Organization of DNA replication origin firing in *Xenopus* egg extracts : the role of intra-S checkpoint

Diletta Ciardo[1], Olivier Haccard[1], Hemalatha Narassimprakash[1], Jean-Michel Arbona[4], Olivier Hyrien[2], Benjamin Audit[3], Kathrin Marheineke [1*], Arach Goldar[1*]

***For correspondence:** arach.goldar@cea.fr; kathrin.marheineke@i2bc.paris-saclay.fr

[1]Institute of Integrative Biology of the Cell (I2BC), CNRS, CEA, University Paris Sud, 1, avenue de la Terrasse, 91190 Gif-sur-Yvette, France; [2]Institut de Biologie de l'Ecole normale supérieure (IBENS), Ecole normale supérieure, CNRS, INSERM, PSL Research University, 75005 Paris, France; [3]Univ. Lyon, ENS de Lyon, Univ. Claude Bernard Lyon 1, CNRS, Laboratoire de Physique, F-69342 Lyon, France; [4]Univ. Lyon, ENS de Lyon, Univ. Claude Bernard Lyon 1, CNRS UMR5239, INSERM U1210,46 Allé d'Italie Site Jacques Monod, 69007 Lyon, France

Abstract During cell division, the duplication of the genome starts at multiple positions called replication origins. Origin firing requires the interaction of rate-limiting factors with potential origins during the S(ynthesis)-phase of the cell cycle. Origins fire as synchronous clusters is proposed to be regulated by the intra-S checkpoint. By modelling either the unchallenged or the checkpoint-inhibited replication pattern of single DNA molecules from *Xenopus* sperm chromatin replicated in egg extracts, we demonstrate that the quantitative modelling of data require: 1) a segmentation of the genome into regions of low and high probability of origin firing; 2) that regions with high probability of origin firing escape intra-S checkpoint regulation; 3) that the intra-S checkpoint controls the firing of replication origins in regions with low probability of firing. This model implies that the intra-S checkpoint is not the main regulator of origin clustering. The minimal nature of the proposed model foresees its use to analyse data from other eukaryotic organisms.

Introduction

Eukaryotic genomes are duplicated in a limited time during the S phase of each cell cycle. Replication starts at multiple origins that are activated (fired) at different times in S phase to establish two diverging replication forks that progress along and duplicate the DNA at fairly constant speed until they meet with converging forks originated from flanking origins (***DePamphilis and Bell, 2010***; ***Machida et al., 2005***). The mechanisms that regulate the origin firing timing remain largely unknown (***Raghuraman, 2001***; ***Heichinger et al., 2006***; ***Eshaghi et al., 2007***; ***Baker et al., 2012***; ***Audit et al., 2013***; ***Rhind and Gilbert, 2013***). The core motor component of the replicative helicase, the MCM2-7 complex, is loaded on chromatin from late mitosis until the end of G1 phase as an inactive head-to-head double hexamer (DH) to form a large excess of potential origins (***DePamphilis et al., 2006***; ***Ticau et al., 2015***). During S phase, only a fraction of the MCM2-7 DHs are activated to form a pair of active Cdc45-MCM2-7-GINS (CMG) helicases and establish bidirectional replisomes (***DePamphilis and Bell, 2010***). MCM2-7 DHs that fail to fire are inactivated by forks emanating from neighboring fired origins (***Blow et al., 2011***). Origin firing requires S-phase cyclin-dependent kinase (CDK) and Dbf4-dependent kinase (DDK) activities as well as the CDK targets Sld2 and Sld3 and the

replisome-maturation scaffolds Dpb11 and Sld7 in *S. cerevisiae*. The six initiation factors Sld2, Sld3, Dpb11, Dbf4, Sld7 and Cdc45 are expressed at concentrations significantly lower than the MCM complex and core replisome components, suggesting that they may be rate-limiting for origin firing (***Mantiero et al., 2011***; ***Tanaka et al., 2011***). Among these six factors, Cdc45 is the only one to travel with the replication fork.

DNA replication initiates without sequence specificity in *Xenopus* eggs (***Harland and Laskey, 1980***; ***Méchali and Kearsey, 1984***), egg extracts (***Mahbubani et al., 1992***; ***Hyrien and Méchali, 1992***; ***Carli et al., 2016***, ***2018***) and early embryos (***Hyrien and Méchali, 1993***; ***Hyrien et al., 1995***) (for review see ***Hyrien et al.*** (***2003***)). To understand how a lack of preferred sequences for replication initiation is compatible with a precise S-phase completion time, investigators have studied replication at the single DNA molecule level using the DNA combing technique (***Lucas et al., 2000***; ***Herrick et al., 2000***; ***Blow et al., 2001***; ***Marheineke and Hyrien, 2001***, ***2004***). In contrast to population based approaches (which average replication characteristics, this technique reveals cell-to-cell differences in origin activation important for understanding how genomes are replicated during S-phase) these experiments did not detect a regular spacing of initiation events but revealed that origin firing rate strongly increases from early to late replication intermediates, speeding up late replication stages (***Lucas et al., 2000***; ***Herrick et al., 2000***). An observation that has been also confirmed for many other model organisms, including human cell lines (***Goldar et al., 2009***).

Mathematical modelling based on the assumption (mean-field hypothesis) that the probability of firing of each replication origin can be replaced by the averaged probability of firing calculated over all degree of freedom of origin firing process (MCM2-7 DH density, genomic position, chromatin compaction, nucleosome density, etc) and agremented with the assumption of independent origins and a constant fork speed, allowed the extraction of a time-dependent rate of replication initiation, $I(t)$, from the measured eye lengths, gap lengths and eye-to-eye distances on combed DNA molecules (***Herrick et al., 2002***). The extracted $I(t)$ markedly increased during S phase. Simulations incorporating this extracted $I(t)$ reproduced the mean eye length, gap length and eye-to-eye distance, but the experimental eye-to-eye distance distribution appeared “peakier” than the simulated one (***Hyrien et al., 2003***; ***Jun et al., 2004***). Modulating origin firing propensity by the probability to form loops between forks and nearby potential origins resulted in a better fit to the data without affecting $I(t)$ (***Jun et al., 2004***).

Importantly, experiments revealed that in *Xenopus*, like in other eukaryotes, replication eyes are not homogeneously distributed over the genome but tend to cluster (***Blow et al., 2001***; ***Marheineke and Hyrien, 2004***). First, a weak correlation between the sizes of neighbouring eyes was observed (***Blow et al., 2001***; ***Marheineke and Hyrien, 2004***; ***Jun et al., 2004***), consistent with firing time correlations. Second, more molecules with no or multiple eyes than expected for spatially uniform initiation were observed in replicating DNA (***Marheineke and Hyrien, 2004***). There are two potential, non-exclusive mechanisms for these spatiotemporal correlations. The first one, compatible with a mean-field hypothesis, is that activation of an origin stimulates nearby origins. The second one, no longer consistent with a mean-field hypothesis, is that the genome is segmented into multi-origin domains that replicate at different times in S phase. This second hypothesis has been explored numerically in human and has been shown to be compatible with the universal bell shaped $I(t)$ profile (***Gindin et al., 2014***).

Interestingly, experiments in *Xenopus* egg extracts revealed that intranuclear replication foci labelled early in one S phase colocalized with those labelled early in the next S phase, whereas the two labels did not coincide at the level of origins or origin clusters were examined (***Labit et al., 2008***). Given the different characteristic sizes of timing domains (1-5 Mb) and origin clusters (50-100 kb) in the *Xenopus* system, it is possible that origin correlations reflect both a programmed replication timing of large domains and a more local origin cross-talk within domains.

It is now well accepted that the intra-S phase checkpoint regulates origin firing during both unperturbed and artificially perturbed S phase (***Marheineke and Hyrien, 2004***; ***Ge and Blow, 2010***; ***Guo et al., 2015***; ***Platel et al., 2015***; ***Forey et al., 2020***). DNA replication stress, through the activation

of the S-phase checkpoint kinase Rad53, can inhibit origin firing by phosphorylating and inhibiting Sld3 and Dbf4 (***Zegerman and Diffley, 2010***). The metazoan functional analogue of Rad53 is Chk1. Experiments in human cells under low replication stress conditions showed that Chk1 inhibits the activation of new replication factories while allowing origin firing to continue within active factories (***Ge and Blow, 2010***). Experiments using *Xenopus* egg extracts suggested that the checkpoint mainly adjusts the rate of DNA synthesis by staggering the firing time of origin clusters (***Marheineke and Hyrien, 2004***). Recently, we showed that even during an unperturbed S phase in *Xenopus* egg extracts, Chk1 inhibits origin firing away from but not near active forks (***Platel et al., 2015***). We used our initial model for DNA replication in *Xenopus* egg extracts (***Goldar et al., 2008***) (which combined time-dependent changes in the availability of a limiting replication factor, and a fork-density dependent affinity of this factor for potential origins) to model the regulation of DNA replication by the intra-S checkpoint. To account for the regulation of DNA replication by the intra-S checkpoint, we replaced the dependency of origin firing on fork density by a Chk1-dependent global inhibition of origin firing with local attenuation close to active forks as was proposed in other contexts (***Trenz et al., 2008***; ***Dimitrova and Gilbert, 2000***; ***Thomson et al., 2010***; ***Ge and Blow, 2010***). This model was able to simultaneously fit the $I(f)$ (the rate of origin firing expressed as a function of each molecule's replicated fraction f) of a control and a UCN-01-inhibited Chk1 replication experiment (***Platel et al., 2015***). However, in that work we did not push further the analysis to verify if our model was able to explain simultaneously $I(f)$ (temporal program) and the eye-to-eye distance distribution (spatial program).

In the present work, using numerical simulations, we quantitatively analyse both the temporal and spatial characteristics of genome replication as measured by DNA combing in the in vitro *Xenopus* system. The use of *Xenopus* egg extracts has been proven to study DNA replication in metazoans (***Hoogenboom et al., 2017***). Rooted on experimental data, we build a general and minimal model of DNA replication able to predict its temporal and spatial characteristics either during an unchallenged or a challenged S phase. By analysing the spatio-temporal pattern of DNA replication under intra-S checkpoint inhibition and comparing it to an unchallenged pattern we disentangle the complex role of the intra-S checkpoint for replication origin firing.

Results

Finding the best integrative model of unperturbed S phase

Our previous model (***Platel et al., 2015***) failed to simultaneously reproduce the eye-to-eye distance distribution and the $I(f)$ of the same control experiment (***Figure 1*** a and b). This discrepancy could be explained if initiation events have a strong tendency to cluster (***Blow et al., 2001***; ***Marheineke and Hyrien, 2004***). Clustering produces an excess of small (intra-cluster) and large (inter-cluster) eye-to-eye distances compared to random initiations, but only the former could be detected on single DNA molecules due to finite length (***Marheineke and Hyrien, 2004***). Chk1 action has been proposed to regulate origins clusters (***Ge and Blow, 2010***). However, Chk1 inhibition by UCN-01 did not result in the broader eye-to-eye distribution predicted by random origin firing (***Figure 1*** c and d), suggesting that other mechanisms than intra-S checkpoint are involved in the origin clustering.

We therefore explored the ability of several nested models with growing complexity (designated MM1 to MM4) (Appendix 1). MM1 corresponds to a mean field hypothesis of origin firing : all potential origins have a constant firing probability P_{out} (***Goldar et al., 2008***; ***Gauthier and Bechhoefer, 2009***). MM2 corresponds to MM1 with a local perturbation, whereby the proximity of forks facilitates origin firing (***Jun et al., 2004***; ***Löb et al., 2016***) over a distance d downstream of an active fork where the probability of origin firing is P_{local}. In MM3 origin firing does not follow mean field hypothesis but assumes that the genome can be segmented into regions of high and low probabilities of origin firing (***Gindin et al., 2014***; ***Löb et al., 2016***) as accepted for most eukaryotes (***McCune et al., 2008***; ***Yang et al., 2010***; ***Rhind and Gilbert, 2013***; ***Boulos et al., 2015***; ***Das et al., 2015***; ***Petryk et al., 2016***; ***Siefert et al., 2017***). In this scenario, the probability of origin firing of potential origins located

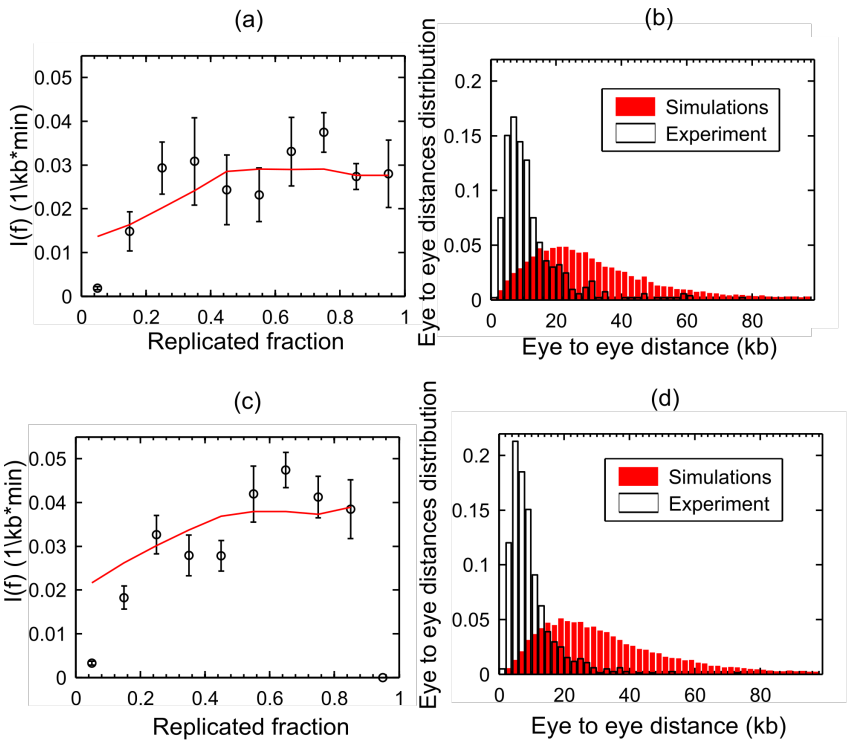


Figure 1. Chk1 does not control origin clustering. The black symbols are experimental data and the red curves are simulations. (a) and (c) Fitting of $I(f)$ data reported in ***Platel et al.*** (***2015***) for control and Chk1 inhibition experiments respectively. (b) and (d) Discrepancy between experimental and simulated distributions of eye-to-eye distances in control and Chk1 inhibition experiments, respectively.

within a fraction θ of the genome, P_{in}, is assumed to be higher than the firing probability P_{out} of potential origins in the complementary fraction $1-\theta$. Lastly, MM4 combines the specific features of MM2 and MM3 into a single model. Furthermore, to verify if the localized nature of potential origins (***Yang et al., 2010***; ***Arbona et al., 2018***) can influence the spatio-temporal program of origin firing, each considered scenario was simulated assuming either a continuous or a discrete distribution of potential origins.

For each model, we coupled dynamic Monte Carlo numerical simulations to a genetic optimization algorithm to find the family of variables that maximized the similarity between the simulated and measured profiles of $I(f)$, replicated fraction of single molecules, global fork density, eye-to-eye distances, gap lengths and eye lengths. MM4 with localized potential origins (***Figure 2***) provided the best fit to the experimental data (Appendix 1, ***Figure 8***). The increase in concordance between MM4 and the data occurs at the expense of increasing the number of parameters, which is justifiable on statistical grounds (Appendix 1,***Table 2***).

Verifying the predictive ability of MM4 model

The real DNA replication process is far more complex than any of the above models. To explore how accurately MM4 can map a more complex process, we built, based on replication process in other eukaryotes (***McCune et al., 2008***; ***Yang et al., 2010***; ***Rhind and Gilbert, 2013***; ***Boulos et al., 2015***; ***Das et al., 2015***; ***Petryk et al., 2016***; ***Siefert et al., 2017***) and our previous model (***Platel et al., 2015***), a more elaborate model (MM5, Appendix 2) to generate *in silico* data with 8%, 19% and 53% global replicated fractions. MM5 assumes that the replication pattern of the genome is reproduced by the coexistence between regions with low probability of origin firing and localised domains with higher probability of origin firing, furthermore MM5 includes explicitly the effect of intra-S checkpoint through supplementary probabilities of origin firing inhibition. However, as during combing experiment the genome is broken randomly into smaller molecules the positional information of each

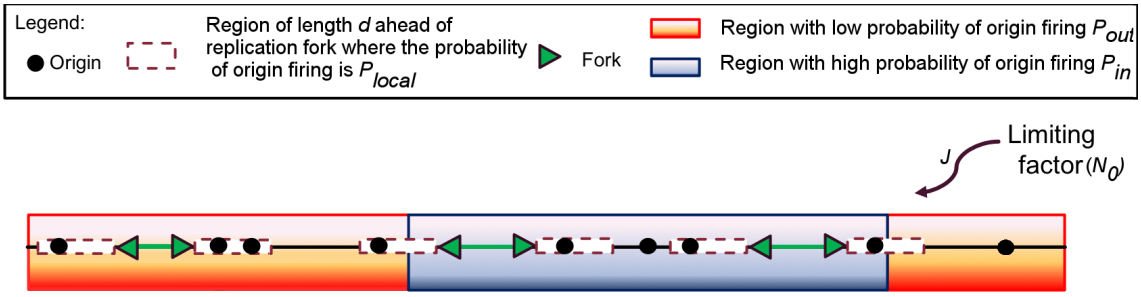


Figure 2. Schematic representation of MM4. Potential replication origins located in a fraction θ of the genome (not necessary contiguous) have a probability of firing P_{in} higher than probability of firing P_{out} of potential origins located in the complementary genome fraction $1-\theta$. The firing of a potential origins requiers its encounter with limiting factors which number $N(t) = N_0 + Jt$ increases as S phase progresses. Potential origins fire with a probability P_{local} over a distance d ahead of a replication fork.

combed single molecule is lost and therefore only genome averaged information can be extracted from a traditional combing experiment. We calculated the expected genome averaged values for each parameter of MM5 (Appendix 2, "Reduction of MM5 to MM4"). Each sample was then fitted with MM4 (Appendix 2 ***Figure 1***, ***Figure 2*** and ***Figure 3***) and we compared the extracted parameters with their expected values after reduction of MM5 to MM4 (***Figure 3***; Appendix 2, ***Table 3***).

For each sample, the mean values of the inferred parameters were statistically similar to the input

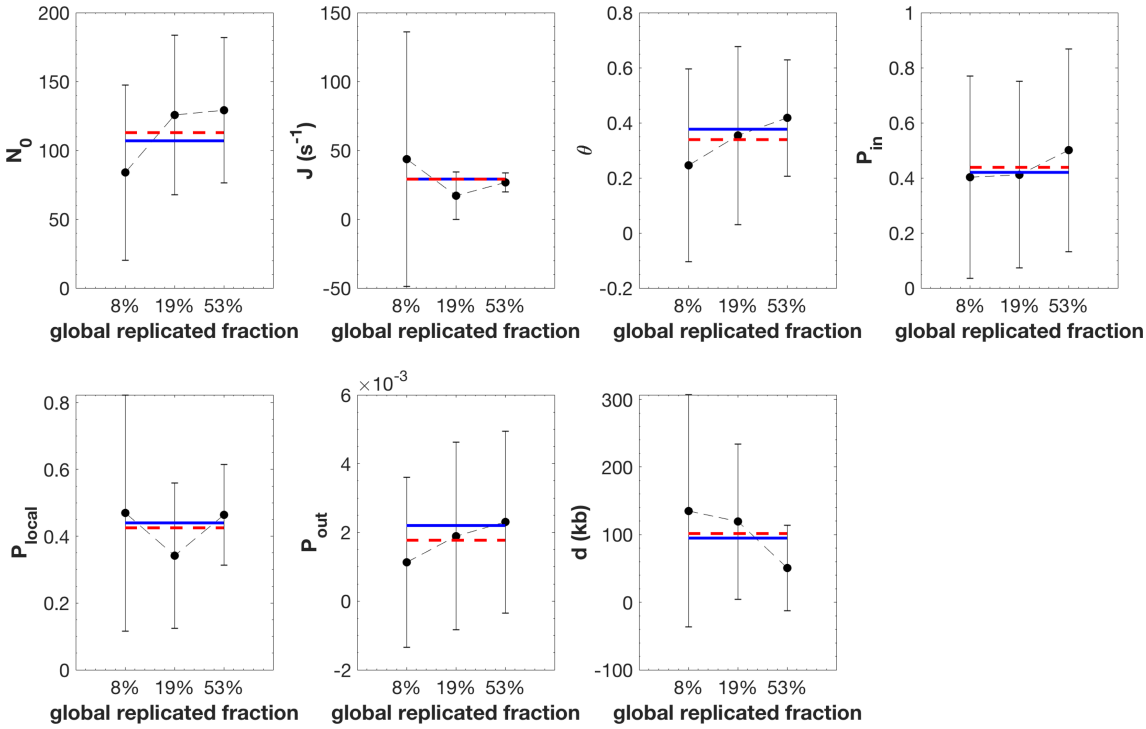


Figure 3. The fitting strategy infers accurately the expected values for the reduced MM5 free parameters. The black circles correspond to the averaged value of the parameter over 100 independent fits and the error bars are the standard-deviations. The solid blue line is the expected value of the parameter as obtained in Appendix 2, ***Table 3***. The red dashed line is the mean value of the parameter obtained by averaging the parameter inferred values over the 3 samples.

ones (Appendix 2, ***Table 3***) and none of the pairwise differences between the predicted parameters values for the 3 considered samples were statistically significant. This demonstrates that our fitting and comparison strategies do not introduce artifactual differences between parameters if their values do not change between different samples (Appendix 2 ***Figure 4***). In conclusion, any

variation in parameter value detected by MM4 when analysing samples at different time points independently can be considered as statistically significant. Therefore, MM4 can adequately model more complex DNA replication dynamics than itself using a reduced number of parameters.

Retrieving the dynamics of an unchallenged S phase using the MM4 model

MM4 faithfully reproduced the temporal and spatial program of DNA replication from unperturbed S phase samples with global replicated fractions of 8%, 19% and 53% (Appendix 1, ***Figure 8***; Appendix 3, ***Figure 1*** and ***Figure 2***). The fitted values of parameters changed as S phase progressed (***Figure 4***). However, only changes in J, θ, P_{out} and d were statistically significant (Appendix 3 ***Figure 3***). In

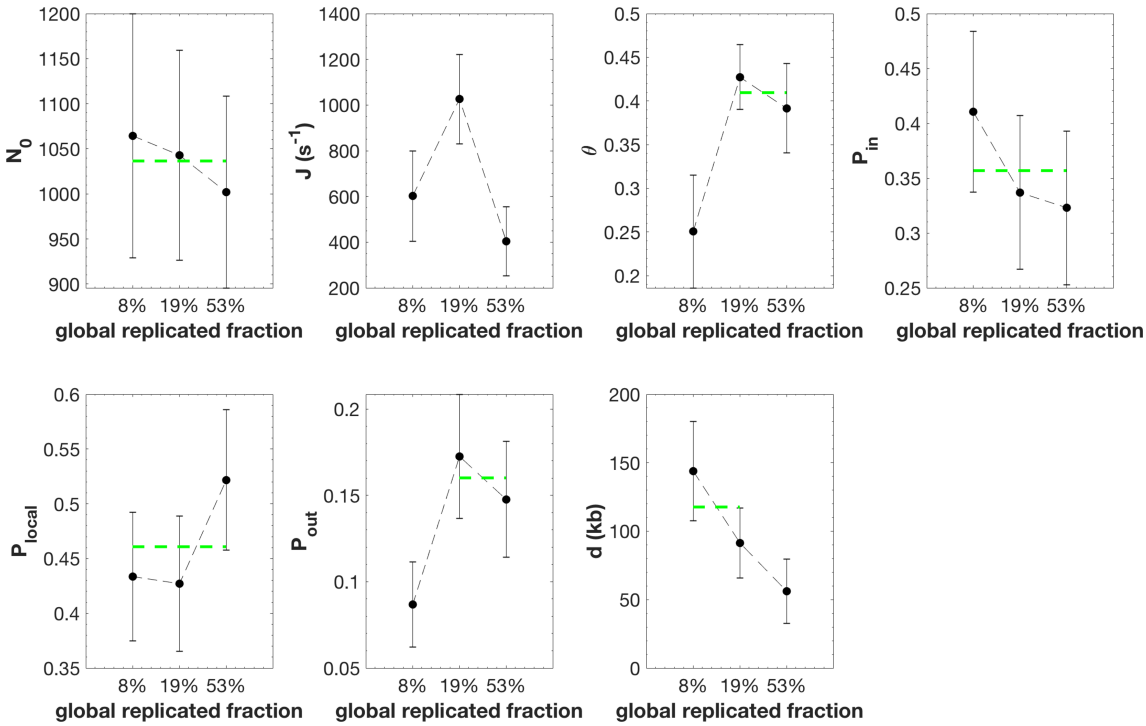


Figure 4. Inferred model parameters by fitting unchallenged S phase data as global replicated fraction increases. The black circles are the averaged value of the parameter over 100 independent fitting processes and the error bars are standard-deviations. The green dashed line is the mean value among consecutive parameters which differences are not statistically significant (Appendix 3 ***Figure 3***).

particular we found that J increased from 8% to 19% replication and then drop back at 53% replication. θ and P_{out} increased only from 8% to 19% replication but not later, while d stayed constant between 8% and 19% replication and decreased at 53% replication.

These observations suggest that during an unchallenged S phase both the fraction (θ) of the genome with high probability of origin firing and the background probability (P_{out}) of origin firing outside that fraction increase as S phase progresses. Interestingly, P_{local} is higher than P_{in} and P_{out}, suggesting that firing of an potential origin significantly favours the firing of nearby potential origins over a distance d, compatible with a chromatin looping process (***Löb et al., 2016***). This fork-related firing process is consistent with the observation that nearby origins tend to fire at similar times, which has been proposed to result from a different regulation of nearby and distant origins by Chk1 (***Ge and Blow, 2010***; ***Platel et al., 2015***).

Modeling DNA replication under Chk1 inhibition

To decipher the regulation of origin firing by Chk1, we examined if the MM4 model could also reproduce the replication program observed when the intra-S phase checkpoint was inhibited by the specific Chk1 inhibitor UCN-01. We analyzed combed fibres from a replicated sample in

the presence of UCN-01 (replicated fraction 22%) that had spent the same interval of time in S phase as the control sample (global replicated fraction of 8%). The MM4 model reproduced the experimental observations very well (Appendix 3, ***Figure 4***, $GoF_{global} = 0.85$). The three

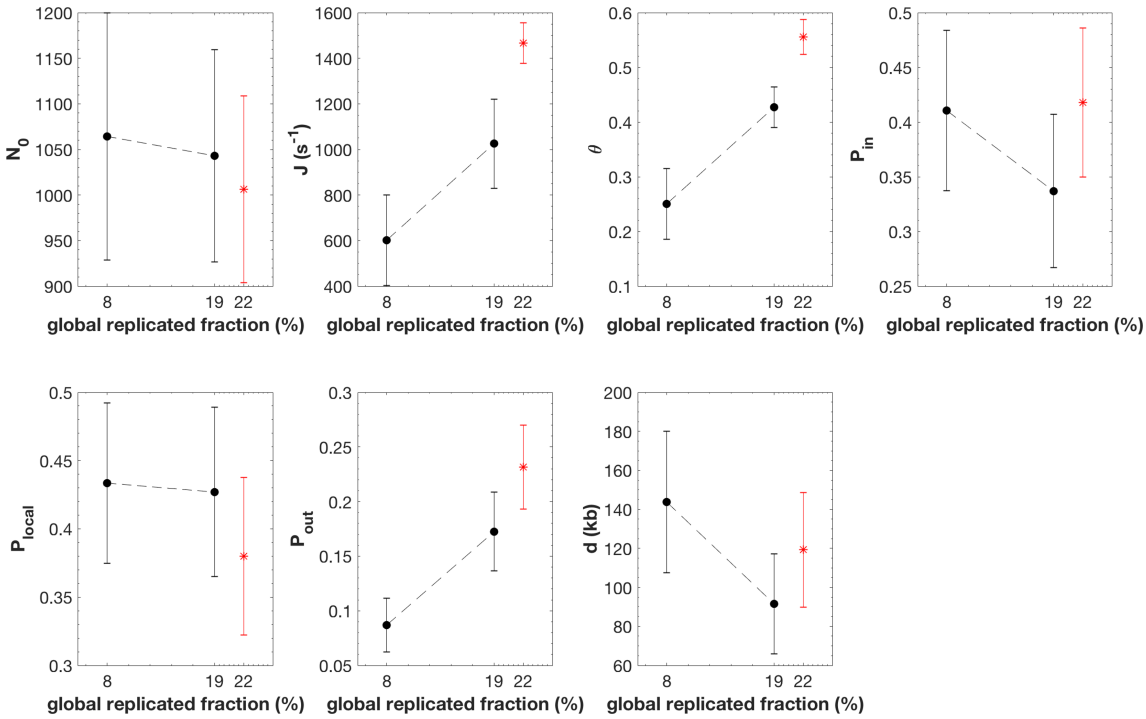


Figure 5. J, θ, and the P_{out} are the only parameters that change when comparing unchallenged and Chk1 inhibited S phase The black circle is the averaged value of the parameter over 100 independent fitting processes of unchallenged S phase and the error bars are standard-deviations. The red star is the averaged value of the parameter over 100 independent fitting processes of Chk1 inhibited sample and the error bars represent the standard-deviations.

parameters J, θ, and the P_{out} were significantly higher in the UCN-01 treated sample than in the control samples with either the same harvesting time or a similar replicated fraction (22% and 19%, respectively) (***Figure 5*** and Appendix 3 ***Figure 5***). The other parameters were unchanged compared to both control samples. These results suggest that upon Chk1 inhibition (i) a fraction θ of the genome, where initiation probability is high, increases during S phase; (ii) the probability of origin firing is insensitive to Chk1 within this fraction (P_{in} is unaltered) but is increased in the rest of the genome (P_{out} is increased) ; (iii) the import/activation rate of the limiting factor, J, is increased, while the starting number of factors, N_0, is unaffected. As was expected, MM4 detected that Chk1 inhibition by UCN-01 increased origin firing (***Platel et al., 2015***; ***Syljuasen et al., 2005***; ***Guo et al., 2015***; ***Michelena et al., 2019***; ***Pommier and Kohn, 2003***; ***Deneke et al., 2016***).
In conclusion , the level of active Chk1 appears to regulate the kinetics of S phase progression (i) by limiting the genome fraction that escapes its inhibitory action, (ii) by down regulating the probability of origin firing outside this fraction (***Syljuasen et al., 2005***; ***Maya-Mendoza et al., 2007***; ***Guo et al., 2015***; ***Michelena et al., 2019***), and (iii) by controlling the import/activation rate of limiting firing factors (***Guo et al., 2015***). However, no significant differences in the strength of origin regulation by nearby forks (P_{local}) was observed after Chk1 inhibition, suggesting that this local action is not mediated by Chk1 (***Trenz et al., 2008***; ***Ge and Blow, 2010***).

Discussion

We explored several biologically plausible scenarios to understand the spatio-temporal organization of replication origin firing in *Xenopus* egg extracts. We used a quantitative approach to objectively

discriminate which model best reproduced the genomic distributions of replication tracks as analyzed by DNA combing at different stages of S phase. We found that model MM4 with discrete potential origins best reproduced the experimental data with a minimal number of adjustable parameters. This model combines five assumptions (**Herrick et al., 2002**; **Goldar et al., 2008**; **Gauthier and Bechhoefer, 2009**; **Blow and Ge, 2009**; **Sekedat et al., 2010**; **Yang et al., 2010**; **Platel et al., 2015**; **Löb et al., 2016**; **Gindin et al., 2014**; **Arbona et al., 2018**): 1) origin firing is stochastic, 2) the availability of a rate-limiting firing factor captures the essential dynamics of the complex network of molecular interactions required for origin firing, 3) the speed of replication forks is constant 4) origins fire in a domino-like fashion in the proximity of active forks (**Guilbaud et al., 2011**; **Löb et al., 2016**) ; 5) the probability of origin firing is heterogeneous along the genome (**Yang et al., 2010**; **Gindin et al., 2014**).

We used MM4 to model DNA combing data from *Xenopus* egg extracts in presence or absence of intra-S checkpoint inhibition. In both conditions, this model was able to match the experimental data in a satisfactory manner. Furthermore, the inferred parameters values indicated that the global probability of origin firing and the rate of activation/import of the limiting firing factor (J) were increased after Chk1 inhibition by UCN-01(**Pommier and Kohn, 2003**; **Seiler et al., 2007**; **Guo et al., 2015**). Importantly, this model assumes a heterogeneous probability of origin firing and suggests that Chk1 exerts a global origin inhibitory action during unperturbed S phase (**Platel et al., 2015**). On the other hand, the constancy of the initial number of limiting factors N_0 in the presence or absence of UCN-01 suggests that Chk1 does not actively control origins before S phase actually starts (**Lupardus et al., 2002**; **Stokes et al., 2002**; **Forey et al., 2020**). These observations indicate that MM4 can deliver a reliable, minimally complex picture of origin firing regulation in *Xenopus* egg extracts.

The global inhibition of origin firing by Chk1

We previously showed that Chk1 is active and limits the firing of some potential origins in an unperturbed S phase (**Platel et al., 2015**). Therefore, the earliest origins must be immune to Chk1 inhibition while later potential origins are strongly inhibited. The comparison between the modelling of Chk1 inhibition and of unperturbed S phase data suggests that i) the probability of origin firing is reduced by active Chk1 in a fraction $1-\theta$ of the genome, ii) in this Chk1-sensitive fraction the probability of origin firing increases as S phase progresses and iii) the probability of origin firing is unaffected by Chk1 inhibition within the Chk1-immune, θ fraction of the genome. Therefore, this model supports the idea that at the start of S phase, some origins fire unimpeded by Chk1, whereas others remain silent. The latter only becomes progressively relieved from Chk1 inhibition as S phase progresses. Indeed, recent works in cultured mammalian cells (**Moiseeva et al., 2019**), *Drosophila* (**Deneke et al., 2016**) and *Xenopus* (**Krasinska et al., 2008**) showed that in unperturbed S phase the global origin firing inhibitory effect (by Chk1 and Rif1) is reduced as S phase progresses. Interestingly, a recent study in unperturbed yeast cells suggests that dNTPs are limiting at the entry into S phase, so that, similar to *Xenopus* (**Zou, 2007**), the firing of the earliest origins creates a replication stress that activates the Rad53 checkpoint which prevents further origin firing. Rad53 activation also stimulates dNTP synthesis, which in turn down regulates the checkpoint and allows later origin firing (**Forey et al., 2020**). However, it remains uncertain if this feed-back loop does also exist in *Xenopus* egg extracts which contain an abundant pool of dNTPs.

A key mechanism of our model is the enhancement of origin firing close to active forks. The necessity to introduce this mechanism supports the domino-like view of DNA replication progression (**Guilbaud et al., 2011**; **Löb et al., 2016**). It was previously shown in *Xenopus* egg extracts that the probability of origin firing could depend on the distance between left and right approaching forks (**Jun et al., 2004**). While this could in principle reflect an origin firing exclusion zone ahead of forks (**Lucas et al., 2000**; **Löb et al., 2016**), our model did not allow for a negative P_{local}. Other proposed mechanisms for origin clustering include the relief of Chk1 inhibition ahead of active forks by checkpoint recovery kinase polo like kinase 1 (Plk1) (**Trenz et al., 2008**; **Platel et al., 2015**).

However, we find that the range, d, and the strength, P_{local}, of origin stimulation by nearby forks, were both insensitive to checkpoint inhibition (***Figure 5*** a and b). Other potential mechanisms such as propagation of a supercoiling wave ahead of forks may better explain this insensitivity to Chk1 inhibition (***Achar et al., 2020***).

Heterogeneous probability of origin firing

In this model, the origin firing process in *Xenopus* egg extracts is not fiably described by a mean-field approximation. In other words, the probability of origin firing is heterogeneous along the genome. Based on this hypothesis, one important outcome of our study is that the genome can be segmented into domains where origin firing probability is either high and immune to Chk1 inhibition or subjected to a tight Chk1 control that attenuates as S phase progresses. This picture challenges the common view that the embryonic *Xenopus in vitro* system would lack the temporal regulation by the intra-S checkpoint at the level of large chromatin domains in contrast to findings in somatic vertebrate cells where Chk1 controls cluster or replication foci activation (***Maya-Mendoza et al., 2007***). However, observations of replicating nuclei in *Xenopus* system have shown that early replication foci are conserved in successive replication cycles, supporting the heterogeneous domain hypothesis (***Labit et al., 2008***). Furthermore, we found that the fraction of the genome covered by these domains increases and that the inhibitory action of Chk1 decreases over time during an unperturbed S phase (***Figure 4*** and ***Figure 5***), consistent with the idea that as S phase progresses more regions of the genome evade the checkpoint inhibition of origins. By comparing samples that have spent the same time interval in S phase or that have reached the same replicated fraction in the absence and presence of UCN-01 (***Figure 5***), we noticed that the probability of origin firing in the Chk1-immune domains (P_{in}) did not change upon Chk1 inhibition. This further suggests that these domains escape actually the regulation of origin firing by Chk1 that rules the rest of the genome.

All together the results of our modelling approach and the existing literature suggest that in the *Xenopus* system the position of early replicating, Chk1-immune domains is conserved in individual nucleus. However, there is no experimental or numerical evidence that the positions of these domains are conserved in a population of nuclei. Assuming that the position of these domains changes randomly from one nucleus to another would result in a flat mean replication timing pattern and involves that each nucleus has its specific replication regulation process. While we cannot reject such a hypothesis objectively, the recent report of a structured replication timing program in zebrafish early embryos (***Siefert et al., 2017***) encourage us to assume that in *Xenopus* early embryos the position of early replication domaines are conserved from one nucleus to an other. Thus, we propose that the mean replication timing pattern of *Xenopus* sperm nuclei in egg extracts is not flat but is structured similarly to other eukaryotic systems (***Baker et al., 2012***; ***Rhind and Gilbert, 2013***; ***Boulos et al., 2015***).

The generality of assumptions and conclusions of our model suggest that it can be used to analyze the dynamics of S phase and its regulation by the intra-S phase checkpoint in other organisms.

Methods and Materials

Monte Carlo simulation of DNA replication process.

A dynamical Monte Carlo method was used to simulate the DNA replication process as before (***Goldar et al., 2008***). We simulate the replicating genome as a one-dimensional lattice of $L = 10^6$ blocks of value 1 for replicated and 0 for unreplicated state, respectively. To match the spatial resolution of DNA combing experiments each block represents 1kb. After one round of calculation an existing replication track grows in a symmetric manner by 2 blocks. Considering that the fork speed $v = 0.5\ kb.min^{-1}$ is constant, one round of calculation corresponds to 2 minutes. In the continuous case we assume that the potential replication origins are continuously distributed on the genome with an average density of one potential origin per 1kb (1 block); in the discrete case we

assume that potential origins are randomly distributed along the genome with an average density of one potential origin per 2.3 kb (***Edwards et al., 2002***). In both cases origins fire stochastically. Origin firing requires an encounter with a trans-acting factor which number $N(t)$ increases as S phase progresses with a rate J, $N(t) = N_0 + Jt$. If an encounter produces an origin firing, the trans-acting factor is sequestrated by replication forks and hence the number of available trans-acting factor is $N_f(t) = N(t) - N_b(t)$, where $N_b(t)$ is the number of bound factors. To ensure that origins do not re-fire during one cycle and are inactivated upon passive replication, only "0" blocks are able to fire. At each round of calculus, each block is randomly assigned 2 independent values between 0 and 1. The first one is compared to θ to decide whether the block belongs to the θ or $1-\theta$ fraction of the genome. The second one to P_{in} or P_{out}, respectively, to decide whether the block may fire. In total, M "0" blocks ($M \leq L$) with value strictly smaller than their reference probability may fire. If $M \leq N_f(t)$ all M blocks may fire, otherwise $N_f(t)$ blocks may fire. Furthermore in MM2 and MM4, we consider that the probability of origin firing P_{local} may be increased downstream of a replication fork over a distance d_{fork}. The trans-acting factors sequestered by forks are released and are made available for new initiation events when forks meet.

Measuring: the replicated fraction $f(t)$, the rate of origin firing $I(t)$, fork density $N_{fork}(t)$, eye-to-eye, eye and gap length distributions.

The genome is represented as an one-dimensional lattice of 10^6 elements $x_i \in \{0, 1\}$. At each round of calculation the replicated fraction is calculated as $f(t) = \langle x \rangle_i$ corresponding to the average value of x_i over the genome.

The rate of origin firing per length of unreplicated genome per time unit (3 min) is calculated at each round of calculation, by counting the number of newly created "1" blocks, N_1 and $I(t) = \frac{N_1}{(1-f(t))L\Delta t}$ where $\Delta t = 3\ min$ and $L = 10^6$. The density of replication forks is calculated at each round of calculation by counting the number of "01" tracks, N_{left}, and "10" tracks, N_{right} and $N_{forks}(t) = \frac{N_{right}+N_{left}}{L}$. The distributions of eye-to-eye distances, eye lengths and unreplicated gap sizes are then computed from the distribution of "0" and "1" tracks after reshaping the data (see below).

Comparing experimental and numerical data.

The simulation results were compared to the DNA combing data from Platel *et al.* (***Platel et al., 2015***). The fluorescence intensities for total DNA and replicated tracks of each fiber were measured and binarized on a Matlab® platform by using a thresholding algorithm. The threshold value was chosen to minimize the difference between the replicated fraction measured by $\alpha 32$P-dATP incorporation and by DNA combing. Replicated tracks larger than 1kb were scored as eyes. Gaps were considered significant if > 1kb, otherwise the two adjacent eyes were merged. The eyes whose lengths span from 1 to 3 kb were considered as new origin firing events. The time interval in which these new detectable events can occur was calculated as $\Delta t = 3min$ assuming a constant replication fork velocity of $v \approx 0.5\ kb.min^{-1}$. This data reshaping protocol was also applied to simulated DNA molecules, in order to match the spatial and temporal resolutions between the experimental and simulated data. The global replicated fraction of each sample was computed as the sum of all eye lengths divided by the sum of all molecule lengths. To minimize finite molecule length effects in comparisons between data and simulations, the experimental molecule length distribution was normalised and considered as probability density of molecule length in the sample and used to weight the random shredding of the simulated genome at each time (***Figure 6***). The global replication fraction of simulated cut molecules was calculated. Only molecules from the simulation time that had the same global replication fraction as the experimental sample were further considered.

Molecules were sorted by replicated fraction $f(t)$. The rate of origin firing and fork density were calculated for each molecule as a function of $f(t)$ ($I(f)$ and $N_{fork}(f)$, respectively) for both simulated and experimental data. The experimental $I(f)$, $N_{fork}(f)$, eye-to-eye distances, eye and

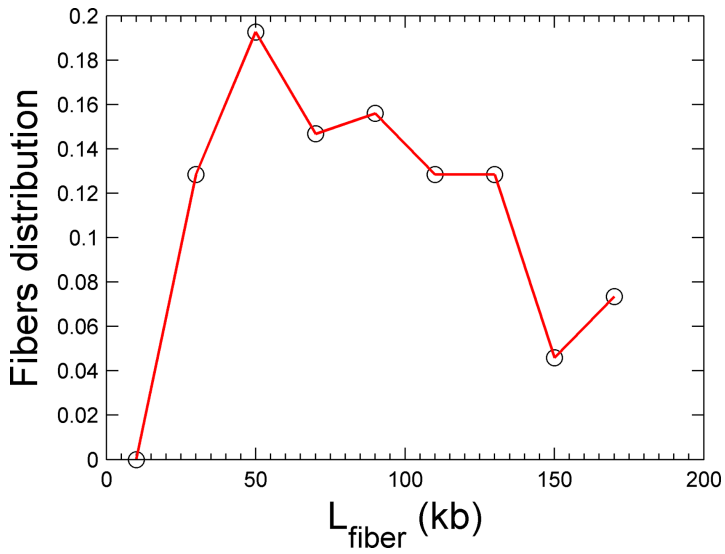


Figure 6. Molecular length distribution (global replicated fraction of 8%). The black open circles are the experimentally measured and the red curve is the simulated cut molecular length distributions, respectively .

gap length distributions were computed as the averaged value of three independent experiments.

Modeling experimental data: parameters optimization.

To estimate the parameters of the model, we fitted the six experimental observables ($I(f)$, $N_{fork}(f)$, replicated fiber, eye-to-eye distances, eye and gap length distribution) using a genetic optimization algorithm (Matlab®). The fitness function was defined as the sum of the square of the differences between experimental and simulated data curves divided by the squared mean of the experimental data curve. The genetic optimization algorithm was set over three subpopulations of 20 individuals with a migration fraction of 0.1 and a migration interval of 5 steps. Each individual defined a set of variables for the simulation and the variables were chosen within the range reported in ***Table 1*** for the model that best fit the data. At each generation, 3 elite children were selected for the next generation. The rest of the population corresponds to a mixture between 60% of children obtained after a scattered crossover between two individuals selected by roulette wheel selection and 40% of children obtained by uniform mutation with a probability of 0.2, leading to a variability of 8%. The genetic algorithm was stopped after 50 generations corresponding to the convergence of the optimization method. As the size of variable space is unknown, we considered a large domain of validity for the variables. This has as an effect to reduce the probability that the optimization process reaches a unique global minimum. For this reason we repeat the genetic optimization method 100 times independently over each data set and consider for each optimization round only the best elite individual.

Table 1. Lower and upper bounds of adjustable variables.

Variable	Lower bound	Upper bound	Significance
N_0	1	2000	Initial number of limiting-factor
$J\ (s^{-1})$	0	4000	Rate at which the number of limiting-factor increases
P_{out}	0	1	Probability of origin firing in the $1-\theta$ fraction
P_{in}	0	1	Probability of origin firing in the θ fraction
P_{local}	0	1	Probability of origin firing ahead of an active replication fork over a distance d
θ	0	1	Fraction of genome where the probability of origin firing is P_{in}
$d\ (kb)$	0	1000	Distance over which a fork acts on the probability of origin firing

Acknowledgments

This article is dedicated to the memory of our colleague and friend Alain Arneodo, who passed away during its elaboration and writing. The authors acknowledge Alain's enthusiasm and constant support. This work was supported by the Fondation pour la Recherche Médicale [FRM DEI201512344404], the Centre National de Recherche Scientifique (CNRS), the department of genome biology of I2BC, by a PhD fellowship of IdEX Paris-Saclay university, the Commissariat à l'énergie atomique (CEA) and the Cancéropole Ile-de-France [PLBIO16-302].

References

Achar YJ, Adhil M, Choudhary R, Gilbert N, Foiani M. Negative Supercoil at Gene Boundaries Modulates Gene Topology. Nature. 2020 Jan; 577(7792):701–705. doi: 10.1038/s41586-020-1934-4.

Arbona JM, Goldar A, Hyrien O, Arneodo A, Audit B. The Eukaryotic Bell-Shaped Temporal Rate of DNA Replication Origin Firing Emanates from a Balance between Origin Activation and Passivation. eLife. 2018 Jun; 7:e35192. doi: 10.7554/eLife.35192.

Audit B, Baker A, Chen CL, Rappailles A, Guilbaud G, Julienne H, Goldar A, d'Aubenton-Carafa Y, Hyrien O, Thermes C, Arneodo A. Multiscale Analysis of Genome-Wide Replication Timing Profiles Using a Wavelet-Based Signal-Processing Algorithm. Nat Protoc. 2013 Jan; 8(1):98–110. doi: 10.1038/nprot.2012.145.

Baker A, Audit B, Chen CL, Moindrot B, Leleu A, Guilbaud G, Rappailles A, Vaillant C, Goldar A, Mongelard F, d'Aubenton-Carafa Y, Hyrien O, Thermes C, Arneodo A. Replication Fork Polarity Gradients Revealed by Megabase-Sized U-Shaped Replication Timing Domains in Human Cell Lines. PLoS Comput Biol. 2012 Apr; 8(4). doi: 10.1371/journal.pcbi.1002443.

Bevington P, Robinson DK. Data Reduction and Error Analysis for the Physical Sciences. McGraw-Hill Education; 2003.

Blow JJ, Ge XQ. A Model for DNA Replication Showing How Dormant Origins Safeguard against Replication Fork Failure. EMBO Rep. 2009 Apr; 10(4):406–412. doi: 10.1038/embor.2009.5.

Blow JJ, Ge XQ, Jackson DA. How Dormant Origins Promote Complete Genome Replication. Trends Biochem Sci. 2011 Aug; 36(8):405–414. doi: 10.1016/j.tibs.2011.05.002.

Blow JJ, Gillespie PJ, Francis D, Jackson DA. Replication Origins in XenopusEgg Extract Are 5–15 Kilobases Apart and Are Activated in Clusters That Fire at Different Times. J Cell Biol. 2001 Jan; 152(1):15–26.

Boulos RE, Drillon G, Argoul F, Arneodo A, Audit B. Structural Organization of Human Replication Timing Domains. FEBS Letters. 2015 Oct; 589(20, Part A):2944–2957. doi: 10.1016/j.febslet.2015.04.015.

Carli FD, Gaggioli V, Millot GA, Hyrien O. Single-Molecule, Antibody-Free Fluorescent Visualisation of Replication Tracts along Barcoded DNA Molecules. Int J Dev Biol. 2016 May; 60(7-8-9):297–304. doi: 10.1387/ijdb.160139oh.

Carli FD, Menezes N, Berrabah W, Barbe V, Genovesio A, Hyrien O. High-Throughput Optical Mapping of Replicating DNA. Small Methods. 2018; 2(9):1800146. doi: 10.1002/smtd.201800146.

Das SP, Borrman T, Liu VWT, Yang SCH, Bechhoefer J, Rhind N. Replication Timing Is Regulated by the Number of MCMs Loaded at Origins. Genome Res. 2015 Dec; 25(12):1886–1892. doi: 10.1101/gr.195305.115.

Deneke VE, Melbinger A, Vergassola M, Di Talia S. Waves of Cdk1 Activity in S Phase Synchronize the Cell Cycle in Drosophila Embryos. Developmental Cell. 2016 Aug; 38(4):399–412. doi: 10.1016/j.devcel.2016.07.023.

DePamphilis M, Bell SD. Genome Duplication. London ; New York: Garland Science; 2010.

DePamphilis ML, Blow JJ, Ghosh S, Saha T, Noguchi K, Vassilev A. Regulating the Licensing of DNA Replication Origins in Metazoa. Current Opinion in Cell Biology. 2006 Jun; 18(3):231–239. doi: 10.1016/j.ceb.2006.04.001.

Dimitrova DS, Gilbert DM. Temporally Coordinated Assembly and Disassembly of Replication Factories in the Absence of DNA Synthesis. Nat Cell Biol. 2000 Oct; 2(10):686–694. doi: 10.1038/35036309.

Edwards MC, Tutter AV, Cvetic C, Gilbert CH, Prokhorova TA, Walter JC. MCM2–7 Complexes Bind Chromatin in a Distributed Pattern Surrounding the Origin Recognition Complex in *Xenopus* Egg Extracts. J Biol Chem. 2002 Sep; 277(36):33049–33057. doi: 10.1074/jbc.M204438200.

Eshaghi M, Karuturi RKM, Li J, Chu Z, Liu ET, Liu J. Global Profiling of DNA Replication Timing and Efficiency Reveals That Efficient Replication/Firing Occurs Late during S-Phase in S. Pombe. PLoS One. 2007 Aug; 2(8). doi: 10.1371/journal.pone.0000722.

Forey R, Poveda A, Sharma S, Barthe A, Padioleau I, Renard C, Lambert R, Skrzypczak M, Ginalski K, Lengronne A, Chabes A, Pardo B, Pasero P. Mec1 Is Activated at the Onset of Normal S Phase by Low-dNTP Pools Impeding DNA Replication. Molecular Cell. 2020 Mar; doi: 10.1016/j.molcel.2020.02.021.

Gauthier MG, Bechhoefer J. Control of DNA Replication by Anomalous Reaction-Diffusion Kinetics. Phys Rev Lett. 2009 Apr; 102(15):158104. doi: 10.1103/PhysRevLett.102.158104.

Ge XQ, Blow JJ. Chk1 Inhibits Replication Factory Activation but Allows Dormant Origin Firing in Existing Factories. The Journal of Cell Biology. 2010 Dec; 191(7):1285–1297. doi: 10.1083/jcb.201007074.

Gindin Y, Valenzuela MS, Aladjem MI, Meltzer PS, Bilke S. A Chromatin Structure-Based Model Accurately Predicts DNA Replication Timing in Human Cells. Mol Syst Biol. 2014 Mar; 10(3):722. doi: 10.1002/msb.134859.

Goldar A, Labit H, Marheineke K, Hyrien O. A Dynamic Stochastic Model for DNA Replication Initiation in Early Embryos. PLoS ONE. 2008 Aug; 3(8):e2919. doi: 10.1371/journal.pone.0002919.

Goldar A, Marsolier-Kergoat MC, Hyrien O. Universal Temporal Profile of Replication Origin Activation in Eukaryotes. PLOS ONE. 2009 Jun; 4(6):e5899. doi: 10.1371/journal.pone.0005899.

Guilbaud G, Rappailles A, Baker A, Chen CL, Arneodo A, Goldar A, d'Aubenton-Carafa Y, Thermes C, Audit B, Hyrien O. Evidence for Sequential and Increasing Activation of Replication Origins along Replication Timing Gradients in the Human Genome. PLOS Computational Biology. 2011 Dec; 7(12):e1002322. doi: 10.1371/journal.pcbi.1002322.

Guo C, Kumagai A, Schlacher K, Shevchenko A, Shevchenko A, Dunphy WG. Interaction of Chk1 with Treslin Negatively Regulates the Initiation of Chromosomal DNA Replication. Molecular Cell. 2015 Feb; 57(3):492–505. doi: 10.1016/j.molcel.2014.12.003.

Harland RM, Laskey RA. Regulated Replication of DNA Microinjected into Eggs of Xenopus Laevis. Cell. 1980 Oct; 21(3):761–771. doi: 10.1016/0092-8674(80)90439-0.

Heichinger C, Penkett CJ, Bähler J, Nurse P. Genome-wide Characterization of Fission Yeast DNA Replication Origins. EMBO J. 2006 Nov; 25(21):5171–5179. doi: 10.1038/sj.emboj.7601390.

Herrick J, Stanislawski P, Hyrien O, Bensimon A. Replication Fork Density Increases during DNA Synthesis in X. Laevis Egg extracts11Edited by M. Yaniv. Journal of Molecular Biology. 2000 Jul; 300(5):1133–1142. doi: 10.1006/jmbi.2000.3930.

Herrick J, Jun S, Bechhoefer J, Bensimon A. Kinetic Model of DNA Replication in Eukaryotic Organisms. Journal of Molecular Biology. 2002 Jul; 320(4):741–750. doi: 10.1016/S0022-2836(02)00522-3.

Hoogenboom WS, Klein Douwel D, Knipscheer P. Xenopus Egg Extract: A Powerful Tool to Study Genome Maintenance Mechanisms. Developmental Biology. 2017 Aug; 428(2):300–309. doi: 10.1016/j.ydbio.2017.03.033.

Hyrien O, Méchali M. Plasmid Replication in Xenopus Eggs and Egg Extracts: A 2D Gel Electrophoretic Analysis. Nucleic Acids Res. 1992 Apr; 20(7):1463–1469.

Hyrien O, Méchali M. Chromosomal Replication Initiates and Terminates at Random Sequences but at Regular Intervals in the Ribosomal DNA of Xenopus Early Embryos. The EMBO Journal. 1993 Dec; 12(12):4511–4520. doi: 10.1002/j.1460-2075.1993.tb06140.x.

Hyrien O, Marheineke K, Goldar A. Paradoxes of Eukaryotic DNA Replication: MCM Proteins and the Random Completion Problem. Bioessays. 2003; 25(2):116–125. doi: 10.1002/bies.10208.

Hyrien O, Maric C, Méchali M. Transition in Specification of Embryonic Metazoan DNA Replication Origins. Science. 1995 Nov; 270(5238):994–997. doi: 10.1126/science.270.5238.994.

Jun S, Herrick J, Bensimon A, Bechhoefer J. Persistence Length of Chromatin Determines Origin Spacing in Xenopus Early-Embryo DNA Replication: Quantitative Comparisons between Theory and Experiment. Cell Cycle. 2004 Feb; 3(2):223–229.

Krasinska L, Besnard E, Cot E, Dohet C, Méchali M, Lemaitre JM, Fisher D. Cdk1 and Cdk2 Activity Levels Determine the Efficiency of Replication Origin Firing in Xenopus. The EMBO Journal. 2008 Mar; 27(5):758–769. doi: 10.1038/emboj.2008.16.

Labit H, Perewoska I, Germe T, Hyrien O, Marheineke K. DNA Replication Timing Is Deterministic at the Level of Chromosomal Domains but Stochastic at the Level of Replicons in Xenopus Egg Extracts. Nucleic Acids Research. 2008 Aug; 36(17):5623–5634. doi: 10.1093/nar/gkn533.

Ljung L. System Identification: Theory for the User. Pearson Education; 1998.

Löb D, Lengert N, Chagin VO, Reinhart M, Casas-Delucchi CS, Cardoso MC, Drossel B. 3D Replicon Distributions Arise from Stochastic Initiation and Domino-like DNA Replication Progression. Nature Communications. 2016 Apr; 7:11207. doi: 10.1038/ncomms11207.

Lucas I, Chevrier-Miller M, Sogo JM, Hyrien O. Mechanisms Ensuring Rapid and Complete DNA Replication despite Random Initiation in Xenopus Early embryos11Edited by M. Yaniv. Journal of Molecular Biology. 2000 Feb; 296(3):769–786. doi: 10.1006/jmbi.2000.3500.

Lupardus PJ, Byun T, Yee Mc, Hekmat-Nejad M, Cimprich KA. A Requirement for Replication in Activation of the ATR-Dependent DNA Damage Checkpoint. Genes Dev. 2002 Sep; 16(18):2327–2332. doi: 10.1101/gad.1013502.

Machida YJ, Hamlin JL, Dutta A. Right Place, Right Time, and Only Once: Replication Initiation in Metazoans. Cell. 2005 Oct; 123(1):13–24. doi: 10.1016/j.cell.2005.09.019.

Mahbubani HM, Paull T, Elder JK, Blow JJ. DNA Replication Initiates at Multiple Sites on Plasmid DNA in Xenopus Egg Extracts. Nucleic Acids Res. 1992 Apr; 20(7):1457–1462.

Mantiero D, Mackenzie A, Donaldson A, Zegerman P. Limiting Replication Initiation Factors Execute the Temporal Programme of Origin Firing in Budding Yeast. EMBO J. 2011 Nov; 30(23):4805–4814. doi: 10.1038/emboj.2011.404.

Marheineke K, Hyrien O. Aphidicolin Triggers a Block to Replication Origin Firing inXenopus Egg Extracts. J Biol Chem. 2001 May; 276(20):17092–17100. doi: 10.1074/jbc.M100271200.

Marheineke K, Hyrien O. Control of Replication Origin Density and Firing Time in Xenopus Egg Extracts ROLE OF A CAFFEINE-SENSITIVE, ATR-DEPENDENT CHECKPOINT. J Biol Chem. 2004 Jul; 279(27):28071–28081. doi: 10.1074/jbc.M401574200.

Maya-Mendoza A, Petermann E, Gillespie DA, Caldecott KW, Jackson DA. Chk1 Regulates the Density of Active Replication Origins during the Vertebrate S Phase. EMBO J. 2007 Jun; 26(11):2719–2731. doi: 10.1038/sj.emboj.7601714.

McCune HJ, Danielson LS, Alvino GM, Collingwood D, Delrow JJ, Fangman WL, Brewer BJ, Raghuraman MK. The Temporal Program of Chromosome Replication: Genomewide Replication in clb5Δ Saccharomyces Cerevisiae. Genetics. 2008 Dec; 180(4):1833–1847. doi: 10.1534/genetics.108.094359.

Méchali M, Kearsey S. Lack of Specific Sequence Requirement for DNA Replication in Xenopus Eggs Compared with High Sequence Specificity in Yeast. Cell. 1984 Aug; 38(1):55–64. doi: 10.1016/0092-8674(84)90526-9.

Michelena J, Gatti M, Teloni F, Imhof R, Altmeyer M. Basal CHK1 Activity Safeguards Its Stability to Maintain Intrinsic S-Phase Checkpoint Functions. J Cell Biol. 2019 Sep; 218(9):2865–2875. doi: 10.1083/jcb.201902085.

Moiseeva TN, Yin Y, Calderon MJ, Qian C, Schamus-Haynes S, Sugitani N, Osmanbeyoglu HU, Rothenberg E, Watkins SC, Bakkenist CJ. An ATR and CHK1 Kinase Signaling Mechanism That Limits Origin Firing during Unperturbed DNA Replication. PNAS. 2019 Jul; 116(27):13374–13383. doi: 10.1073/pnas.1903418116.

Petryk N, Kahli M, d'Aubenton-Carafa Y, Jaszczyszyn Y, Shen Y, Silvain M, Thermes C, Chen CL, Hyrien O. Replication Landscape of the Human Genome. Nat Commun. 2016 Jan; 7(1):10208. doi: 10.1038/ncomms10208.

Platel M, Goldar A, Wiggins JM, Barbosa P, Libeau P, Priam P, Narassimprakash H, Grodzenski X, Marheineke K. Tight Chk1 Levels Control Replication Cluster Activation in Xenopus. PLoS One. 2015 Jun; 10(6). doi: 10.1371/journal.pone.0129090.

Pommier Y, Kohn KW. Cycle cellulaire et points de contrôle en oncologie : nouvelles cibles thérapeutiques. Med Sci (Paris). 2003 Feb; 19(2):173–186. doi: 10.1051/medsci/2003192173.

Raghuraman MK. Replication Dynamics of the Yeast Genome. Science. 2001 Oct; 294(5540):115–121. doi: 10.1126/science.294.5540.115.

Rhind N, Gilbert DM. DNA Replication Timing. Cold Spring Harb Perspect Biol. 2013 Aug; 5(8):a010132. doi: 10.1101/cshperspect.a010132.

Seiler JA, Conti C, Syed A, Aladjem MI, Pommier Y. The Intra-S-Phase Checkpoint Affects Both DNA Replication Initiation and Elongation: Single-Cell and -DNA Fiber Analyses. MCB. 2007 Aug; 27(16):5806–5818. doi: 10.1128/MCB.02278-06.

Sekedat MD, Fenyö D, Rogers RS, Tackett AJ, Aitchison JD, Chait BT. GINS Motion Reveals Replication Fork Progression Is Remarkably Uniform throughout the Yeast Genome. Molecular Systems Biology. 2010 Jan; 6(1). doi: 10.1038/msb.2010.8.

Siefert JC, Georgescu C, Wren JD, Koren A, Sansam CL. DNA Replication Timing during Development Anticipates Transcriptional Programs and Parallels Enhancer Activation. Genome Res. 2017 Aug; 27(8):1406–1416. doi: 10.1101/gr.218602.116.

Stokes MP, Van Hatten R, Lindsay HD, Michael WM. DNA Replication Is Required for the Checkpoint Response to Damaged DNA in Xenopus Egg Extracts. J Cell Biol. 2002 Sep; 158(5):863–872. doi: 10.1083/jcb.200204127.

Syljuasen RG, Sorensen CS, Hansen LT, Fugger K, Lundin C, Johansson F, Helleday T, Sehested M, Lukas J, Bartek J. Inhibition of Human Chk1 Causes Increased Initiation of DNA Replication, Phosphorylation of ATR Targets, and DNA Breakage. MCB. 2005 May; 25(9):3553–3562. doi: 10.1128/MCB.25.9.3553-3562.2005.

Tanaka S, Nakato R, Katou Y, Shirahige K, Araki H. Origin Association of Sld3, Sld7, and Cdc45 Proteins Is a Key Step for Determination of Origin-Firing Timing. Current Biology. 2011 Dec; 21(24):2055–2063. doi: 10.1016/j.cub.2011.11.038.

Thomson AM, Gillespie PJ, Blow JJ. Replication Factory Activation Can Be Decoupled from the Replication Timing Program by Modulating Cdk Levels. The Journal of Cell Biology. 2010 Jan; 188(2):209–221. doi: 10.1083/jcb.200911037.

Ticau S, Friedman LJ, Ivica NA, Gelles J, Bell SP. Single-Molecule Studies of Origin Licensing Reveal Mechanisms Ensuring Bidirectional Helicase Loading. Cell. 2015 Apr; 161(3):513–525. doi: 10.1016/j.cell.2015.03.012.

Trenz K, Errico A, Costanzo V. Plx1 Is Required for Chromosomal DNA Replication under Stressful Conditions. The EMBO Journal. 2008 Mar; 27(6):876–885. doi: 10.1038/emboj.2008.29.

Yang SCH, Rhind N, Bechhoefer J. Modeling Genome-Wide Replication Kinetics Reveals a Mechanism for Regulation of Replication Timing. Molecular Systems Biology. 2010 Jan; 6(1):404. doi: 10.1038/msb.2010.61.

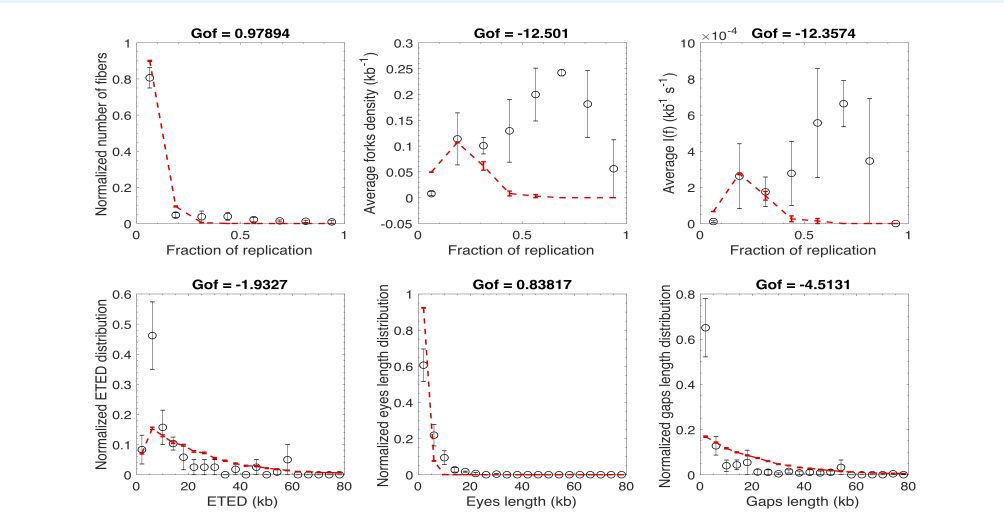
Zegerman P, Diffley JFX. Checkpoint Dependent Inhibition of DNA Replication Initiation by Sld3 and Dbf4 Phosphorylation. Nature. 2010 Sep; 467(7314):474–478. doi: 10.1038/nature09373.

Zou L. Single- and Double-Stranded DNA: Building a Trigger of ATR-Mediated DNA Damage Response. Genes Dev. 2007 Apr; 21(8):879–885. doi: 10.1101/gad.1550307.

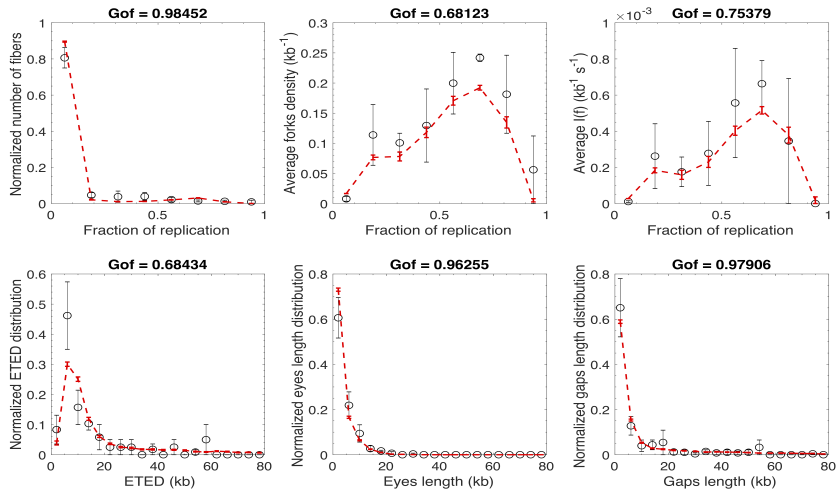
Appendix 1

Different models

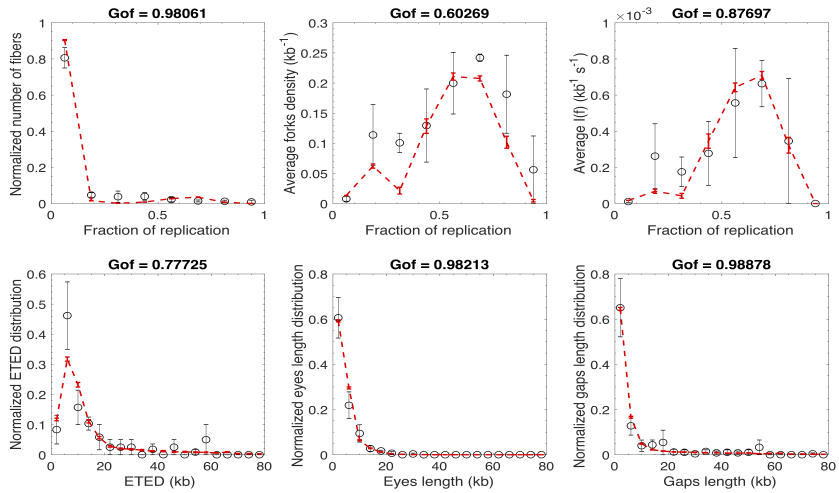
To model experimental observations a series of nested models were compared with experimental data. Below are the fits of each model to experimental sample with 8% global replicated fraction. To assess the goodness of the fit (GoF) we considered the normalised mean square error between the simulated profile and the fitted entity as the indicator of likelihood ($GoF = 1 - \frac{||y_{fit} - y_{exp}||^2}{||y_{exp} - \langle y_{exp} \rangle||^2}$). GoF costs vary between $-\infty$ (bad fit) to 1 (perfect fit). If $GoF = 0$, y_{fit} is no better than a straight line at matching experimental data. The global cost is calculated as $GoF_{global} = \frac{1}{6}\sum_1^6 GoF_i$ where i represents one fitted entity. All models reproduce with the same accuracy the distribution of replicated fibres, gaps lengths and eyes lengths distributions. The major contributions to score values come from residuals of average fork density, average $I(f)$ and eye-to-eye distances distribution fits. From the value of GoF_{global} (Appendix1, ***Table 1***), the model that best described the whole data set is the MM4 with localized distribution of potential origins: its GoF_{global} value is closest to one. However, MM4 also has the highest number of fitting variables (7) compared to other models (MM1 has 3 fitting variables, MM2 and MM3 have 5 fitting variables), and facilitating fit to the data.



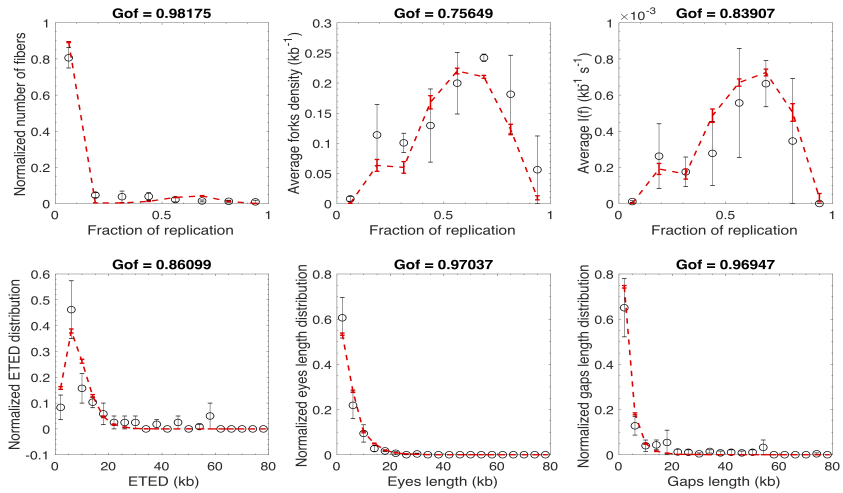
Appendix 1 Figure 1. Modeling experimental data with MM1 model in the case where the potential origins are continuously distributed along the genome. Open circles are experimental data and the red dashed line is the fit.



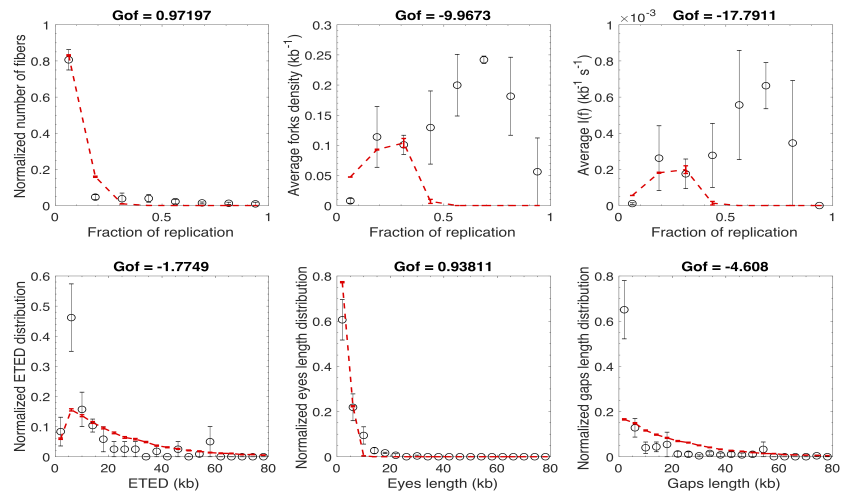
Appendix 1 Figure 2. Modeling experimental data with MM2 model in the case where the potential origins are continuously distributed along the genome. Open circles are experimental data and the red dashed line is the fit.



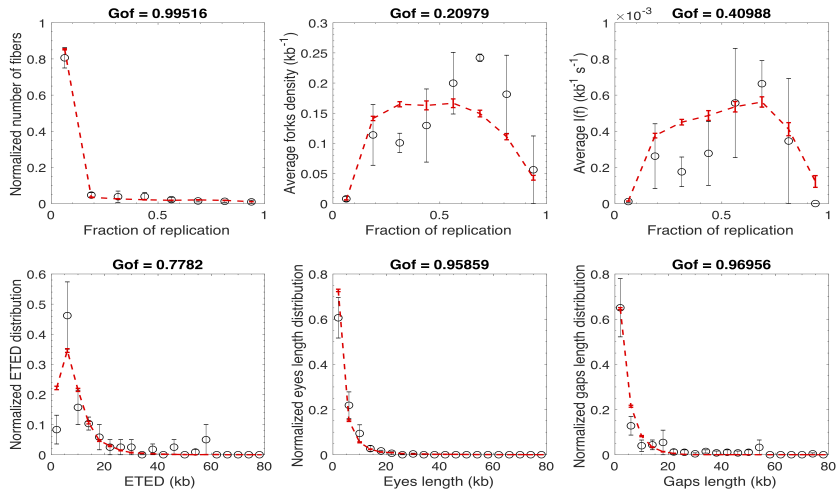
Appendix 1 Figure 3. Modeling experimental data with MM3 model in the case where the potential origins are continuously distributed along the genome. Open circles are experimental data and the red dashed line is the fit.



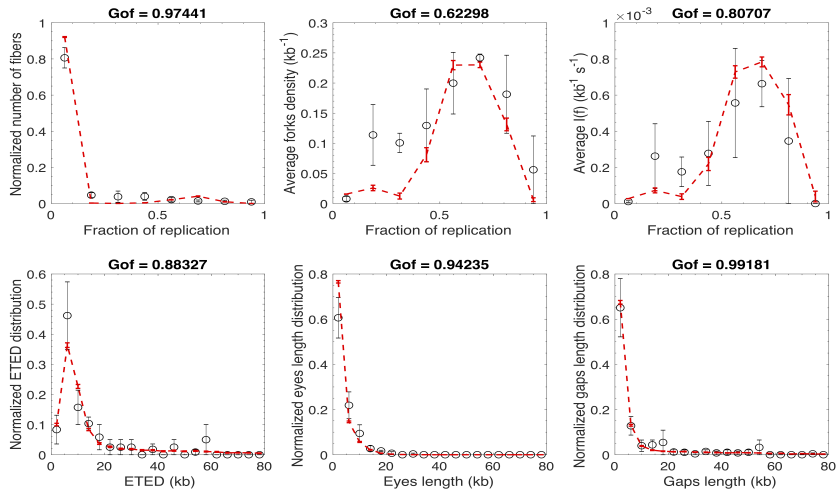
Appendix 1 Figure 4. Modeling experimental data with MM4 model in the case where the potential origins are continuously distributed along the genome. Open circles are experimental data and the red dashed line is the fit.



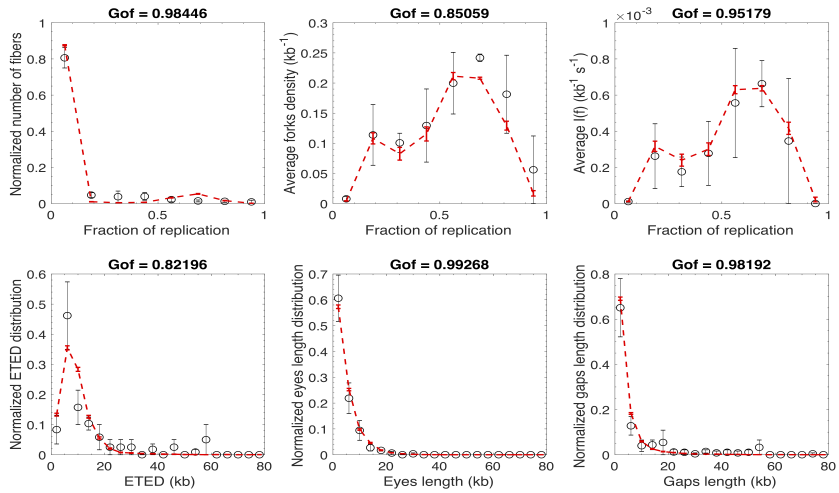
Appendix 1 Figure 5. Modeling experimental data with MM1 model in the case where the potential origins form a discrete set along the genome. Open circles are experimental data and the red dashed line is the fit.



Appendix 1 Figure 6. Modeling experimental data with MM2 model in the case where the potential origins form a discrete set along the genome. Open circles are experimental data and the red dashed line is the fit.



Appendix 1 Figure 7. Modeling experimental data with MM3 model in the case where the potential origins form a discrete set along the genome. Open circles are experimental data and the red dashed line is the fit.



Appendix 1 Figure 8. Modeling experimental data with MM4 model in the case where the potential origins form a discrete set along the genome. Open circles are experimental data and the red dashed line is the fit.

	Continuous	Discrete	Continuous	Discrete
model	GoF_{global}	GoF_{global}	$(y_{exp} - y_{fit})^2$	$(y_{exp} - y_{fit})^2$
MM1	-0.95	-5.28	0.66	0.56
MM2	0.85	0.72	0.08	0.10
MM3	0.87	0.88	0.08	0.09
MM4	0.90	0.92	0.08	0.05

Appendix 1 Table 1. Values of GoF_{global} and fitting residual norm ($(y_{exp} - y_{fit})^2$) for each model.

Models comparison

To address whether the better data fit with MM4 is solely due to the higher degree of complexity of the model, we used two different approaches : a traditional statistical hypothesis testing: the extra sum of squares F test (***Bevington and Robinson, 2003***) and the Akaike's criterion (ΔAIC) that is based on information theory (***Ljung, 1998***). We can objectively reject MM1 as it did not reproduce in a satisfactory manner the averaged fork density, $I(f)$ and eye-to-eye distances distributions (Appendix 1, ***Figure 1*** and ***Figure 5***). MM2 and MM3 satisfactorily reproduced all measured quantities (Appendix 1, ***Figure 2*** , ***Figure 3***, ***Figure 6*** and ***Figure 7***) but with lower GoF_{global} value than the MM4 models (Appandix1, ***Table 1***). The discrete MM4 model has higher GoF_{global} value than the continuous one, whereas the continuous MM2 and MM3 models were better than or equal to their discrete version, respectively (Appandix1, ***Table 1***). To choose the best model, we compared the discrete MM4 model, continuous MM2, MM3 and MM4 corresponding to fits with highest GoF_{global} values (Appendix1, ***Table 1***). Comparing the discrete MM4 with the continuous MM2, MM3 and MM4 models led in all cases to $F > 1$ with p-values $p < 10^{-6}$ and negative ΔAIC values (Appendix1, ***Table 2***). The discrete MM4 model is therefore the best model and the observed increase in GoF_{global} does not reflect an overfitting of the data.

model	F	p	ΔAIC
Continuous MM2	19.3	1.5×10^{-7}	-30.2
Continuous MM3	16.9	8.3×10^{-7}	-26.6
Continuous MM4	∞	Not defined	-31.1

Appendix 1 Table 2. Values of F-test, the associated p-value (p) and the ΔAIC when the discrete MM4 model is compared with continuous MM2, MM3 and MM4 model.

Appendix 2

The MM5 model used to generate the *in silico* data

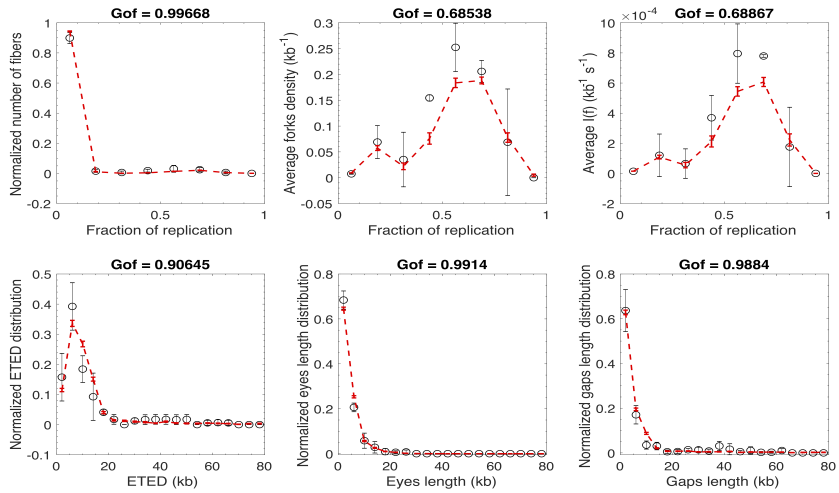
In MM5, localized potential origins were distributed with a uniform density $\rho = 1\ kb^{-1}$ and N_{dom} domains of size l_{dom} were randomly positioned along a genome of length $L = 10^5\ kb$. As in previous works, we assumed that at the start to S phase N_0 limiting factors were available for origin firing and their number, $N(t)$, increased during the cours of S phase as $N(t) = N_0 + Jt$, and that each factor was sequestrated by new forks upon origin activation and released and made available again for origin firing upon coalescence of converging forks. Forks progressed at a constant velocity $v = 0.5\ kb.min^{-1}$. The probability of origin firing by encounter with a limiting factor was higher inside the domains ($P_0 + P_{dom}$) than outside them (P_0). In addition, origins outside but not inside the domains had a non-null probability P_{inhib} of being inhibited. Two local effects were allowed to act within a distance d_{fork} from active forks: P_0 was enhanced by P_{fork} and origin inhibition was relieved with a probability $P_{deinhib}$. We simulated 300 complete S phases using the 10 parameter values listed in Appendix 2, ***Table 1***, and extracted snapshots at 8%, 19% and 53% global replicated fractions. Each snapshot was considered as an independent sample and for each of them: i) the genome was randomly cut following the molecule length distribution presented in ***Figure 6***, ii) the data were reshaped as described in material and methods to account for the finite experimental resolution and iii) the distributions of $I(f)$, replicated fraction of single fibres, global fork density, eye-to-eye distances, gap lengths and eye lengths were determined.

Parameter	Value
N_0	107
$J(s^{-1})$	29
P_0	0.11
P_{inhib}	0.96
P_{fork}	0.28
$d(kb)$	94.91
N_{dom}	196
l_{dom}	192.39
$P_{deinhib}$	0.06
P_{dom}	0.73

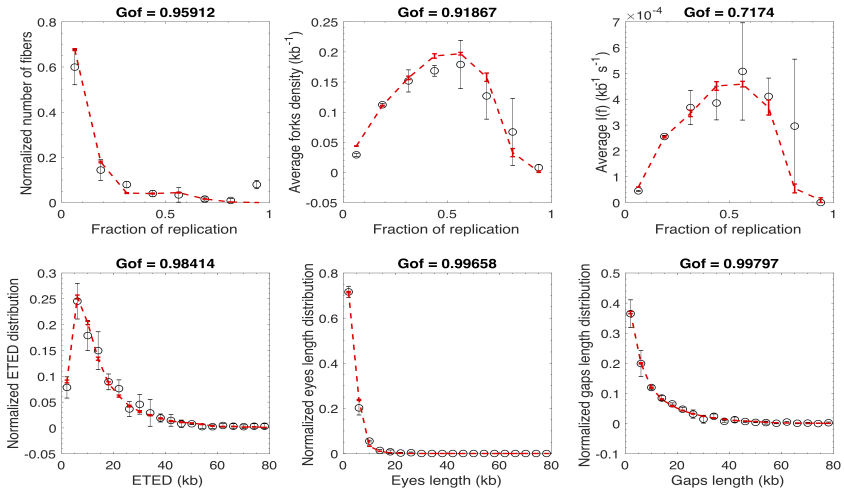
Appendix 2 Table 1. Values of MM5's parameters. These values are chosen arbitrarly

Fitting the *in silico* data by MM4 model

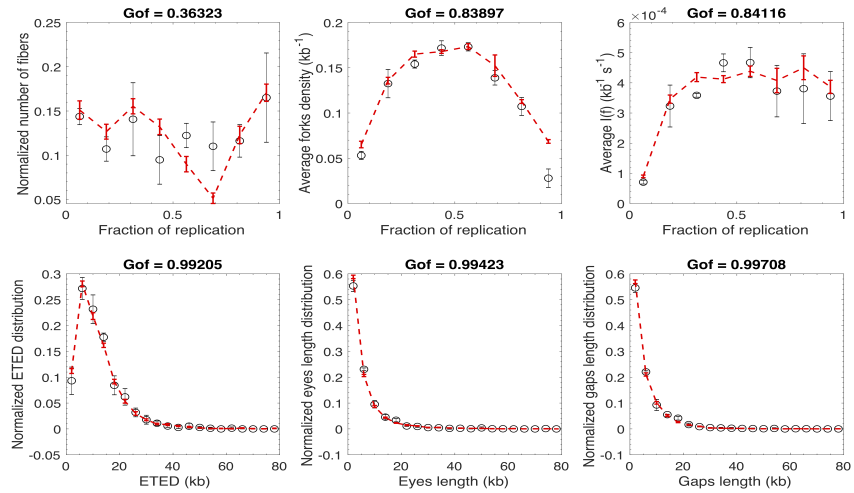
By independently fitting the simulated profiles of each global replicated fraction, we implicitly assume that samples could originate from separated experiments, hence MM4 parameters values are possibly different for each global replicated fraction. This allows us to accurately reproduce observations from each sample (Appendix2 ***Figure 1***, ***Figure 2*** and ***Figure 3***).



Appendix 2 Figure 1. Modeling 8% global replicated fraction simulated data with discrete MM4 model. Open circles are simulated data and the red dashed line is the fit. $GoF_{global} = 0.96$



Appendix 2 Figure 2. Modeling 19% global replicated fraction simulated data with discrete MM4 model. Open circles are simulated data and the red dashed line is the fit. $GoF_{global} = 0.97$



Appendix 2 Figure 3. Modeling 53% global replicated fraction simulated data with discrete MM4 model. Open circles are simulated data and the red dashed line is the fit. $GoF_{global} = 0.82$

Reduction of MM5 to MM4

In the MM5 model origins fire globally with two origin firing probabilities (P_0 and $P_0 + P_{dom}$) eventually increased by a local origin firing probability (P_{fork}) close to an active fork, and the genome is divided into domains that either support or escape some inhibitory probability of firing (assumed to represent inhibition by the intra-S checkpoint). As the position of these domains is not identical between repeated simulations, we can reduce their description by specifying a fraction θ ($\theta = \frac{N_{dom} l_{dom}}{L}$) of the genome where origins escape checkpoint inhibition. In these domains, the global origin firing probability $P_{in} = \frac{1}{2}\left(P_0 + P_{dom}\right)$, with the $\frac{1}{2}$ pre-factor being due to normalization considerations. The local probability of origin firing (close to a fork) inside a domain is $P^{in}_{local} = \frac{1}{2}\left(P_0 + P_{dom} + P_{fork}\right)$. Outside these domains, the global probability of origin firing is modulated by the probability of origin inhibition $P_{out} = \frac{1}{2} P_0 \left(1 - P_{inhib}\right)$. In the same manner the local probability of origin firing is modulated by the action of intra-S checkpoint and the local cancellation of inhibition process $P^{out}_{local} = \frac{1}{2}\left(P_0 + P_{fork}\right)\left[1 + P_{inhib}\left(P_{deinhib} - 1\right)\right]$. Local probabilities of origin firing only influence origins over a distance d_{fork} downstream of a fork. The MM4 model contains a unique local probability of origin firing, that corresponds to the average value of the two local probabilities of origin firing, $P_{local} = \theta P^{in}_{local} + (1-\theta) P^{out}_{local}$. Therefore, by considering the essential ingredients of the MM5 model, we combined the parameters of the model to retrieve the parameters of MM4 (***Table 2***).

MM4	equivalence with MM5
N_0	N_0
$J\ (s^{-1})$	J
θ	$\frac{N_{dom} l_{dom}}{L}$
P_{in}	$\frac{1}{2}\left(P_0 + P_{dom}\right)$
P_{local}	$\frac{1}{2}\left(P_0 + P_{fork}\right)\left[1 + (1-\theta) P_{inhib}\left(P_{deinhib} - 1\right)\right] + \theta P_{dom}$
P_{out}	$\frac{1}{2} P_0 \left(1 - P_{inhib}\right)$
$d\ (kb)$	d

Appendix 2 Table 2. Reducing MM5 to MM4.

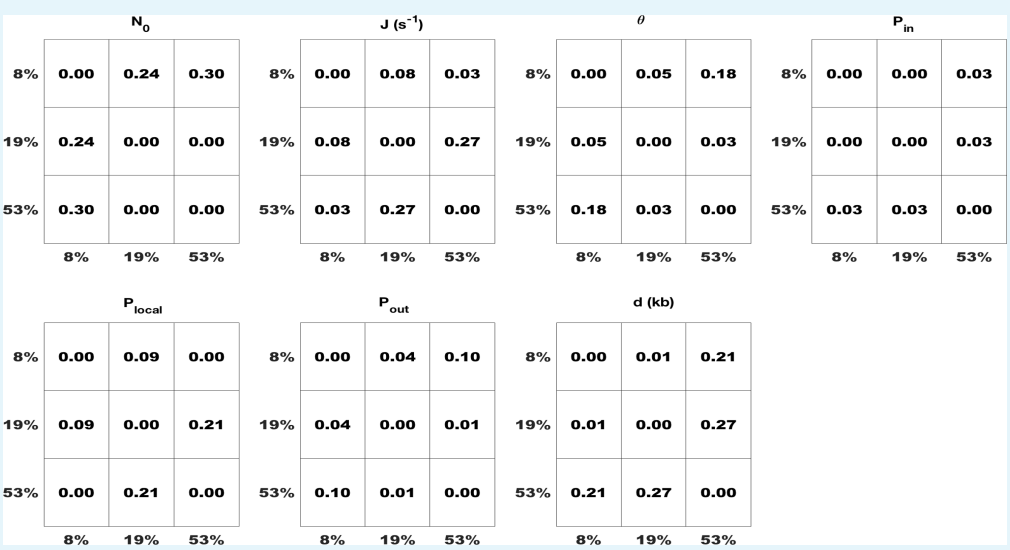
The values of these parameters can be compared directly to parameters of MM4 model

obtained from the fitting of the simulated data for each sample (***Table 3***). To assess if the difference between the expected and the inferred value of a parameter is statistically significant we calculate $t = \frac{(\text{expected value} - \text{inferred value})^2}{\text{error}^2}$, for $t \geq 1$ the difference is statistically significant otherwise it is not. The values of parameters changed as the global replicated fraction increased (Appendix 2, ***Table 3***). To assess the level of significance of these variations we calculated $\chi^2 = \frac{(\text{parameter}_1 - \text{parameter}_2)^2}{\text{error}_1^2 + \text{error}_2^2}$ coefficient between the values of the same parameter obtained for different global replicated fraction. If $\chi^2 < 1$ the difference between the two values was not statistically significant otherwise it was significant. Appendix 2, ***Figure 4*** shows that the differences of predicted parameters values among the 3 considered samples were not statistically significant, as was expected.

MM4	Input	8%	19%	53%
N_0	107	83.86 ± 32 ($t < 1$)	125 ± 29 ($t < 1$)	129 ± 26 ($t < 1$)
J (s^{-1})	29	43.6 ± 46 ($t < 1$)	17 ± 9 ($t < 1$)	27 ± 3.4 ($t < 1$)
θ	0.38	0.25 ± 0.2 ($t < 1$)	0.35 ± 0.16 ($t < 1$)	0.42 ± 0.1 ($t < 1$)
P_{in}	0.42	0.4 ± 0.2 ($t < 1$)	0.41 ± 0.17 ($t < 1$)	0.5 ± 0.2 ($t < 1$)
P_{local}	0.22	0.23 ± 0.09 ($t < 1$)	0.17 ± 0.05 ($t < 1$)	0.23 ± 0.04 ($t < 1$)
P_{out} ($\times 10^{-3}$)	2.2	1.1 ± 1 ($t < 1$)	1.9 ± 1 ($t < 1$)	2.3 ± 1 ($t < 1$)
d (kb)	94.91.	135 ± 86 ($t < 1$)	119 ± 57 ($t < 1$)	51 ± 32 ($t < 1$)

Appendix 2 Table 3. Comparison between the expected and inferred values of MM4 parameters.

All $t < 1$ and $\chi^2 < 1$ (Appendix 2, ***Figure 4***), meaning the constancy of parameters values for all three samples. Therefore, we conclude that the optimization procedure was able to circumscribe the expected parameters values in an accurate manner for each sample. It should be noted that we choose a very conservative criterion to assess if two parameters are different or not. The conditions of $\chi^2 = 1$ or $t = 1$ are equivalent to a confidence level of $\alpha = 10^{-7}$ in the case of a two sided and one sided t statistics. In other words, with our criterion the probability to find that the values of two parameters are different by chance is smaller than 10^{-7}.



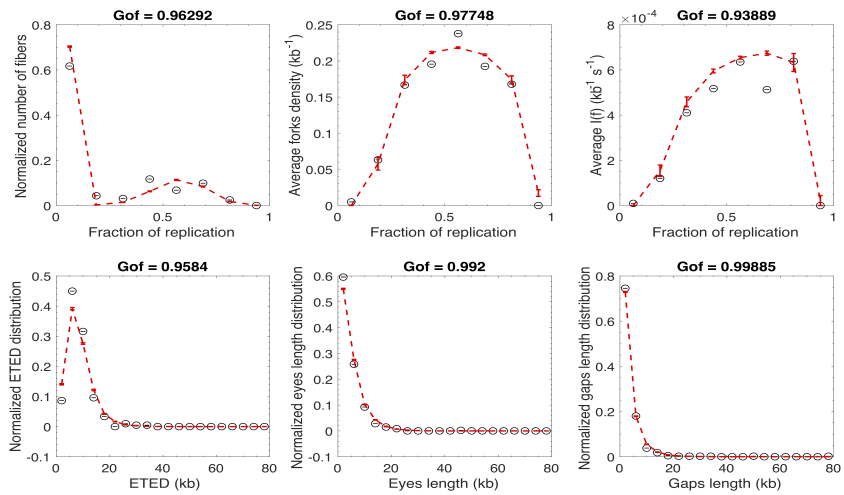
Appendix 2 Figure 4. The values of each MM4 model parameter were compared pair-wise between samples with different global replicated fraction. The statistical significance of their difference was assessed by χ^2 test and represented as a binary heat map where not statistically significant differences are couloured in white and statistically significant difference are coloured in blue. The number in each box is the χ^2 coefficient.

The ability of the fitting procedure i) to circumscribe the values of MM4 model parameters close to the expected ones (Appendix 2, ***Table 3***) and ii) to retrieve the constancy of these parameter's values as the global degree of replication increases (Appendix 2, ***Figure 4***) demonstrates the adequacy of our fitting strategy to recover the dynamic of DNA replication during S phase in the framework of MM4 model by setting the null hypothesis as : the values of MM4 parameters do not change as S phase progresses. Therefore, rejection of this hypothesis for a considered parameter means its variation during S phase.

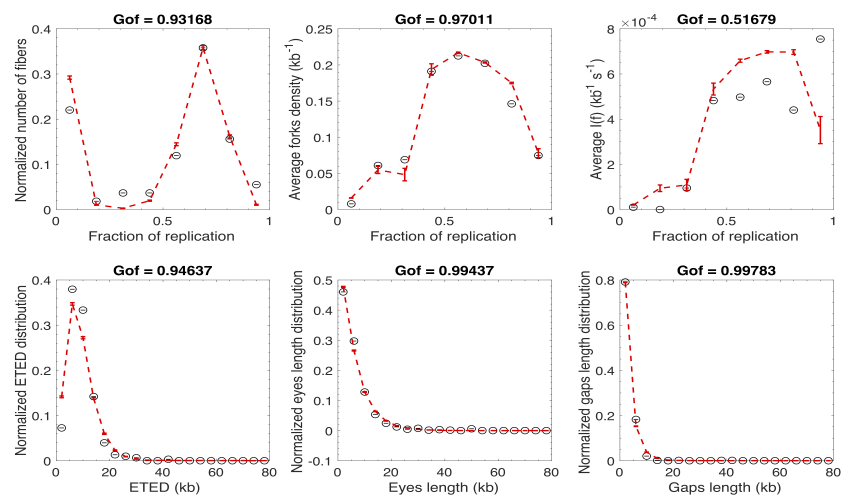
Appendix 3

Fitting the experimental profiles by MM4 model : Unchallenged S phase

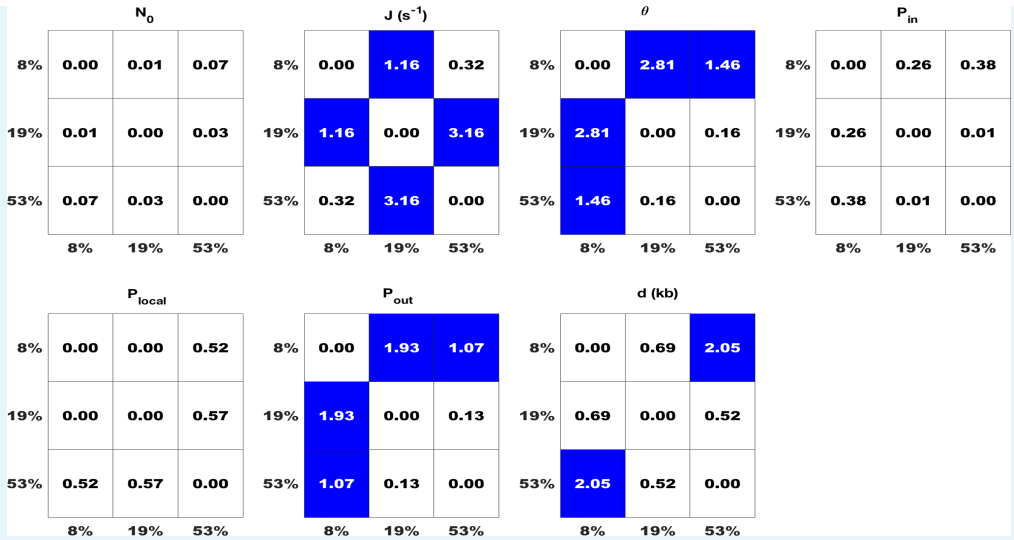
We fitted independently the measured profiles for each global replicated fraction by discrete MM4 model. The fits of observations from 8% global replicated fraction are presented in Appendix 1***Figure 8*** and those of 19% and 53% are presented Appendix 3 ***Figure 1*** and ***Figure 2*** respectively. In Appendix 3 ***Table 1*** we give the value of the fitted parameters. The reliability of observed differences among inferred MM4 parameters are assessed statistically by using χ^2 coefficient as defined in Appendix 2 (Appendix 3 ***Figure 3***)



Appendix 3 Figure 1. Modeling measured sample with 19% global replicated fraction with the discrete MM4 model. Open circles are simulated data and the red dashed line is the fit. $GoF_{global} = 0.96$



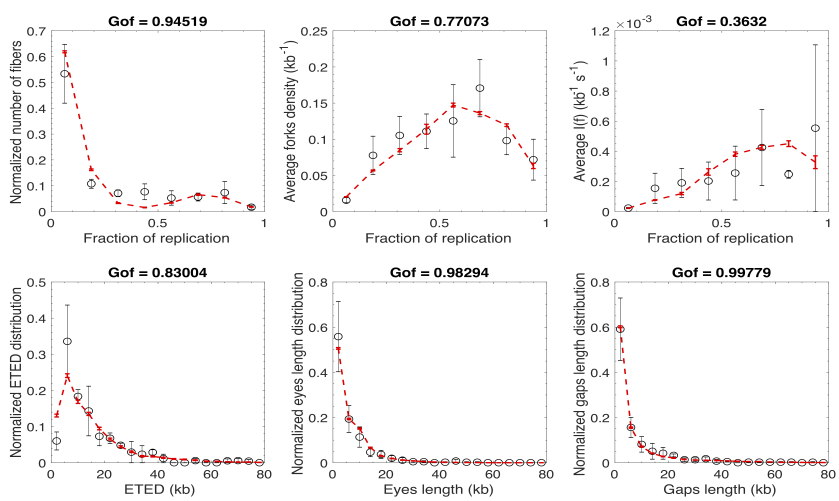
Appendix 3 Figure 2. Modeling measured sample with 53% global replicated fraction with the discrete MM4 model. Open circles are simulated data and the red dashed line is the fit. $GoF_{global} = 0.90$



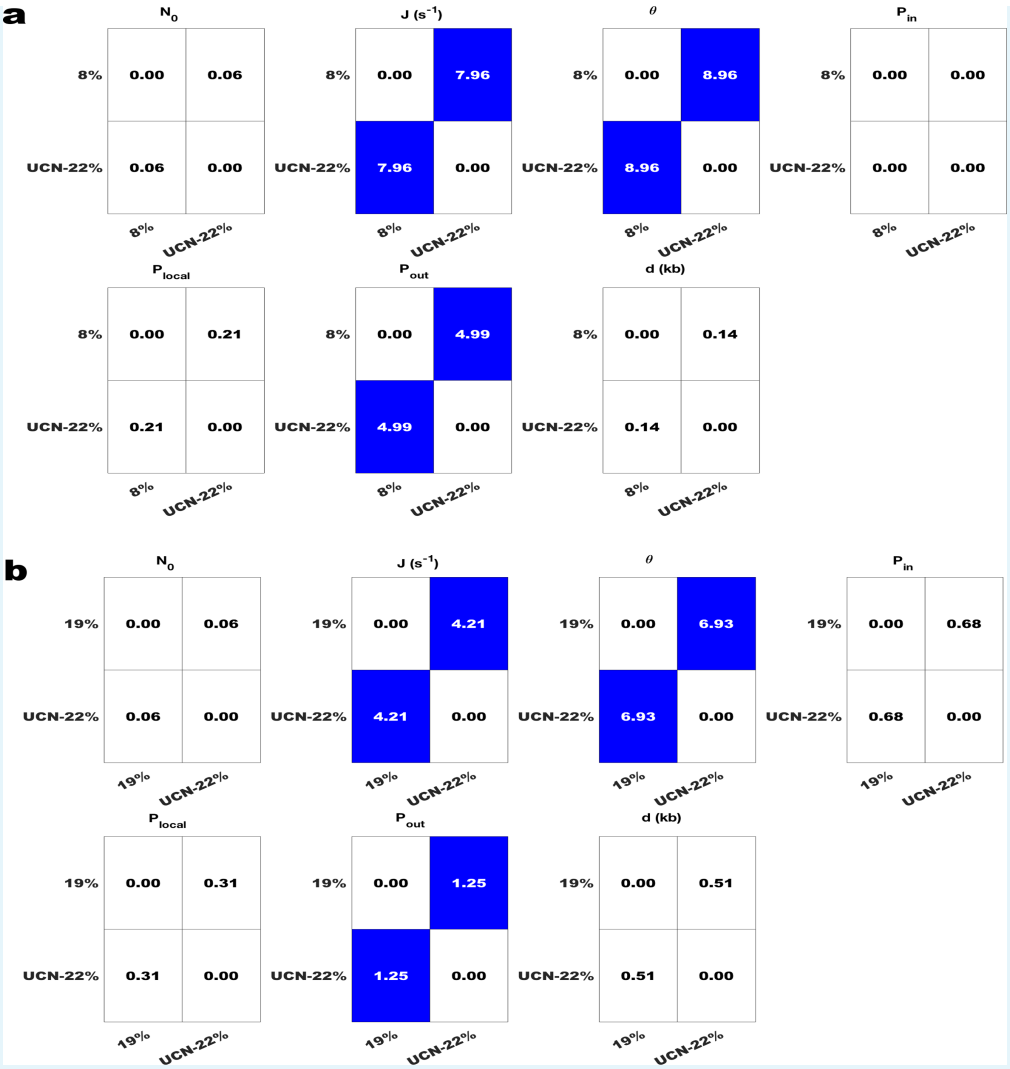
Appendix 3 Figure 3. The values of each MM4 model parameter were compared pair-wise between samples with different global replicated fraction. The statistical significance of their difference was assessed by χ^2 test and represented as a binary heat map where the white colour represents no statistically significant difference and the blue colour represents statistically significant difference. The number in each box is the χ^2 coefficient.

Fitting the experimental profiles by MM4 model : Chk1 inhibited S phase

We fitted with the discrete MM4 model a sample that had spent in the presence of UCN-01 the same time interval in S phase as the control sample with 8% global replicated fraction. The global replicated fraction of the of the UCN-01 sample was 22%. The fits are presented in Appendix 3 ***Figure 4*** and the obtained parameters values are given in Appendix 3 ***Table 1***. The reliability of observed differences among inferred MM4 parameters between controls and Chk1 inhibited sample are assessed statistically by using χ^2 coefficient as defined in Appendix 2 (Appendix 3 ***Figure 5***)



Appendix 3 Figure 4. Modeling a measured sample with 22% global replicated fraction in presence of UCN-01with discrete MM4 model. Open circles are simulated data and the red dashed line is the fit. $GoF_{global} = 0.85$



Appendix 3 Figure 5. a. Comparing samples that have spent the same time interval in S phase. **b**. Comparing samples that have similar global replication fractions. The values of each MM4 model parameter were compared pair-wise between samples with different global replicated fraction. The statistical significance of their difference was assessed by χ^2 test and represented as a binary heat map where the white colour represents no statistically significant difference and the blue colour represents statistically significant difference. The number in each box is the χ^2 coefficient.

MM4	unchallenged: 8%	unchallenged: 19%	unchallenged: 53%	UCN-01: 22%
N_0	1064 ± 135	1043 ± 116	1002 ± 106	1006 ±102
$J\ (s^{-1})$	601 ± 198	1026 ± 196	404 ± 151	1467 ± 89
θ	0.25 ± 0.06	0.43 ± 0.04	0.39 ± 0.05	0.56 ± 0.032
P_{in}	0.41 ± 0.07	0.34 ± 0.07	0.32 ± 0.07	0.42 ± 0.07
P_{local}	0.43 ± 0.06	0.43 ± 0.06	0.52 ± 0.06	0.38 ± 0.06
P_{out}	0.09± 0.02	0.17 ± 0.04	0.15 ± 0.03	0.23 ± 0.04
$d\ (kb)$	143.8 ± 36.3	91.5 ± 25.6	56.1 ± 23.6	119.3 ± 29.3

Appendix 3 Table 1. Values and the corresponding errors of MM4 parameters for the best fit of each sample and each condition.