DEVELOPMENTAL BIOLOGY

Correction for "Progressive lengthening of 3′ untranslated regions of mRNAs by alternative polyadenylation during mouse embryonic development," by Zhe Ji, Ju Youn Lee, Zhenhua Pan, Bingjun Jiang, and Bin Tian, which appeared in issue 17, April 28, 2009, of *Proc Natl Acad Sci USA* (106:7028–7033; first published April 16, 2009; 10.1073/pnas.0900028106).

The authors note that due to a printer's error, Fig. 1*B* appeared incorrectly on page 7028. The corrected figure and its legend appear below.

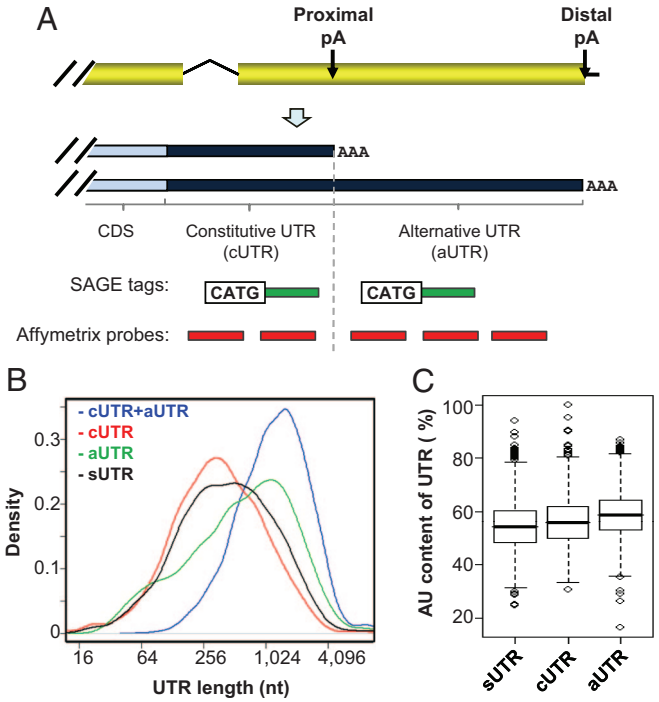

Fig. 1. Alternative polyadenylation (APA) leading to alternative 3′ UTRs. (*A*) Schematics of APA in the 3′-most exon and using SAGE tags and Affymetrix GeneChip probes to examine APA. Two transcript variants resulting from proximal and distal poly(A) sites are shown. The dotted vertical line separates 2 UTR regions, i.e., constitutive UTR (cUTR) and alternative UTR (aUTR). CDS, coding sequence; pA, poly(A) site; AAA, poly(A) tail. (*B*) Distribution of length for different UTR groups. The full 3′ UTRs for genes with APA are shown by cUTR + aUTR [4,139 genes, median = 1,288 nucleotide (nt)]. Constitutive and alternative regions are represented by cUTR (median = 358 nt) and aUTR (median = 685 nt), respectively. sUTRs are 3′ UTRs not affected by APA (7,242 genes, median = 439 nt). (*C*) AU content for different UTR sequences.

www.pnas.org/cgi/doi/10.1073/pnas.0904454106

CELL BIOLOGY, COMPUTER SCIENCE

Correction for "Stochastic hybrid modeling of DNA replication across a complete genome," by J. Lygeros, K. Koutroumpas, S. Dimopoulos, I. Legouras, P. Kouretas, C. Heichinger, P. Nurse, and Z. Lygerou, which appeared in issue 34, August 26, 2008, of *Proc Natl Acad Sci USA* (105:12295–12300; first published August 19, 2008; 10.1073/pnas.0805549105).

The authors wish to note the following: "We wish to add direct references to a stochastic model of DNA replication previously applied to the *Xenopus laevis* early embryonic divisions. That model was applied to molecular combing experiments on cell-free extracts from *Xenopus laevis* embryos." The additional references appear below.

39. Herrick J, Jun S, Bechhoefer J, Bensimon A (2002) Kinetic model of DNA replication in eukaryotic organisms. *J Mol Biol* 320:741–750.
40. Jun S, Bechhoefer J (2005) Nucleation and growth in one dimension. II. Application to DNA replication kinetics. *Phys Rev E Stat Nonlin Soft Matter Phys* 71:011909.
41. Bechhoefer J, Marshall B (2007) How *Xenopus laevis* replicates DNA reliably even though its origins of replication are located and initiated stochastically. *Phys Rev Lett* 98:098105.

www.pnas.org/cgi/doi/10.1073/pnas.0905246106

MICROBIOLOGY

Correction for "Electrically conductive bacterial nanowires produced by *Shewanella oneidensis* strain MR-1 and other microorganisms," by Yuri A. Gorby, Svetlana Yanina, Jeffrey S. McLean, Kevin M. Rosso, Dianne Moyles, Alice Dohnalkova, Terry J. Beveridge, In Seop Chang, Byung Hong Kim, Kyung Shik Kim, David E. Culley, Samantha B. Reed, Margaret F. Romine, Daad A. Saffarini, Eric A. Hill, Liang Shi, Dwayne A. Elias, David W. Kennedy, Grigoriy Pinchuk, Kazuya Watanabe, Shun'ichi Ishii, Bruce Logan, Kenneth H. Nealson, and Jim K. Fredrickson, which appeared in issue 30, July 25, 2006, of *Proc Natl Acad Sci USA* (103:11358–11363; first published July 18, 2006; 10.1073/pnas.0604517103).

The authors note that in Table 1 on page 11360, the units for "Electrochemical activity," designated as mA, should have been designated as **μA**. This error does not affect the conclusions of the article. The corrected table appears below.

Table 1. Metal-reduction and electrochemical properties of *S. oneidensis* MR-1, GSPD and ΔMTRC/OMCA

Strain	Ferrihydrite-reduction-extractable Fe(II),* mM	Electrochemical activity, μA
MR-1	2.17 ± 0.19	68.0 ± 7.8
ΔMTRC/OMCA	0.67 ± 0.13	14.3 ± 2.08
GSPD	0.42 ± 0.02	12.3 ± 0.58

*HFO (20 mM) reduction sampled at 24 h and extracted with 0.5 N HCl overnight. Fe(II) was determined by using the ferrozine assay (30).

www.pnas.org/cgi/doi/10.1073/pnas.0903426106

Stochastic hybrid modeling of DNA replication across a complete genome

J. Lygeros*†, K. Koutroumpas*, S. Dimopoulos*‡, I. Legouras‡, P. Kouretas*, C. Heichinger§¶, P. Nurse§¶, and Z. Lygerou‡‖

*Automatic Control Laboratory, ETH 8092 Zurich, Switzerland; ‡Laboratory of General Biology, School of Medicine, University of Patras, 26504 Rio, Patras, Greece; §Laboratory of Yeast Genetics and Cell Biology, The Rockefeller University, 1230 York Avenue, New York, NY 10065; and ¶Cancer Research United Kingdom, London Laboratories, London WC2A 3PX, England

Communicated by Fotis C. Kafatos, Imperial College London, London, United Kingdom, June 6, 2008 (received for review December 6, 2007)

DNA replication in eukaryotic cells initiates from hundreds of origins along their genomes, leading to complete duplication of genetic information before cell division. The large number of potential origins, coupled with system uncertainty, dictates the need for new analytical tools to capture spatial and temporal patterns of DNA replication genome-wide. We have developed a stochastic hybrid model that reproduces DNA replication throughout a complete genome. The model can capture different modes of DNA replication and is applicable to various organisms. Using genome-wide data on the location and firing efficiencies of origins in the fission yeast, we show how the DNA replication process evolves during S-phase in the presence of stochastic origin firing. Simulations reveal small regions of the genome that extend S-phase to three times its reported duration. The low levels of late replication predicted by the model are below the detection limit of techniques used to measure S-phase length. Parameter sensitivity analysis shows that increased replication fork speeds genome-wide, or additional origins are not sufficient to reduce S-phase to its reported length. We model the redistribution of a limiting initiation factor during S-phase and show that it could shorten S-phase to the reported duration. Alternatively, S-phase may be extended, and what has traditionally been defined as G2 may be occupied by low levels of DNA synthesis with the onset of mitosis delayed by activation of the G2/M checkpoint.

cell cycle | fission yeast | *Schizosaccharomyces pombe* | random gap problem | systems biology

DNA replication, the process of duplication of the cell's genetic material, must be carried out before every cell division. In bacteria and viruses, replication initiates from a single genomic site in every cell, called the origin of replication (1). Eukaryotic DNA replication is characterized by a higher degree of uncertainty; hundreds to thousands of potential origins exist along the genome that fire with varying efficiencies and at different times during S-phase (2–5). Although the spatial and temporal pattern of DNA replication is not strictly defined in eukaryotic cells, each cell in a population must complete the replication process in an accurate and timely manner. Failure to replicate even a small part of the genome would disrupt proper segregation of the genetic material to the two daughter cells during mitosis, leading to genomic instability.

Initial models of eukaryotic DNA replication, influenced by the replication of bacterial genomes, postulated that defined regions in the genome would act as origins of replication in every cell cycle (6). Indeed, work in the budding yeast permitted the identification of specific sequences that could act as origins with high efficiency (7, 8). However, a simple deterministic model of origin selection could not accommodate the complexity of the replication process observed in Metazoa. Different organisms and cell types exhibit variations on the DNA replication process but a common feature is a degree of uncertainty in origin selection. In *Drosophila* or *Xenopus* preblastula embryos, DNA replication initiates at short intervals (8–15 kb) without apparent sequence specificity (9, 10). Similarly, DNA replication can initiate from any DNA fragment introduced on a plasmid in cultured mammalian cells and at random sites within each fragment (11). In the genomic context, specific regions have origin activity (12, 13), but these lack clear sequence characteristics and fire in only a fraction of the cells (2). As a result, initiation events differ between cells in a population. Stochastic selection of active origins from a large pool of putative origins could lead to the random generation of large interorigin gaps, which may take too long to replicate. This complication of stochastic DNA replication initiation is known as the random gap problem (14). Different solutions to this problem have been suggested, including a spacing mechanism for the generation of defined interorigin gaps (15, 16), coordination of fork velocity with interorigin distance (17), and an increase in origin efficiency during S-phase (14, 18).

Schizosaccharomyces pombe (fission yeast) is an attractive model system for studying genome-wide DNA replication (19). In this organism, similar to Metazoan cells, replication initiates from specific regions, and no origin defining consensus sequences have been characterized (20). Following publication of the *S. pombe* genome (21), bioinformatics (22) and high-throughput analyses (23–25) made it possible to map origins of replication along the complete genome. Dai *et al.* (26) provided evidence that approximately half of the 5.000 intergenic regions present in the *S. pombe* genome may exhibit origin activity and suggested that the origins active in any given S-phase are recruited stochastically among these intergenic regions. Single-molecule analysis has provided evidence for a stochastic pattern of initiation events in fission yeast; DNA combing experiments showed that interorigin distances have an exponential distribution, indicative of stochastic firing (27). Heichinger *et al.* (24) used comparative genomic hybridization to map the sites of replication origins genome-wide and to measure the firing efficiency of each origin. They showed that ≈400 origins have a firing efficiency of 10–60%, while ≈500 more putative origins can be detected, which exhibit efficiencies below 10%. This analysis is in good agreement with microarray analysis based on chromatin immunoprecipitation (25), with single molecule analysis by DNA combing (27), and with 2D gel analysis of replication intermediates (22), and suggests that ≈160 origins out of a total of 900 would be active in any S-phase.

Here, we employ new analytical tools developed in the area of stochastic hybrid control (28) to create a versatile model able to capture different modes of DNA replication in different organ-

Author contributions: J.L., P.N., and Z.L. designed research; J.L., K.K., S.D., I.L., and P.K. performed research; K.K. and C.H. analyzed data; and J.L., P.N., and Z.L. wrote the paper.

The authors declare no conflict of interest.

†To whom correspondence may be addressed at: ETH Zurich, Automatic Control Laboratory, ETL I 22, Physikstrasse 3, 8092 Zurich, Switzerland. E-mail: lygeros@control.ee.ethz.ch.

‖To whom correspondence may be addressed. E-mail lygerou@med.upatras.gr.

This article contains supporting information online at www.pnas.org/cgi/content/full/0805549105/DCSupplemental.

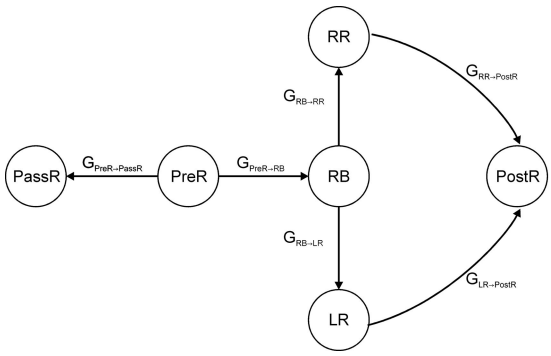

Fig. 1. A stochastic hybrid model of DNA replication. Discrete system states. Each origin *i* along the genome can be in one of 6 different states. Guards (G) govern transitions between states.

isms. Model instantiation using full genome data from fission yeast allows us to simulate DNA replication across the entire genome. Strikingly, *in silico* analysis reveals that a small part of the genome (<5%) may remain unreplicated for a long time, extending into the cell cycle period traditionally defined as G2-phase. The model's flexibility allows us to explore possible solutions to the problems brought about by random origin selection.

Results

A Stochastic Hybrid Model of DNA Replication. Stochastic hybrid systems are used in areas such as mathematical finance, transportation, and telecommunications to model systems that involve the interaction of discrete phenomena, continuous phenomena, and uncertainty (28). DNA replication is a stochastic hybrid process, where discrete transitions between origin states are coupled with the continuous movement of replication forks along the genome, while initiation events are characterized by uncertainty in time and space. We have built a stochastic hybrid model that reproduces the DNA replication process genome-wide. The basic features of the model are presented here; a full description including formal definitions is given in supporting information (SI) *Text*, see also (28) and (29) for the mathematical foundations on which the model is based.

The discrete dynamics of the model capture instantaneous changes in the state of the system, such as origin firing or fork conversion. The state of each origin i is represented by a variable, S_i, that takes one of six values (*PreR*, *PassR*, *RB*, *RL*, *RR*, *PostR*) (Fig. 1). All origins start at the Pre-Replicative state (*PreR*) and end either in the passively replicated (*PassR*) or the Post-Replicating (*PostR*) states. The remaining states discriminate origins from which active forks emanate on both directions (*RB*), only to the left (*LR*) or only to the right (*RL*). Transitions between states are governed by guards (G), logical statements that depend on the continuous and stochastic dynamics of the system and determine when the transition will take place. The continuous dynamics of the model capture the evolution of replication forks along the genome and are represented by two differential equations for each active origin, denoting the number of bases replicated by forks moving away from origin i toward the left (L_i) or right (R_i). Notice that there is a tight coupling between the discrete and continuous states of the process, a defining feature of hybrid systems. Origin firing is characterized by a degree of uncertainty in both the location and time of firing of different origins, captured by the stochastic dynamics of the model. For example, the time T_i at which each origin i will fire in any given cell in a population can be randomly extracted based on its experimentally defined firing probability. Note that each origin is characterized by an intrinsic propensity to fire (referred to hereafter as firing propensity), which denotes the instantaneous probability of firing in the absence of passive replication. The firing propensity differs from the observed firing efficiency (referred to hereafter as observed efficiency), which also depends on the probability of passive replication for this particular origin.

Model Flexibility. The model developed is generally applicable and can, in principle, capture the DNA replication process of any organism for which knowledge of the process of origin selection and activation is available. The model requires as input:

1. The length of the genome in base pairs (L).
2. The number (N) of putative origins of replication and the location X_i of each putative origin i along the genome. The number and location of putative origins can be either deterministic based on experimental data, or extracted randomly based on information such as the average number or average spacing of origins.
3. The firing propensity λ_i of each origin i. This is related to the probability of the origin firing in a unit of time in the absence of passive replication. The model can accommodate different modes of origin firing in time and space (such as stochasticity in space versus a spacing mechanism, stochasticity in time versus a timing mechanism, or constant firing propensity during S-phase versus increasing or decreasing firing propensity as a function of time). We will provide an example of this below.
4. The speed $v(x)$ with which replication forks progress when they go over position x of the genome. Speed can be deterministically defined, or depend on system parameters, such as interorigin spacing.

Model Instantiation. We have used experimental data from the fission yeast *Schizosaccharomyces pombe* for model instantiation. In this organism, there is experimental support for a stochastic mode of origin activation provided at the single cell level (27). High-throughput analyses have permitted the location of the majority of putative origins to be specified along the genome (23–25). Additionally, the firing propensity of each origin along the genome has been determined based on the experimentally defined observed efficiencies in the presence of the drug hydroxyurea, which inhibits fork progression, thereby eliminating passive replication (24). In fission yeast, only a small number of origins genome-wide appear intrinsically late and do not fire during a long hydroxyurea block (24, 27), while for the vast majority of origins, the timing of firing appears correlated to their firing propensities (24), suggesting that stochastic determination of firing times is the prevailing mechanism in the fission yeast (18). Fission yeast offers therefore a good opportunity to capture full genome replication based on experimental data coupled to an understanding of the underlining biology.

The number of origins ($N = 893$), their exact locations ($X_1, X_2, \ldots, X_N$) along the three chromosomes of *S. pombe* (total length $L = 12{,}039{,}987$ bases), and the firing propensity of each origin (FP_i, for $i = 1, 2, \ldots, N$, defined as the fraction of cells in which each origin was observed to fire in hydroxyurea) were used as input (24), see Table S1. A constant speed of $v(x) = 3000$ bases/minute was used genome-wide, based on experimental estimates (24). The firing time, T_i (in minutes), of each origin i was extracted according to an exponential distribution whose rate λ_i depends on FP_i (*SI Text*). Initially, the firing times of different origins in the absence of passive replication were assumed to be statistically independent. This is a valid assumption for fission yeast, based on available experimental data (24). In addition, the firing propensity of each origin as a function of time (λ_i) was assumed to be constant during S-phase. This means that the total remaining firing propensity of the system decreases as S-phase progresses. An alternative model that relaxes both

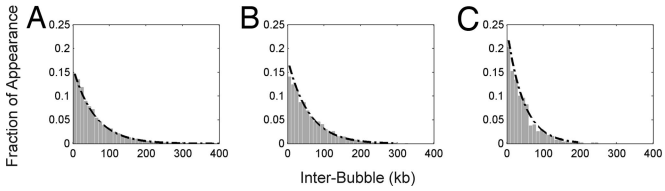

Fig. 2. Model Validation: *In silico* analysis of replication intermediates. Following 2,000 Monte Carlo simulations of full genome DNA replication, >5,000 genomic pieces matching is size the DNA fibres analyzed in (27) were extracted and distances from the end of one replication bubble to the beginning of the next (inter-bubble distance) calculated. Histograms of the fraction of inter-bubble distances falling within a given range (in bins of 10 kb) are shown for: (*A*) Hydroxyurea, (*B*) Early S-phase, and (*C*) Late S-phase. For (*C*), only replication bubbles <30 kb were included, following (27). The curves best-fitting the experimental data of (27) have been over-laid for comparison.

assumptions will be discussed below (Propensity Redistribution Model).

Simulation Results. The model was coded in the Matlab environment. Monte-Carlo simulations were used to capture the process of DNA replication at a single cell, extract useful statistics on model predictions at a population level, and validate them against experimental data. Examples of simulation runs are given in Fig. S1 and show the stochastic nature of the process. Diagnostic tests were carried out which showed that observed fluctuations in continuous and discrete system parameters during the evolution of DNA replication behave as predicted and each part of the genome is copied once and only once (see *SI Text* and Fig. S2).

To validate the model using an independent source of experimental data, the output of simulations was compared to data derived by the analysis of replication intermediated at the single molecule level by DNA combing (27). In this method, replication intermediates are labeled *in vivo*, genomic DNA is extracted, DNA fibers stretched and visualized. To reproduce *in silico* the experimental conditions of (27), random genomic pieces with a size distribution matching the one of the fibers analyzed in (27) were extracted from 2,000 simulations under three circumstances: assuming that the first 160 origins (24) scheduled to fire in each simulation would arrest after 5 kb [mimicking the hydroxyurea block of (27)], when 30% of the genome had been replicated in each simulation (early S-phase) and when 70% of the genome had been replicated in each simulation (late S-phase). In Fig. 2, the distribution of distances between replication bubbles (inter-bubble distances) is shown for each case. The exponential curves generated by (27) fitting their experimental data have been overlaid for comparison. In all cases, simulated results closely matched the experimentally derived curves (R^2 values of 0.9937, 0.9794, and 0.9749, respectively).

Model output was used to analyze genome-wide replication kinetics *in silico*. In Fig. 3*A*, the fraction of unreplicated DNA as a function of time is shown, with each curve representing one simulation (single cell). While in all simulations the process of conversion from fully unreplicated (1) to fully replicated (0) follows similar kinetics, differences in individual cells due to stochastic phenomena are evident. Strikingly, while the bulk of DNA replication takes place quickly during the first half of the process, a very long tail is evident in the second half, showing that completion of DNA replication of a remaining small fraction of the genome takes a long time. This is also evident in the rate of new DNA synthesis (Fig. 3*B*), which shows the expected dumb-bell shape observed experimentally, but continues with a long tail. As a consequence, the mean completion time of the process (Fig. 3*C*) is unexpectedly long [67 min, as opposed to experimental estimates of ≈20 min (30)]. This is not due to

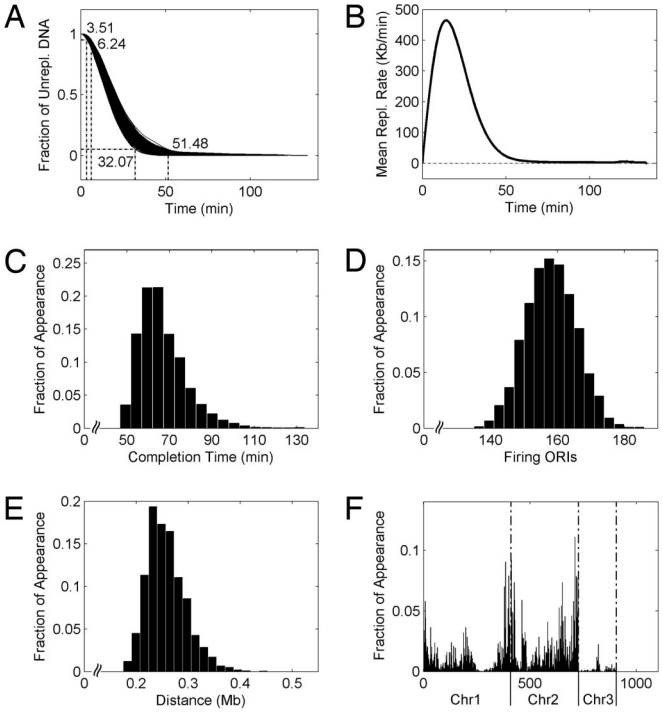

Fig. 3. A stochastic mode of DNA replication across the fission yeast genome results in long S-phase completion times. (*A*) Evolution of replication. The fraction of unreplicated DNA is plotted as a function of time for 100 different simulations. The dashed lines show where the percentage of unreplicated DNA becomes 5% and 95%, and the numbers indicate the minimum and maximum time where these values are reached (in min). Note the differences in the form of the curves resulting from stochasticity and the long tail that delays completion of the process. (*B*) Mean replication rate (kb/min) as a function of time. Note again the long tail of the curve. The end of the curve corresponds to the point where replication has completed in all cells in the sample. (*C*) Histogram of DNA replication completion time. (*D*) Histogram of the number of origins that are observed to fire in each simulation. (*E*) Histogram of the maximum observed distance between adjacent firing origins in each simulation. (*F*) Histogram of interorigin intervals that complete replication after 99% of the genome has been replicated. Each potential interorigin interval along the fission yeast genome is marked by a vertical line, whose height reflects the fraction of simulations where it is observed to replicate late. Chromosomes are separated with dashed vertical lines. Note an accumulation of late replicating regions close to the ends of chromosomes 1 and 2. In (*C–E*) Fraction of appearance is the fraction of simulations that exhibit each abscissa value. Results in (*B–F*) based on 2,000 Monte Carlo simulations.

less origins firing in each simulation than expected (Fig. 3*D*, average 159 origins firing in each simulation, very close to the experimental estimate of ≈160). In each simulation, at least one region of the genome exhibits a distance between adjacent active origins that is very long (Fig. 3*E*, mean maximum interorigin distance of 258 kb, replication of which would last 43 min with a fork speed of 3,000 bases/minute, assuming both origins fire at the same time). Different locations along each chromosome delay the process in different simulations (Fig. 3*F*, showing regions that terminate replication after 99% of the genome has replicated) and even the most problematic regions do not delay the process in >12% of the simulations. This is explained by the random distribution of active origins along the genome (random gap problem). The model predicts that there is a small percentage (<5%) of DNA that remains unreplicated for a long time, that the rate of DNA synthesis at late time points is low and that there is an infrequent appearance within the population of each problematic region. Such a delay in S-phase completion is difficult to detect experimentally and could well have gone unnoticed.

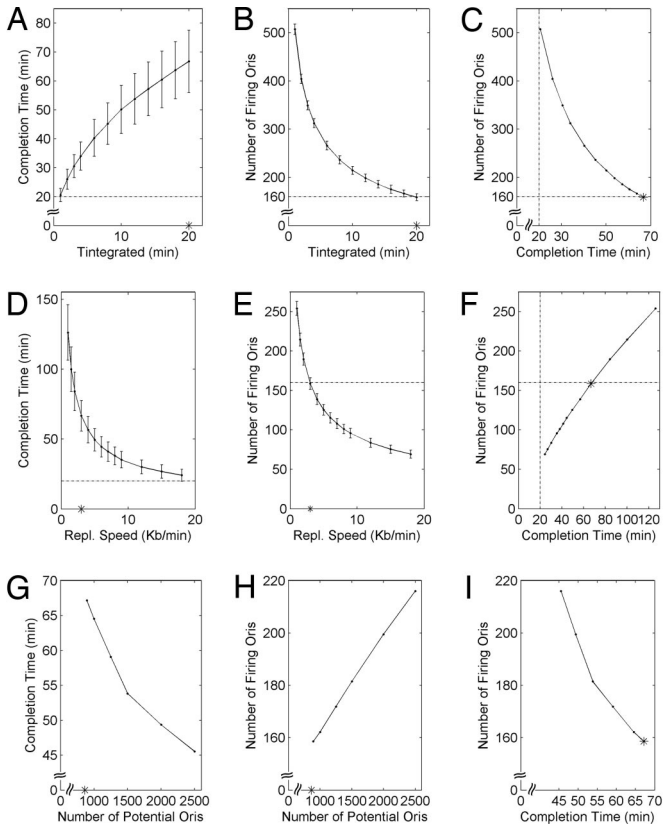

Fig. 4. Sensitivity Analysis. Simulation results obtained for single parameter variations of the base model. Predicted average S-phase duration (Completion Time), average number of firing origins and S-phase duration versus average number of firing origins as: the parameter T_f ranges from 1 to 20 min (*A–C* respectively); fork velocity varies from 1,000 to 18,000 bases/minute (*D–F*); and. the total number of origins is increased (to 1,000, 1,250, 1,500, 2,000, and 2,500, *G–I*). In all cases, mean values from 2,000 Monte Carlo simulation runs are shown. Vertical lines indicate the minimum and maximum values obtained in the same sample. We interpolate linearly between adjacent points. The experimentally expected value for each quantity is plotted as a dotted line in (*A–F*). The nominal parameter values used in Fig. 3 are indicated by stars.

Sensitivity Analysis. Given the unexpectedly long S-phase completion times revealed by full genome simulations, we proceeded to vary the input parameters to test their effect on system behavior. Fig. 4 shows how the predicted S-phase duration and the number of firing origins depends on three of the model parameters: the intrinsic firing propensity of each origin; the fork speed $v(x)$; and the number of origins N.

To systematically increase the firing propensity of all origins we decreased the parameter T_f (a system parameter inversely correlated to the firing propensity of each origin, see *SI Text*) from 20 min down to 1 min (Fig. 4 *A–C*). As firing propensities increase, the predicted average S-phase duration decreases, from 67 to 21 min. As expected, however, the predicted average number of firing origins also increases from 18% to 57%. While small values of T_f (high firing propensities) lead to reasonable S-phase completion times, the high number of firing origins predicted for these values is not supported by experimental data (22, 24, 27, 31).

As fork speed $v(x)$ increases from 1 kb/min to ≈6 kb/min, S-phase completion times are quickly reduced (from an average of 126 min to an average of 44 min, Fig. 4*D*). However, the fraction of firing origins is similarly reduced (Fig. 4*E*); moreover, further increases in speed have smaller effects on the overall completion time. For S-phase completion time to reach the expected 20 min, speeds of >18 kb/min would be required, which are unrealistically high.

Finally, the number of potential origins, N, was increased by introducing more low efficiency origins at specific locations along the genome (Fig. 4 *G–I*). The experimental method used for origin identification (24) could not reliably detect origins with <10% efficiency. New origins with efficiencies below this experimental threshold were introduced into biologically relevant chromosomal locations, predicted by bioinformatics analysis to fulfill characteristics of known origin regions (*SI Text*). Criteria depending on AT content and size were used to gradually introduce additional origins with observed efficiencies of $FP_i = 8\%$. The addition of new origins causes a decrease in the predicted average S-phase duration, however even after adding all potential origins predicted by our analysis (≈50% of all intergenes in the fission yeast genome), S-phase completion times still remained long (46 min).

Firing Propensity Redistribution Model. Our analysis leads to two alternatives: Either S-phase duration in the presence of stochastic firing is longer than experimentally defined, or a further mechanism exists that succeeds in limiting S-phase duration circumventing the random gap problem. One proposed solution to the random gap problem is that the efficiency of unreplicated origins may increase during the course of S-phase (14, 18). Stochasticity could be brought about by a limiting factor that allows firing of a fraction of putative origins. The redistribution of this factor after an origin fires or gets passively replicated would imply a gradual increase in firing propensity of origins still in the prereplicative state as S-phase progresses (18). This biological scenario can be captured in our model by keeping the total system firing propensity (given by the sum of the firing propensities of all origins) constant throughout S-phase. Each origin that fires or becomes passively replicated releases its firing propensity λ_i, which is redistributed to the remaining *PreR* origins proportionally to their initial firing propensities. We call the resulting variant of the model the "Firing Propensity Redistribution Model" and we describe it in detail in the *SI Text*.

Fig. 5*A* shows the evolution of unreplicated DNA over time predicted by the firing propensity redistribution model. While the kinetics of bulk DNA replication are similar to the basic model, the curve has lost its long tail. A similar change is observed in the curve depicting the rate of newly synthesized DNA over time (Fig. 5*B*). Consequently, the average time of completion of DNA replication has decreased from 67 min in the basic model to 33 min (Fig. 5*C*), concomitant with a decrease in the mean maximum distance observed between active origins in each simulation (from 258 kb to 135 kb, Fig. 5*E*). The number of firing origins has increased to a mean of 286 (Fig. 5*D*). The regions that replicate last in the process are shown in Fig. 5*F*. Two additional variants of the redistribution model (redistribution upon fork conversion proportional to initial firing propensity and redistribution upon firing proportional to current firing propensity) are discussed in the *SI Text*. The simulation results for the two variants (data not shown) were similar to the results of Fig. 5, with redistribution upon fork conversion leading to slightly longer S-phase duration (average replication completion time of 37 min).

We conclude that the redistribution during S-phase of a limiting factor acting at the level of initiation of DNA replication is one way by which the problems arising due to stochastic origin activation can be circumvented.

Discussion

Simulating Full Genome DNA Replication. A general and flexible stochastic hybrid model was developed to capture the DNA replication process during the S-phase of the cell cycle. The model in principle applies to any organism; here it was instan-

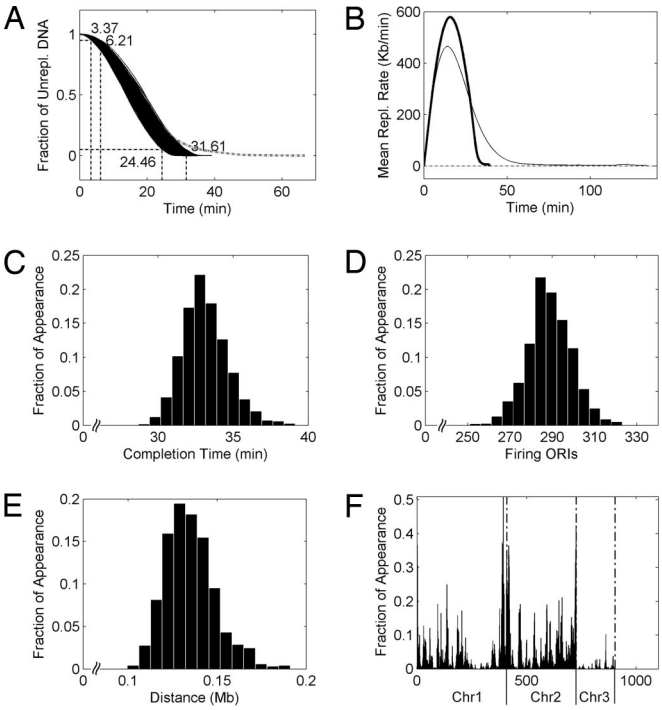

Fig. 5. Firing propensity redistribution model. (*A*) Evolution of replication. The fraction of unreplicated DNA is plotted as a function of time for 100 different simulations. Base case average (Fig. 3*A*) is shown as thick, gray dotted line for comparison. Note the considerably shorter tail. (*B*) Mean replication rate (bases/min) as a function of time. Base case (Fig. 3*B*) is shown as a thin line for comparison. Note again the shorter tail for the propensity redistribution model. (*C*) Histogram of DNA replication completion time. (*D*) Histogram of the number of origins that fire in each simulation. (*E*) Histogram of the maximum distance between adjacent firing origins. (*F*) Histogram of interorigin intervals that complete replication after 99% of the genome has been replicated. In (*C–E*) Fraction of appearance is the fraction of simulations that exhibit each abscissa value. Results in (*B–F*) based on 2,000 Monte Carlo simulations.

tiated using experimental data for *Schizosaccharomyces pombe* (24). The resulting model was coded in simulation and tested against independent genome-wide data on fission yeast DNA replication (27), showing that it captures well our current understanding of the DNA replication process. The predictions of the model suggest that given stochastic firing, S-phase duration would be considerably longer than generally accepted for this organism [on average 67 min and reaching 113 min in some cells, instead of 20 min (30)], due to randomly generated large interorigin gaps.

The Effect of Parameter Values. The model's flexibility allowed us to introduce modifications to test possible explanations for this difference. Input parameters (intrinsic firing propensities of all origins, fork speed and number of potential origins) were varied to investigate their effect on S-phase duration. While the parameters tested have a direct effect on S-phase duration, none of these changes alone seems capable of explaining in a biologically realistic way the discrepancy between experimentally measured S-phase duration and the predictions of the model. Local effects, such as an increase in fork speed at problematic regions, could conceivably speed up DNA replication completion. However, different regions create problems in different cells (see Fig. 3*F*) and local effects at a number of these problematic regions would be required for an effect at the population level. Combining changes in input parameters could potentially reveal a biologically plausible solution, although we have been unable to pinpoint a likely combination (data not shown). More complex patterns of origin firing, such as coordination of fork velocity with interorigin distance (17) or defined spacing of active origins (15, 16) could potentially decrease S-phase duration. Such conjectures can be tested by our model and we have made use of the model's versatility to capture one such possibility.

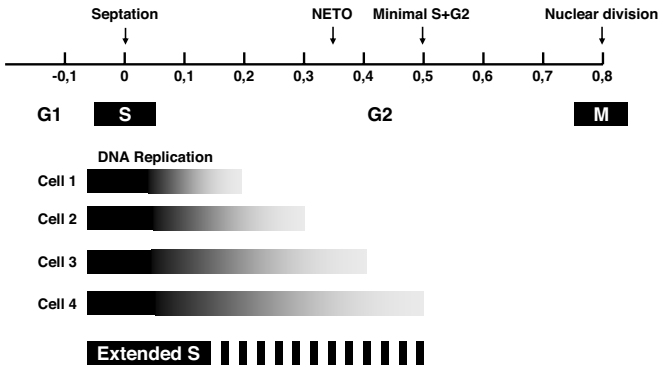

Fig. 6. A model for an extended S-phase in the fission yeast cell cycle. See text for details

Firing Propensity Redistribution. The model was modified to accommodate a specific biological hypothesis for the reason behind the probabilistic nature of the firing of origins: the presence of a limiting initiation factor, which is released following origin firing and binds again to origins still at the prereplicative state (18). Since the number of origins in the prereplicative state decreases as S-phase progresses, the probability of the limiting factor binding to a particular prereplicative origin (and therefore the firing probability of this origin) increases along S-phase (18). The firing propensity redistribution model was developed to code this hypothesis, assuming that the pool of available factor remains roughly constant during S-phase. Using the initial model parameters, the firing propensity redistribution model predicts a considerable reduction in S-phase duration, while also maintaining the average fraction of origins that fire during S-phase in a reasonable range. Strikingly, while bulk replication progresses with dynamics similar to the basic model, the long tail of replication has disappeared. The firing propensity redistribution model reproduces the experimental data of (27) as well as the base-case model does (data not shown), consistent with the bulk of replication progressing in a similar manner in the two models. Firing propensity redistribution therefore offers a possible solution to the random completion problem, without the need to postulate a spacing mechanism for positioning active origins at appropriate locations. This scenario is indeed biologically plausible. The limiting factor should act at the level of origin activation, and not origin licensing that is globally inhibited following S-phase onset. It could be released at firing (Fig. 5) or upon fork conversion (data not shown). Cyclin-dependent kinases or the Dbf4-dependent kinase (DDK), which are required for initiation, are possible candidates for the first option, while Cdc45, previously shown to be limiting for replication in *Xenopus* (32), is a candidate for the second option.

"G2" Replication? An alternative explanation is that DNA replication takes longer than has been generally accepted. The predicted low percentage of unreplicated DNA (<5%, Fig. 3*A*), the very low rate of replication at late time points (Fig. 3*B*), and the fact that problematic regions (those that delay the overall process) differ in different cells (Fig. 3*F*), suggest that such a prolonged phase of DNA replication could have gone unnoticed. The inability of techniques used up to now to detect small

amounts of unreplicated DNA scattered across different regions in different cells means that the length of S-phase may be longer than our current estimates.

An S-phase that extends into what we now define as G2-phase is consistent with known properties of the fission yeast cell cycle and the available experimental data (Fig. 6). G2 makes up the largest part of the fission yeast cell cycle, and it would be inconsequential if small parts of the genome continued replication for an extended time while the cell grows in size. The checkpoint that arrests entry into mitosis in the presence of unreplicated DNA would become essential if the G2/M size control was inoperative. This alternative explanation can explain a number of previous observations in fission yeast. In fission yeast, S+G2 is not shortened beneath a certain minimal period (0.5 of the cell cycle) even in large cells that have reached the critical size for mitosis (33, 34). Furthermore, a wee1 mutant in which the G2/M size control is inoperative and, as a consequence, advances into mitosis at a small size (33, 35) also has a S+G2 length of approximately 0.5 of the cell cycle. Significantly, wee1 is synthetically lethal with a rad3 mutant, defective in the G2/M checkpoint control (36). We postulate that in wee1 mutant cells, the extended S-phase occupies all of the previously designated "G2" period and mitosis is restrained by the Rad3-dependent G2/M checkpoint. When the checkpoint is inactivated, cells cannot restrain entry into mitosis even though S is not complete, and this leads to lethality. The minimal S+G2 (65–120 min depending on growth conditions), is similar to the extended S-phase suggested here (Fig. 3*A*). Additional support for the extended S-phase hypothesis comes from studies of New-End Take-Off (NETO), a growth transition that is dependent on the completion of DNA replication and takes place in early G2. This occurs at approximately 0.35 of the cycle (37) rather than 0.1 of the cycle, the time when DNA replication is believed to be completed (30). NETO shows heterogeneity among cells in a population (38), consistent with the results presented here.

Any organism with an active checkpoint inhibiting entry into mitosis until DNA replication has been completed might not require a further specific mechanism to inhibit the random gap problem. In this light, part or all of what we now experimentally define as G2-phase might be better described as the period required for the cell to complete replication of the last remaining parts of its genome. Proving or disproving the extended S-phase hypothesis will have to await the development of more sensitive methods, able to detect the low amounts of late replication suggested by the model described here.

ACKNOWLEDGMENTS. We are grateful to Jacqueline Hayles for critical reading of the manuscript and helpful suggestions. This work was supported by the European Commission under the project HYGEIA (FP6-NEST-004995)

1. Kornberg A, Baker T (1992) *DNA replication* (WH Freeman, New York) 2nd Ed.
2. DePamphilis ME (2006) *DNA Replication and human disease* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York).
3. Gilbert DM (2001) Making sense of eukaryotic DNA replication origins. *Science* 294:96–100.
4. Costa S, Blow JJ (2007) The elusive determinants of replication origins. *EMBO Rep* 8:332–334.
5. DePamphilis ML (1999) Replication origins in metazoan chromosomes: Fact or fiction? *Bioessays* 21:5–16.
6. Gilbert DM (2004) In search of the holy replicator. *Nat Rev Mol Cell Biol* 5:848–855.
7. Marahrens Y, Stillman B (1992) A yeast chromosomal origin of DNA replication defined by multiple functional elements. *Science* 255:817–823.
8. Theis JF, Newlon CS (1997) The ARS309 chromosomal replicator of Saccharomyces cerevisiae depends on an exceptional ARS consensus sequence. *Proc Natl Acad Sci USA* 94:10786–10791.
9. Harland RM, Laskey RA (1980) Regulated replication of DNA microinjected into eggs of *Xenopus laevis*. *Cell* 21:761–771.
10. Mechali M, Kearsey S (1984) Lack of specific sequence requirement for DNA replication in *Xenopus* eggs compared with high sequence specificity in yeast. *Cell* 38:55–64.
11. Krysan PJ, Calos MP (1991) Replication initiates at multiple locations on an autonomously replicating plasmid in human cells. *Mol Cell Biol* 11:1464–1472.
12. Heintz NH, Hamlin JL (1982) An amplified chromosomal sequence that includes the gene for dihydrofolate reductase initiates replication within specific restriction fragments. *Proc Natl Acad Sci USA* 79:4083–4087.
13. Kitsberg D, Selig S, Keshet I, Cedar H (1993) Replication structure of the human beta-globin gene domain. *Nature* 366:588–590.
14. Hyrien O, Marheineke K, Goldar A (2003) Paradoxes of eukaryotic DNA replication: MCM proteins and the random completion problem. *Bioessays* 25:116–125.
15. Blow JJ, Gillespie PJ, Francis D, Jackson DA (2001) Replication origins in *Xenopus* egg extract Are 5–15 kilobases apart and are activated in clusters that fire at different times. *J Cell Biol* 152:15–25.
16. Hyrien O, Mechali M (1993) Chromosomal replication initiates and terminates at random sequences but at regular intervals in the ribosomal DNA of *Xenopus* early embryos. *EMBO J* 12:4511–4520.
17. Conti C, *et al.* (2007) Replication fork velocities at adjacent replication origins are coordinately modified during DNA replication in human cells. *Mol Biol Cell* 18:3059–3067.
18. Rhind N (2006) DNA replication timing: Random thoughts about origin firing. *Nat Cell Biol* 8:1313–1316.
19. Legouras I, Xouri G, Dimopoulos S, Lygeros J, Lygerou Z (2006) DNA replication in the fission yeast: Robustness in the face of uncertainty. *Yeast* 23:951–962.
20. Dubey DD, Kim SM, Todorov IT, Huberman JA (1996) Large, complex modular structure of a fission yeast DNA replication origin. *Curr Biol* 6:467–473.
21. Wood V, *et al.* (2002) The genome sequence of Schizosaccharomyces pombe. *Nature* 415:871–880.
22. Segurado M, de Luis A, Antequera F (2003) Genome-wide distribution of DNA replication origins at A+T-rich islands in Schizosaccharomyces pombe. *EMBO Rep* 4:1048–1053.
23. Feng W, *et al.* (2006) Genomic mapping of single-stranded DNA in hydroxyurea-challenged yeasts identifies origins of replication. *Nat Cell Biol* 8:148–155.
24. Heichinger C, Penkett CJ, Bahler J, Nurse P (2006) Genome-wide characterization of fission yeast DNA replication origins. *EMBO J* 25:5171–5179.
25. Hayashi M, *et al.* (2007) Genome-wide localization of pre-RC sites and identification of replication origins in fission yeast. *EMBO J* 26:2821.
26. Dai J, Chuang RY, Kelly TJ (2005) DNA replication origins in the Schizosaccharomyces pombe genome. *Proc Natl Acad Sci USA* 102:337–342.
27. Patel PK, Arcangioli B, Baker SP, Bensimon A, Rhind N (2006) DNA replication origins fire stochastically in fission yeast. *Mol Biol Cell* 17:308–316.
28. Cassandras C, Lygeros JE (2006) *Stochastic Hybrid Systems* (CRC Press, Boca Raton).
29. Davis MHA (1993) *Markov Models and Optimization* (Chapman and Hall, London).
30. Nasmyth K, Nurse P, Fraser RS (1979) The effect of cell mass on the cell cycle timing and duration of S-phase in fission yeast. *J Cell Sci* 39:215–233.
31. Kim SM, Huberman JA (2001) Regulation of replication timing in fission yeast. *EMBO J* 20:6115–6126.
32. Edwards MC, *et al.* (2002) MCM2–7 complexes bind chromatin in a distributed pattern surrounding the origin recognition complex in *Xenopus* egg extracts. *J Biol Chem* 277:33049–33057.
33. Fantes PA, Nurse P (1978) Control of the timing of cell division in fission yeast. *Exp Cell Res* 115:317–329.
34. Fantes PA (1977) Control of cell size and cycle time in Schizosaccharomyces pombe. *J Cell Sci* 24:51–67.
35. Nurse P (1975) Genetic control of cell size at cell division in yeast. *Nature* 256:547–551.
36. al-Khodairy F, Carr AM (1992) DNA repair mutants defining G2 checkpoint pathways in Schizosaccharomyces pombe. *EMBO J* 11:1343–1350.
37. Mitchison JM, Nurse P (1985) Growth in cell length in the fission yeast Schizosaccharomyces pombe. *J Cell Sci* 75:357–376.
38. Sveiczer A, Novak B, Mitchison JM (1996) The size control of fission yeast revisited. *J Cell Sci* 109:2947–2957.