Defects and DNA Replication

Michel G. Gauthier, John Herrick, and John Bechhoefer*

Department of Physics, Simon Fraser University, Burnaby, British Columbia, Canada, V5A 1S6

(Received 12 December 2009; published 28 May 2010)

We introduce a rate-equation formalism to study DNA replication kinetics in the presence of defects resulting from DNA damage and find a crossover between two regimes: a *normal* regime, where the influence of defects is local, and an *initiation-limited* regime. In the latter, defects have a global impact on replication, whose progress is set by the rate at which origins of replication are activated, or initiated. Normal, healthy cells have defect densities in the normal regime. Our model can explain an observed correlation between interorigin separation and rate of DNA replication.

DOI: 10.1103/PhysRevLett.104.218104 PACS numbers: 87.15.A−, 87.14.gk, 87.17.Ee

In higher organisms, DNA replication is initiated at distinct sites called *replication origins*, where pairs of *replication forks* begin to duplicate DNA bidirectionally outward from the origin site until they eventually coalesce with another fork. Unfortunately, at natural *pause sites* and at externally caused defects along the DNA replication forks can slow, or stall [1,2]. Such fork stalls can trigger local and global checkpoints that can suspend replication. Even in the absence of checkpoint activity, fork stalls slow the replication program. In embryonic cells, such delays may lead to cell death, since replication is decoupled from cell division [3], while in mature (somatic) cells, DNA damage is an initial step in the development of cancer [4]. However, although fork stalls clearly play an important role during replication, they have not been considered in previous modeling approaches [5–7]. Two exceptions are discussed below [8,9].

In this Letter, we show how to accommodate defects within an analytical DNA replication model. At low defect densities or short repair times, we find that replication kinetics is negligibly affected by fork blocks. However, when the density of fork blocks exceeds a critical crossover value, the density of initiated origins increases sharply. The transition from normal to initiation-limited replication is analogous to a percolation threshold, above which defect interactions span arbitrarily long distances.

Because DNA lesions that lead to fork blocks are present in great number, even in normal cells [4], cells have developed various response mechanisms to address these stalling events [10]. One mechanism that has been observed is a correlation between an apparent slowing down of forks with an increase in the origin firing density [11]. It has been suggested that active control mechanisms are responsible for this correlation [12,13]. Here, we show that fork stalls alone can induce such correlations, without recourse to explicit control mechanisms.

The KJMA approach.—In previous work, our group modeled the replication kinetics of long DNA chains using a formal analogy with the Kolmogorov-Johnson-Mehl-Avrami (KJMA) theory of phase-transition kinetics in one spatial dimension [5]. That work was based on Sekimoto's rate-equation (RE) formulation [14], which we extend here to account for the presence of fork blocks.

The KJMA model incorporates three features of replication: initiation at multiple sites; bidirectional movement of replication forks away from their origin; and coalescence as merging forks annihilate. One important result is that for large, defect-free genomes, the fraction of replicated DNA is $f_0(t) = 1 - \exp[-I v t^2]$, where t is the time elapsed since the beginning of replication, I the initiation rate of replication domains per time per unreplicated length of DNA, and v the growth rate of replication domains (forks) [5]. Here, we assume $I(x, t) = I$ and $v(x, t) = v$ are homogeneous and constant, which corresponds to averaging over the genome and the cell cycle. The product $I v t^2$ is the number of origins that are expected to fire in the shaded triangle in Fig. 1(a). This *causal area* is the space-time region in which an initiation would cause the point x to be replicated by time t.

We define a defect to be a position along the genome at which replication forks stall for an average time τ (the repair time). We generalize the previous expression for $f_0(t)$ to study the impact of a single defect on the replication kinetics of an infinitely large genome by calculating the probability to have no initiation in a reduced causal area. The space-time signature of a single defect at x_d is

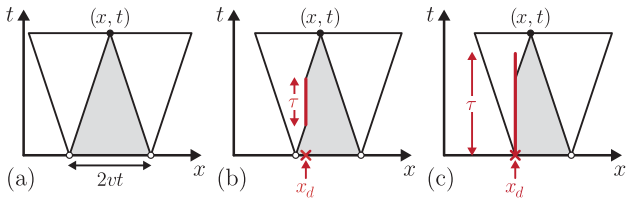

FIG. 1 (color online). Space-time diagram of replication. (a) During normal replication, the point x is replicated at time $\leq t$ only if a fork is activated within the shaded triangular area. (b), (c) Fork blocks (defects) near x can reduce the relevant space-time area and thereby delay replication at x. (b) For small τ, forks can initiate to the left of the defect ($x < x_d$), stall for a time τ, and still replicate x before t. (c) For large τ, forks to the left of the defect cannot replicate x before t.

illustrated in Figs. 1(b) and 1(c), for short and long repair times, respectively. In both cases, we can calculate the reduced causal area and the replication fraction as function of space and time. Figures 2(b) and 2(c) show examples of $f_0(t)$ and $f(x, t)$, respectively, for a single, fixed defect. Incoming forks stall at x_d, and replication is complete on both sides of the defect when a second fork arrives from the other side. The replication fraction $f(x, t)$ in Fig. 2(b) is given by $1 - \exp[-\iint_{x,t\in \mathrm{RCA}} I(x, t)dxdt]$ which represents the probability that no initiation occurred in the reduced causal area (RCA) depicted in Fig. 1(c) when $\tau \to \infty$. Figure 2(c) depicts contours of equal replication delay, $\Delta f = f_0 - f$, which characterize the average suppression of replication in the space-time area surrounding a defect. The shape of the defect contours will differ for finite repair time τ. The important qualitative point is that defects are local. $\Delta f(x, t)$ decays as a Gaussian in space and time, with a characteristic length $(\sqrt{v/I})$ and time $(1/\sqrt{Iv})$, implying two distinct regimes: a low-density regime, where defects are isolated and have only local impact, and a high-density regime, where defects overlap (percolate) and have global impact. Such a picture is confirmed by analyzing the case of many defects.

Rate equations (REs).—Although the replication delay produced by a single defect is easy to describe, analytical expressions become too complex in the realistic case of many defects. Systems with a few defects can be studied using computer simulations, as Blow and Ge did [8], but such methods limit the system size that can be explored.

As an alternative to lengthy Monte Carlo simulations, we introduce REs to calculate the replicated fraction of DNA and the number of replication forks as a function of time in the presence of many defects. Meneghini and de Mello Filho used a simpler version of such an approach to calculate the rate of DNA synthesis in the presence of UV lesions [9]. We assume that lesions are separated by a mean spacing δ and that forks are blocked for an average time τ before being repaired. As before, we assume that v, I, τ and δ are all constant throughout both space and time [15]. We also assume that initiation, stalling, and repair events are independent of one another and thus can be modeled as

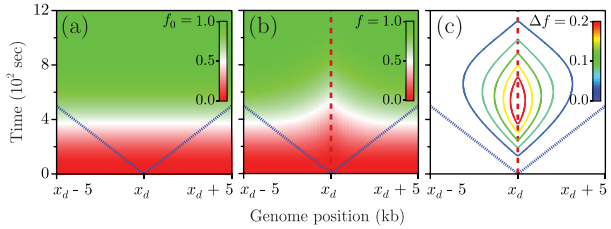

FIG. 2 (color online). (a) Replication fraction $f_0(t)$ for a section of an infinite genome ($v = 10^{-2}$ kb/sec and $I = 5 \times 10^{-4}$ kb^{-1} sec^{-1} [29]). (b) Replication fraction $f(x, t)$ with a defect at $x = x_d$ and an infinite repair time ($\tau \to \infty$). (c) $\Delta f = f_0 - f$. The defect influences the replication kinetics in the space-time area above the V-shaped line ($vt = |x - x_d|$).

Poisson processes. We denote by $\rho(t)$ and $\rho'(t)$ the time-dependent densities of moving and stalled forks. Although stalled forks do not move, their concentration changes as repairs and new stalls occur. Both densities decrease as forks coalesce. We then have

$$\frac{df(t)}{dt} = 2v\rho(t), \tag{1}$$

which asserts that the average rate of replication is given by the local fork density times the fork speed. The factor of 2 reflects the bidirectional growth of replication domains [16]. The density of moving forks is given by

$$\frac{d\rho(t)}{dt} = I[1 - f(t)] - \frac{v\rho(t)}{\delta} + \frac{\rho'(t)}{\tau} - \frac{v[2\rho^2(t) + \rho(t)\rho'(t)]}{1 - f(t)}, \tag{2}$$

with four contributions: (1) replication forks are created at a rate I at unreplicated parts of the genome; (2) moving forks stall at a rate v/δ; (3) stalled forks are repaired and restart at a rate τ^{-1}; (4) moving forks annihilate when they coalesce with opposite forks (moving or stalled) [17]. An analog of Eq. (2) has been used to study the growth of one-dimensional crystal lamella [18]. Finally, the density of blocked forks $\rho'(t)$ is

$$\frac{d\rho'(t)}{dt} = \frac{v\rho(t)}{\delta} - \frac{\rho'(t)}{\tau} - \frac{v\rho(t)\rho'(t)}{1 - f(t)}, \tag{3}$$

where the three terms represent blocking, repair, and coalescence, respectively. The differential equations [Eqs. (1)–(3)] can be solved numerically to find a mean-field solution for $f(t)$, $\rho(t)$ and $\rho'(t)$, which corresponds to the spatial average of an infinitely large genome over an infinite number of realizations. As the repair time τ turns out to play a marginal role in the replication kinetics of normal, healthy cells, we focus on replication kinetics for the worst-case scenario, an infinite repair time.

Two regimes of replication.—We now show that, for $\tau \to \infty$, Eqs. (1)–(3) imply two regimes of replication. Let ℓ_i be the average distance between two origins of replication, given by $\ell_i^{-1} = \int_0^\infty I[1 - f(t)]dt$. We scale lengths by $\ell_0 = 2\sqrt{v/I\pi}$, the average interorigin distance in the absence of defects (from the expression for $f_0(x, t)$) and times by $t_0 = \ell_0/v$. In Fig. 3, we present a parametric study of the average interorigin distance ℓ_i as a function of the dimensionless defect density $\lambda = \ell_0/\delta$. All solutions collapse to a line that closely approximates $\ell_i/\ell_0 = 1/\sqrt[n]{1 + \lambda^n}$ with $n = 1.6$. These results clearly show a crossover between replication regimes at $\lambda = 1$ (see Fig. 3) that can be understood as the intersection of two asymptotic limits. For $\lambda \to 0$ (no defects), we must have $\ell_i \to \ell_0$. For $\lambda \to \infty$, we have $\ell_i \to \delta$, since the DNA between two defects is replicated independently of the rest of the genome. In this regime, the numbers of origins

TABLE I. Experimental values of the replication parameters for various organisms and estimates of the scaled density λ.

organism	$\delta(kb)$	$\ell_i(kb)$	λ
Xenopus embryo	375 [4]	6 [19]	0.016[a]
S. cerevisiae	600 [4]	44 [20]	0.074[a]
human (genome wide)	1200 [4]	110 [11]	0.093[a]
human (amplicon)	—	85 [11]	0.68[b]
human ($^-$DDR, $^-$Ras)	1200 [4]	246 [21]	0.22[a]
human ($^-$DDR, $^+$Ras)	—	186 [21]	0.80[b]

[a]From $\ell_i = \ell_0/\sqrt[n]{1+\lambda^n}$, $\lambda = 1/\sqrt[n]{(\delta/\ell_i)^n - 1}$. We used $n = 1.6$.
[b]Here, δ is unknown and ℓ_0 is assumed to be the same as the normal corresponding cell. Then $\lambda = \sqrt[n]{(\ell_0/\ell_i)^n - 1}$.

and replication forks required to complete replication increase drastically, and cells would require much larger amounts of replication factors such as DNA polymerase. The crossover occurs when the density of defects equals the density of firing origins in a defect-free genome.

Where do various living organisms lie on Fig. 3? For Table I, we combine measurements of interorigin distances with estimates of defect densities. Although we assume that the repair time is infinite, a more detailed analysis shows that a finite repair time merely reduces the effective value of λ by a minor amount [22]. Then, frog embryos, which have a simplified cell cycle that focuses on replicating as fast as possible, have $\lambda = 0.016$. Normal budding yeast cells and human somatic cells have $\lambda \approx 0.1$. The value for frog embryos is markedly lower than for somatic cells. Perhaps embryos need a larger "safety margin" because they lack the active checkpoint mechanism that mature cells have that can delay the cell cycle in order to repair DNA damage [23]. Next, a human cell mutant in which the DNA damage checkpoint response, DDR, has been inactivated but is otherwise normal gives a slightly higher λ, 0.22. All these values are well below the crossover at $\lambda = 1$.

We also include two cases that approach the crossover point. The first is from a specific region of the human genome (amplicon) that is known, qualitatively, to have more endogenous fork pauses [11]. The second is from DDR-inactivated cells where the Ras oncogene has been expressed, an event that is associated with increases in fork pausing and is a first step towards cancer [21]. Although there is unfortunately no direct data on the increase in defect density, we can estimate the value of λ via the measured reduction in interorigin distances and the model relation between interorigin distance shown in Fig. 3. These give $\lambda \lesssim 1$. In summary, embryos have $\lambda \approx 10^{-2}$, ordinary cells have $\lambda \approx 10^{-1}$, and cells with problems have $\lambda \approx 1$.

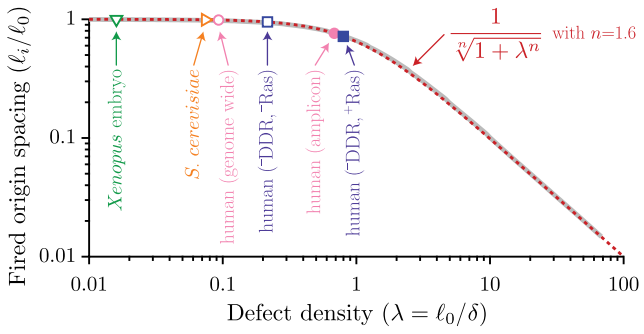

FIG. 3 (color online). Interorigin distance vs defect density for $\tau \to \infty$. The solid gray line is the numerical solution to the scaled REs. Symbols indicate estimates from Table I.

Correlation between fork velocity and origin spacing.— Experimentally, average fork velocities appear to be positively correlated with replication origin spacings [24]. Fork-velocity measurements are typically obtained from pulse-labeling experiments, where replication is observed by incorporating fluorescent markers during the synthesis of DNA [11]. The markers are added to a population of cells with asynchronous cell cycles for a fixed time and then flushed. The reported fork velocity is the observed genomic length of incorporated fluorescent bases divided by the labeling time. The integrated contribution of moving and stalled forks implies that the measured fork velocity corresponds to an effective replication velocity, v_{eff}:

$$\frac{v_{\text{eff}}}{v} = \frac{1}{C}\int_0^\infty \frac{\rho(t)}{\rho(t) + \rho'(t)}[1 - f(t)]dt, \qquad (4)$$

where the normalization constant $C = \int_0^\infty [1 - f(t)]dt$ and the weighting factor $1 - f$ accounts for the asynchrony of cell cycles during the labeling period [11]. Cells that begin replication when fluorescent markers are incorporated have more unreplicated DNA available for labeling than cells that have nearly ended. Thus, each cell contributes to the fork-velocity measurements in proportion to its unreplicated DNA fraction, $1 - f$.

Figure 4 presents the average interorigin distance as a function of the effective velocity for arbitrary repair time. All solutions fall in a narrow region below the line $\ell_i/\ell_0 = \sqrt{v_{\text{eff}}/v}$ (the $\tau \to 0$ solution). Thus, even though initiation

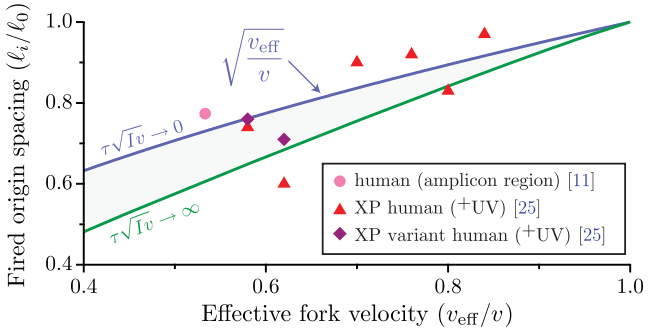

FIG. 4 (color online). Interorigin distance vs effective velocity. The two solid lines are the numerical solutions to the scaled REs when $\tau/t_0 \to 0$ and $\tau/t_0 \to \infty$. Symbols correspond to measurements on different human cells.

and fork propagation are modeled as independent processes, the effective velocity in the presence of defects is strongly correlated with the initiation rate.

In Fig. 4, we also plot the correlated reductions of fork speed and interorigin distances in the amplicon region, scaling both quantities by their "normal" genome-wide values. We have included data for xeroderma pigmentosum (XP) human cells, which lack the ability to repair UV-induced damage [25]. Exposing such cells to UV light leads to pyrimidine dimers that can block forks [2,26]. Thus one expects, and Fig. 4 confirms, a slowing of average fork speed and a decrease in interorigin spacing.

Points from the three data sets fall close to the expected shaded region. Given the measurement uncertainties and simplifying assumptions in our model (constant I and v), the agreement is satisfactory, and the data are consistent with the effects of fork blocks in an otherwise-unaltered replication scenario. Biologists have often speculated on a link between observed fork slowdown, as observed in measurements of average fork velocities, and fork pauses. Our model implements this link quantitatively, suggesting that a passive control mechanism can explain the correlation between fork speeds and interorigin distances.

Conclusion.—We have introduced a mean-field, rate-equation approach to study DNA replication kinetics in the presence of replication defects and found that fork stalls lead to a crossover threshold as the density of defects increases. Normal human cells operate below this critical point. Our results also suggest that the observed correlation between average fork speed and origin firing density can be due to fork stalls, in cases where the pauses are endogenously or exogenously induced.

We can use our model to estimate other quantities, such as the replication time or the number of origins initiated, or extend it to spatially varying I and v, to allow comparison with experiments that go beyond genome- and cell-cycle-averaged data [27]. We can also incorporate explicit DNA checkpoint response mechanisms by dynamically updating I and v appropriately. For example, we can include the effect of double-strand breaks that result when a moving fork collides with a stalled fork [28]. This last scenario is particularly interesting: if repair of double-strand breaks is limited by the number of "repair agents," our model implies a minimum replication time as the fork speed increases. Perhaps, then, this gives a natural "speed limit" for replication, where going too fast only leads to more fork stalls and double-strand breaks, which slows down replication until repaired.

We thank N. Rhind, F. d'Adda di Fagagna, and A. Cerutti for valuable comments. This work was supported by NSERC (Canada) and by HFSP.

*Corresponding author; johnb@sfu.ca

[1] D. Branzei and M. Foiani, Curr. Opin. Cell Biol. **17**, 568 (2005).
[2] W. K. Kaufmann, Carcinogenesis **31**, 751 (2010).
[3] C. Hensey and J. Gautier, Mechanisms of Development **69**, 183 (1997).
[4] M. M. Vilenchik and A. G. Knudson, Proc. Natl. Acad. Sci. U.S.A. **100**, 12871 (2003).
[5] S. Jun, H. Zhang, and J. Bechhoefer, Phys. Rev. E **71**, 011908 (2005); S. Jun and J. Bechhoefer, Phys. Rev. E **71**, 011909 (2005).
[6] J. Lygeros *et al.*, Proc. Natl. Acad. Sci. U.S.A. **105**, 12 295 (2008).
[7] A. Goldar *et al.*, PLoS ONE **3**, e2919 (2008).
[8] J. J. Blow and X. Q. Ge, EMBO J. **10**, 406 (2009).
[9] R. Meneghini and A. C. de Mello Filho, J. Theor. Biol. **100**, 359 (1983).
[10] B. Michel, G. Grompone, M.-J. Florès, and V. Bidnenko, Proc. Natl. Acad. Sci. U.S.A. **101**, 12 783 (2004).
[11] C. Conti, J. Herrick, and A. Bensimon, Genes, Chromosomes and Cancer **46**, 724 (2007).
[12] D. Shechter, V. Costanzo, and J. Gautier, Nat. Cell Biol. **6**, 648 (2004).
[13] K. Marheineke and O. Hyrien, J. Biol. Chem. **279**, 28 071 (2004).
[14] K. Sekimoto, Physica (Amsterdam) **125A** , 261 (1984); **135**, 328A (1986); Int. J. Mod. Phys. B **5**, 1843 (1991).
[15] We have also derived REs for heterogeneous parameters. Using a more realistic, variable-rate initiation function does not drastically change our observations [M. G. Gauthier and J. Bechhoefer, Phys. Rev. Lett. **102**, 158104 (2009)].
[16] For finite genomes or space-dependent v or I, we define densities of left- and right-moving forks, $\rho_L(t)$ and $\rho_R(t)$.
[17] The coalescence rate is the fork density $\times$ the relative fork velocity ($2v$ for moving-moving, v for moving-stalled), normalized by the probability of not being replicated.
[18] F. C. Frank, J. Cryst. Growth **22**, 233 (1974).
[19] J. Herrick *et al.*, J. Mol. Biol. **320**, 741 (2002).
[20] G. M. Alvino *et al.*, Mol. Cell. Biol. **27**, 6396 (2007).
[21] R. D. Micco *et al.*, Nature (London) **444**, 638 (2006).
[22] The reduced, effective value of λ is $\lambda_{\text{eff}} = \lambda/\sqrt{1 + \lambda t_0/\tau}$.
[23] A. M. Friedel, B. L. Pike, and S. M. Gasser, Curr. Opin. Cell Biol. **21**, 237 (2009).
[24] J. Herrick and A. Bensimon, Chromosoma **117**, 243 (2008).
[25] T. D. Griffiths and S. Y. Ling, Mutat. Res. **218**, 87 (1989); Mutagenesis **6**, 247 (1991).
[26] The probability to bypass a defect, P_{bypass}, can be taken into account using $\delta_{\text{eff}} = \delta/(1 - P_{\text{bypass}})$ instead of δ.
[27] P. Norio *et al.*, Mol. Cell **20**, 575 (2005).
[28] A. Kuzminov, Proc. Natl. Acad. Sci. U.S.A. **98**, 8241 (2001).
[29] Parameters for *Xenopus* embryos from K. Marheineke and O. Hyrien, J. Biol. Chem. **279**, 28 071 (2004).