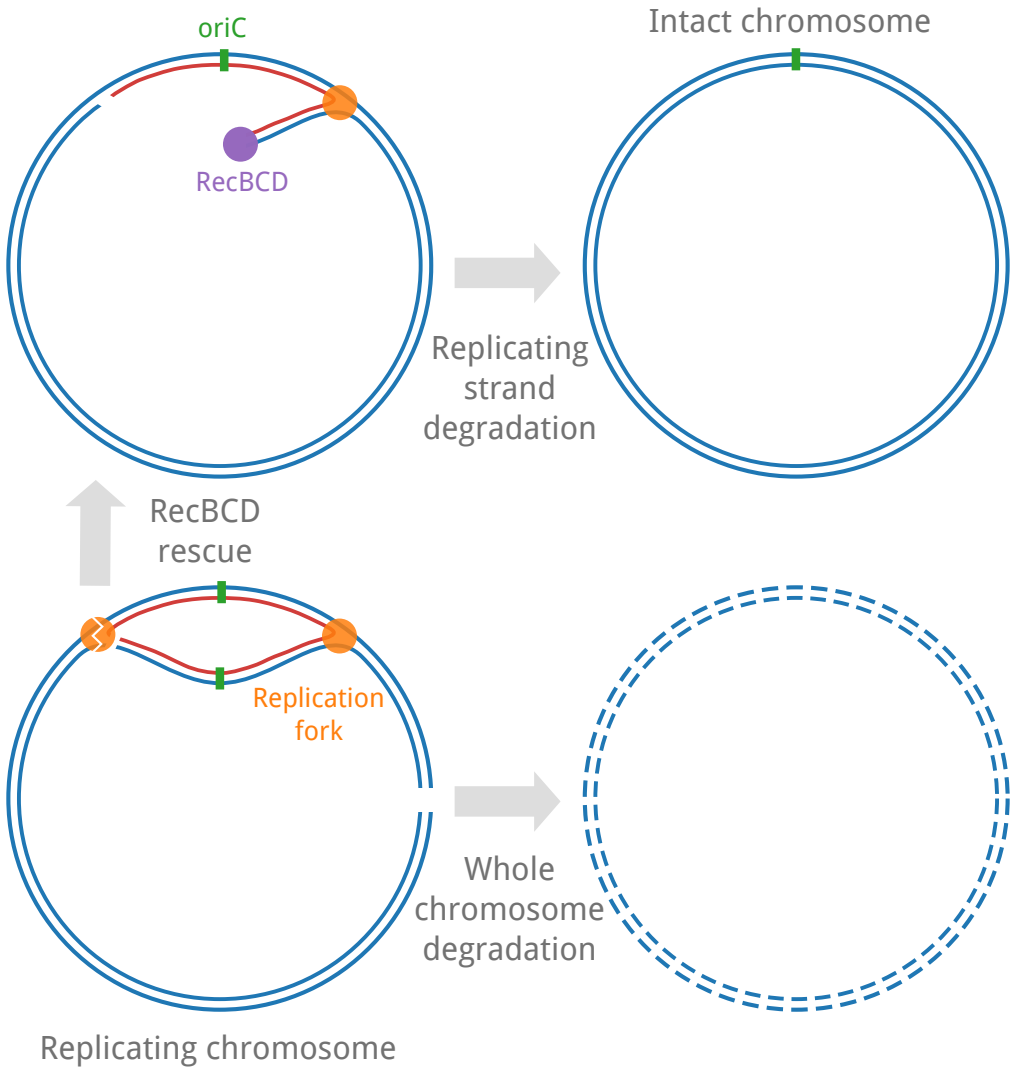

Figure S6. Summary of the simplified description of DNA double stranded damage responses used in our simulations to represent the consequence of *recA* mutation. Green bars are the OriC, orange circles are the replication fork bubbles, the purple circle is the RecBCD enzyme, each blue line is a single chromosome strand, and each red line represents a newly replicated DNA strand. As described in the text, our simulation captures two potential outcomes following DNA damage. In the first scenario, a double-stranded break occurs at the replication fork, which leads to the collapse of the replication fork (broken orange circle), which is subsequently rescued by RecBCD, returning the chromosome to its original form. In the second scenario, the double-stranded break occurs downstream of the replication fork, which instead leads to degradation of the whole chromosome.

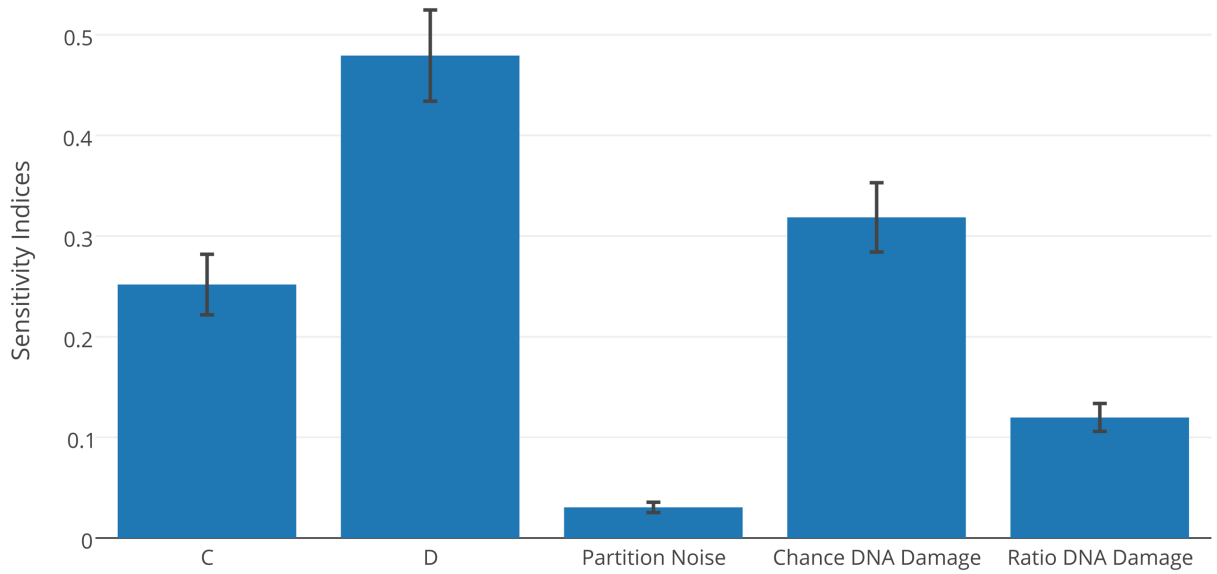

Figure S7. Total-order index sensitivity analysis for the optimized parameters. These sensitivities were calculated based upon an initial ensemble (Sobol sequence) of 12,000 parameter sets. The error bars represent the 90% confidence intervals. For this analysis, a base case was generated by simulating exponential growth, using the following parameters: C: 40.0 min, D: 20.0 min, Partition Noise: 0.24, Chance of DNA Damage: 3.034, Ratio of DNA damage: -1.652. Therefore, the sensitivity indices plotted here represent the degree to which changing any model parameter shifts the simulated DNA distribution (under exponential growth) compared to the base case.

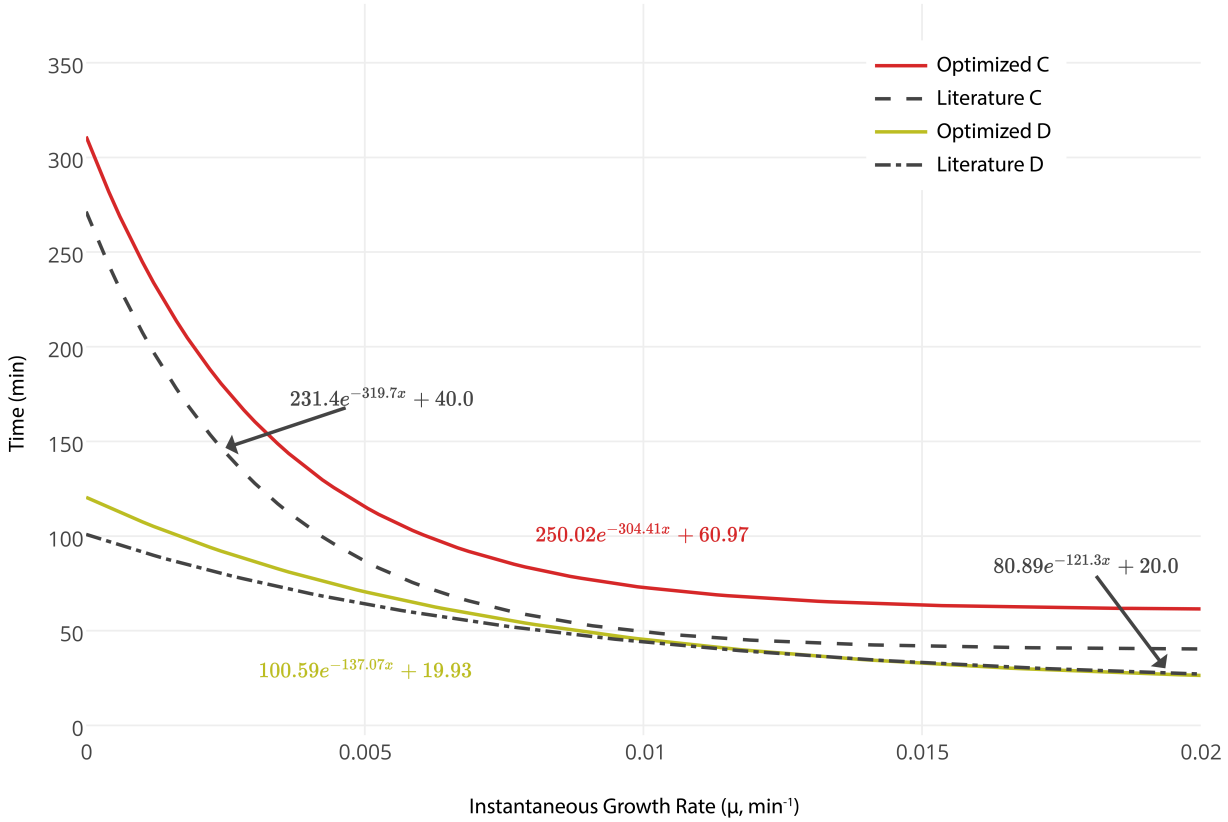

Figure S8: Optimized C and D functions used in the HMG simulations compared with functions previously reported in the literature based upon analysis of exponential growth[3].

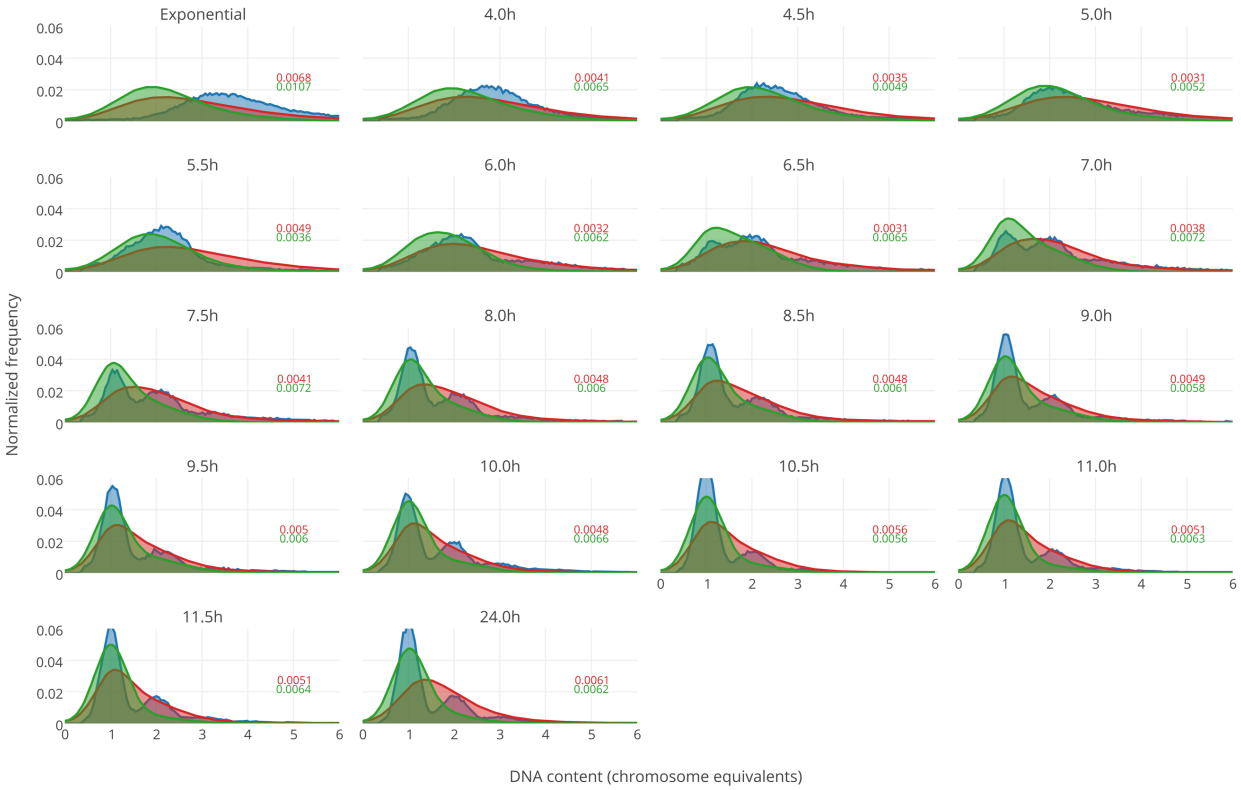

Figure S9. DNA distributions for cells growth in LB, shaken at 230RPM. These histograms correspond to the data plotted as heat maps in Figure 3. In each panel, the blue histogram represents the measured DNA distribution, the green histogram represents the distribution generated by simulation using the model of [3], and the red histogram represents the distribution generated by simulation with the new model including optimized parameterization. Each panel is also annotated with the corresponding similarity scores comparing the measured DNA distribution with the distributions simulated via each approach.

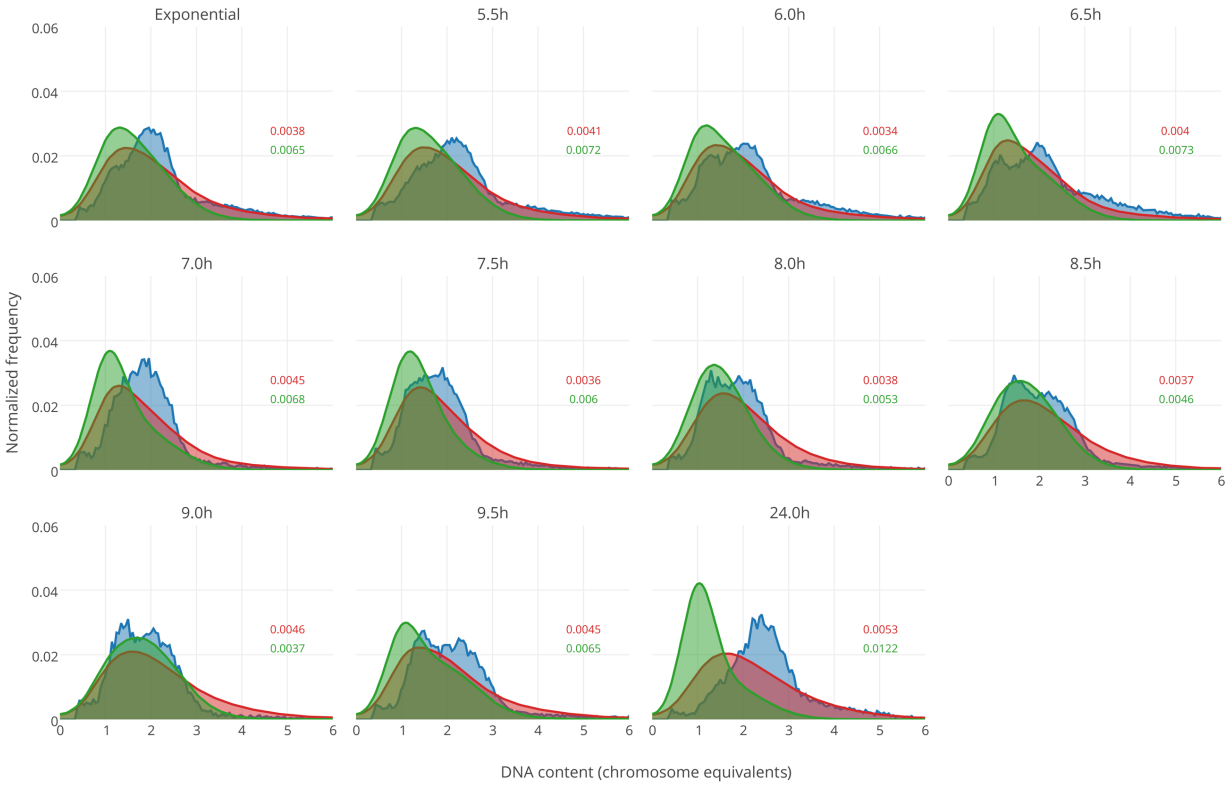

Figure S10. DNA distributions for cells growth in LB, shaken at 23RPM. These histograms correspond to the data plotted as heat maps in Figure 3. In each panel, the blue histogram represents the measured DNA distribution, the green histogram represents the distribution generated by simulation using the model of [3], and the red histogram represents the distribution generated by simulation with the new model including optimized parameterization. Each panel is also annotated with the corresponding similarity scores comparing the measured DNA distribution with the distributions simulated via each approach.

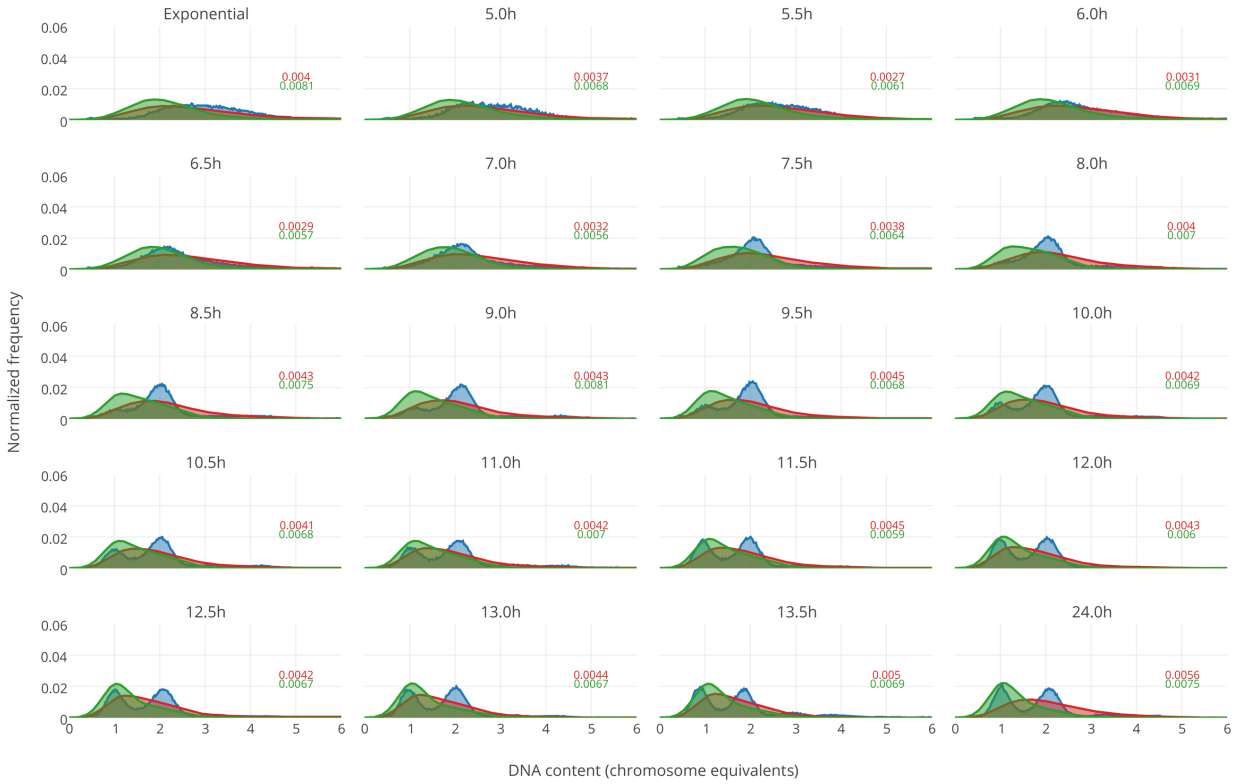

Figure S11. DNA distributions for cells growth in M9, shaken at 230RPM. These histograms correspond to the data plotted as heat maps in Figure 3. In each panel, the blue histogram represents the measured DNA distribution, the green histogram represents the distribution generated by simulation using the model of [3], and the red histogram represents the distribution generated by simulation with the new model including optimized parameterization. Each panel is also annotated with the corresponding similarity scores comparing the measured DNA distribution with the distributions simulated via each approach.

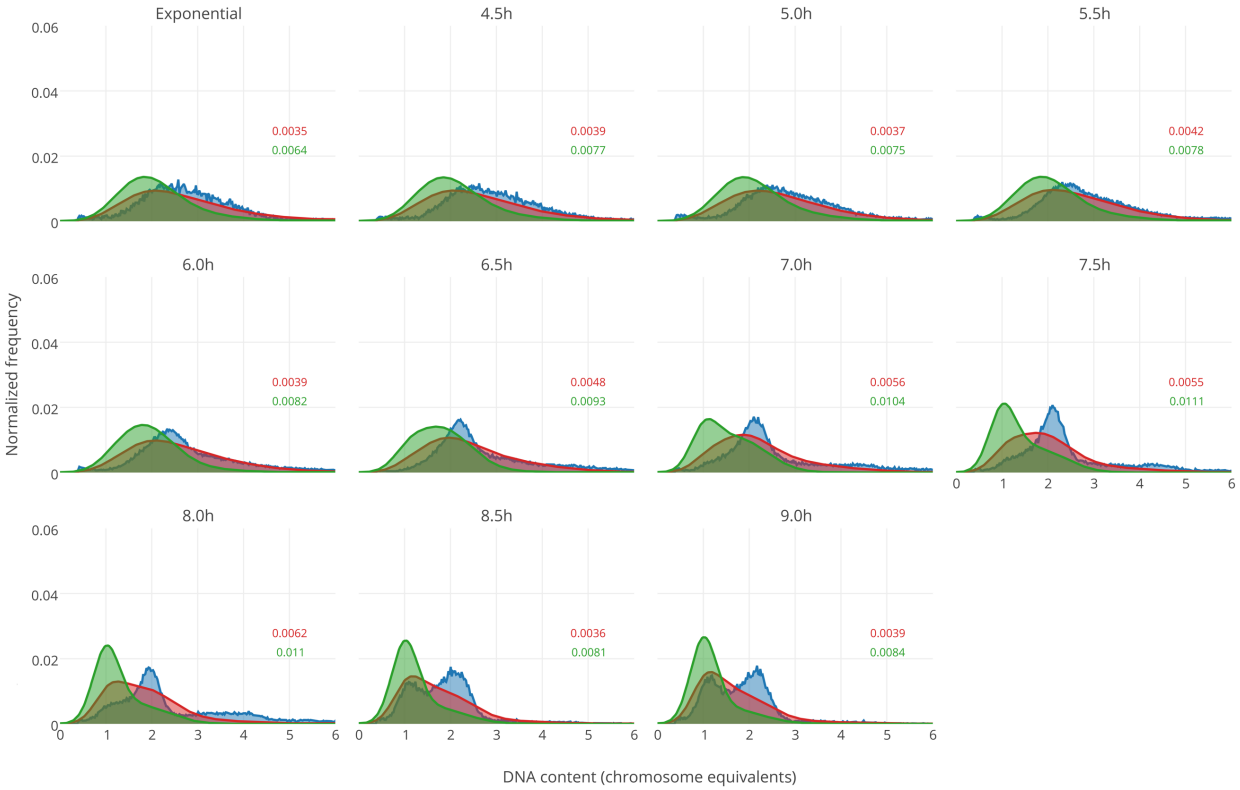

Figure S12. DNA distributions for cells growth in M9, shaken at 23RPM. These histograms correspond to the data plotted as heat maps in Figure 3. In each panel, the blue histogram represents the measured DNA distribution, the green histogram represents the distribution generated by simulation using the model of [3], and the red histogram represents the distribution generated by simulation with the new model including optimized parameterization. Each panel is also annotated with the corresponding similarity scores comparing the measured DNA distribution with the distributions simulated via each approach.

III. REFERENCES CITED IN SUPPORTING INFORMATION

1. Wallden, M.; Fange, D.; Baltekin, z.; Elf, J., Fluctuations in growth rates determine the generation time and size distributions of E. coli cells. *arXiv preprint arXiv:1504.03145* **2015**.
2. Browning, S. T.; Castellanos, M.; Shuler, M. L., Robust control of initiation of prokaryotic chromosome replication: essential considerations for a minimal cell. *Biotechnol Bioeng* **2004,** *88* (5), 575-84.
3. Keasling, J. D.; Kuo, H.; Vahanian, G., A Monte Carlo simulation of the Escherichia coli cell cycle. *J Theor Biol* **1995,** *176* (3), 411-30.
4. Michelsen, O.; Teixeira de Mattos, M. J.; Jensen, P. R.; Hansen, F. G., Precise determinations of C and D periods by flow cytometry in Escherichia coli K-12 and B/r. *Microbiology* **2003,** *149* (Pt 4), 1001-10.
5. Stokke, C.; Flatten, I.; Skarstad, K., An easy-to-use simulation program demonstrates variations in bacterial cell cycle parameters depending on medium and temperature. *PLoS One* **2012,** *7* (2), e30981.
6. Abner, K.; Aaviksaar, T.; Adamberg, K.; Vilu, R., Single-cell model of prokaryotic cell cycle. *J Theor Biol* **2014,** *341*, 78-87.
7. Bryant, F. R., Construction of a recombinase-deficient mutant recA protein that retains single-stranded DNA-dependent ATPase activity. *J Biol Chem* **1988,** *263* (18), 8716-23.
8. Skarstad, K., Degradation of Individual Chromosome in recA Mutants of Escherichia coli. **1993**.
9. De Boor, C.; De Boor, C.; Mathématicien, E.-U.; De Boor, C.; De Boor, C., *A practical guide to splines*. Springer-Verlag New York: 1978; Vol. 27.
10. Buchanan RL, W. R. D. W., When is simple good enough: a comparison of the Gompertz, Baranyi, and three-phase linear models for fitting bacterial growth curves. *Food Microbiol* **1997,** *14* (4), 313-326.
11. Skarstad, K.; Steen, H. A. R. A. L. D. B.; Boye, E., Escherichia coli DNA distributions measured by flow cytometry and compared with theoretical computer simulations. *Journal of bacteriology* **1985,** *163* (2), 661-668.
12. Volkmer, B.; Heinemann, M., Condition-dependent cell volume and concentration of Escherichia coli to facilitate data conversion for systems biology modeling. *PLoS One* **2011,** *6* (7), e23126.
13. Courcelle, J.; Hanawalt, P. C., RecA-dependent recovery of arrested DNA replication forks. *Annu Rev Genet* **2003,** *37*, 611-46.
14. Kuzminov, A., Collapse and repair of replication forks in Escherichia coli. *Mol Microbiol* **1995,** *16* (3), 373-84.
15. Michel, B., Replication fork arrest and DNA recombination. *Trends Biochem Sci* **2000,** *25* (4), 173-8.
16. Little, J. W.; Edmiston, S. H.; Pacelli, L. Z.; Mount, D. W., Cleavage of the Escherichia coli lexA protein by the recA protease. *Proc Natl Acad Sci U S A* **1980,** *77* (6), 3225-9.
17. Kuzminov, A., Instability of inhibited replication forks in E. coli. *Bioessays* **1995,** *17* (8), 733-41.
18. Zyskind, J. W.; Svitil, A. L.; Stine, W. B.; Biery, M. C.; Smith, D. W., RecA protein of Escherichia coli and chromosome partitioning. *Mol Microbiol* **1992,** *6* (17), 2525-37.
19. Horii, DEGRADATION OF THE DNA OF ESCHERICHIA COLI K12 REC-(JC1569b) AFTER IRRADIATION WITH ULTRAVIOLET LIGHT. **1968**.
20. Kuzminov, A.; Stahl, F. W., Stability of linear DNA in recA mutant Escherichia coli cells reflects ongoing chromosomal DNA degradation. *J Bacteriol* **1997,** *179* (3), 880-8.
21. Kuzminov, A., Recombinational repair of DNA damage in Escherichia coli and bacteriophage lambda. *Microbiol Mol Biol Rev* **1999,** *63* (4), 751-813, table of contents.

22. Courcelle, J.; Donaldson, J. R.; Chow, K. H.; Courcelle, C. T., DNA damage-induced replication fork regression and processing in Escherichia coli. *Science* **2003,** *299* (5609), 1064-7.
23. Miranda, A.; Kuzminov, A., Chromosomal lesion suppression and removal in Escherichia coli via linear DNA degradation. *Genetics* **2003,** *163* (4), 1255-71.
24. Sobol', I. y. M., On the distribution of points in a cube and the approximate evaluation of integrals. *Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki* **1967,** *7* (4), 784-802.