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[2], Olivier Hyrien[3], Benjamin Audit[4], Kathrin Marheineke[1*], Arach Goldar[1*],

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

* arach.goldar@cea.fr, kathrin.marheineke@i2bc.paris-saclay.fr

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 which is proposed to be regulated by the intra-S checkpoint. By modelling the unchallenged, the checkpoint-inhibited and the checkpoint protein Chk1 over-expressed replication pattern of single DNA molecules from *Xenopus* sperm chromatin replicated in egg extracts, we demonstrate that the quantitative modelling of data requires: 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 and 3) the variability of the rate of DNA synthesis close to replication forks is a necessary ingredient that should be taken in to account in order to describe the dynamic of replication origin firing. This model implies that the observed origin clustering emerges from the apparent synchrony of origin firing in regions with high probability of origin firing and challenge the assumption that the intra-S checkpoint is the main regulator of origin clustering.

Author summary

DNA replication is one of the fundamental cell functions. The genome of eukaryotic organisms is duplicated from multiple positions named replication origins. Single molecule experiments allow to visualise the dynamics of spatio-temporal patterns created during replication process. The dynamic of replication process is regulated by checkpoints. However, the influence and the role of checkpoint regulation in the dynamics of spatio-temporal patterns of DNA replication is not understood. In this work we build a minimal, process-based and data rooted numerical model that allows to decipher the impact of checkpoint regulation on the dynamics of spatio-temporal pattern of DNA replication.

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 [1, 2]. The mechanisms that regulate the timing of origin firing remain largely unknown [3–8].

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 [9, 10]. 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 [1]. MCM2-7 DHs that fail to fire are inactivated by forks emanating from neighboring fired origins [11]. 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 (RecQ4 in *Xenopus*), Sld3 (Tresline in *Xenopus*), Dpb11 (TopBP1 in *Xenopus*), Dbf4 (Drf1 in *Xenopus*), Sld7 (MTBP in *Xenopus*) 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 [12, 13]. Among these six factors, Cdc45 is the only one to travel with the replication fork.

DNA replication initiates without sequence specificity in *Xenopus* eggs [14, 15], egg extracts [16–19] and early embryos [20, 21] (for review see [22]). 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 [23–27] which contrast to population based approaches, that average replication characteristics. DNA combing 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 the origin firing rate strongly increases from early to late replication intermediates, speeding up late replication stages [23, 24]. An observation that has been also confirmed in many other model organisms, including human cell lines [28].

A mathematical model based on the assumptions i) 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 named "mean-field hypothesis"), ii) that firing of origins are independent events and ii) that fork speed is constant was proposed [29]. This model 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 (Fig 1.a) [29]. 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 [22, 30]. 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 $I(t)$ [30].

Importantly, experiments revealed that in *Xenopus*, like in other eukaryotes, replication eyes are not homogeneously distributed over the genome but tend to cluster [25, 27]. First, a weak correlation between the sizes of neighbouring eyes was observed [25, 27, 30], consistent with firing time correlations. Second, more molecules with no or multiple eyes than expected for spatially uniform initiations were observed in replicating DNA [27]. 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 [31].

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 [32]. 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 [27, 33–36]. DNA replication stress, through the activation of the S-phase checkpoint kinase Rad53, can inhibit origin firing by phosphorylating and inhibiting Sld3 and Dbf4 [37]. 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 [33]. Experiments using *Xenopus* egg extracts suggested that the checkpoint mainly adjusts the rate of DNA synthesis by staggering the firing time of origin clusters [27]. Our first model for DNA replication in *Xenopus* egg extracts [38] (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) was used to model the regulation of DNA replication by the intra-S checkpoint [35]. We showed that even during an unperturbed S phase in *Xenopus* egg extracts, Chk1 inhibits origin firing away from but not near active forks [35]. 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 [33, 39–41]. 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 both a control and a UCN-01-inhibited Chk1 replication experiment [35]. 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 analysed both the temporal and spatial characteristics of genome replication as measured by DNA combing in the *in vitro Xenopus* system. *Xenopus* egg extracts have been successfully used since over three decades now to study DNA replication in metazoans [42]. Rooted on experimental data, we build a general and minimal model of DNA replication able to predict both the temporal and the spatial characteristics either during an unchallenged or a challenged S phase. We use the experimental data from [35] where the experimental mean chosen for activating or inhibiting (manipulating) the checkpoint was respectively to overexpress Chk1 protein or to inhibit its activity using the specific inhibitor UCN-01. By analysing the spatio-temporal pattern of DNA replication after inhibition or activation of intra-S checkpoint and comparing them to an unchallenged pattern we disentangle the complex role of the intra-S checkpoint for replication origin firing.

Materials and methods

Monte Carlo simulation of DNA replication process.

A dynamical Monte Carlo method was used to simulate the DNA replication process as detailed before [38]. 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 (except in MM2 where the value of $v(\ kb.min^{-1})$ for each active fork is randomly chosen at each round of calculus from {0,1,2,3} ($kb.min^{-1}$)), 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). As it is also considered that potential replication origins are discrete objects and as a consequence are distributed in a heterogeneous manner on the genome [43, 44] we also simulate the case where the distribution of potential origins is discrete. 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 [45]. 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 leads to an origin firing event, the trans-acting factor is sequestrated by replication forks and hence the number of available trans-acting factors 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 (not replicated) are able to fire. Before the beginning of replication process the one-dimensional lattice is randomly segmented into θL blocks where the probability of origin firing is P_{in} and $(1-\theta)L$ blocks where the probability of origin firing is P_{out}. After the start of replication process, at each round of calculus, each block is randomly assigned a value between 0 and 1. This value is compared to P_{in} or P_{out} (depending to which category the block belongs) 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 MM3 and MM5, we consider that the probability of origin firing P_{local} may be increased downstream of a replication fork over a distance d. 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 a 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.* [35]. 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 α^{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 (Fig 1.b). 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.

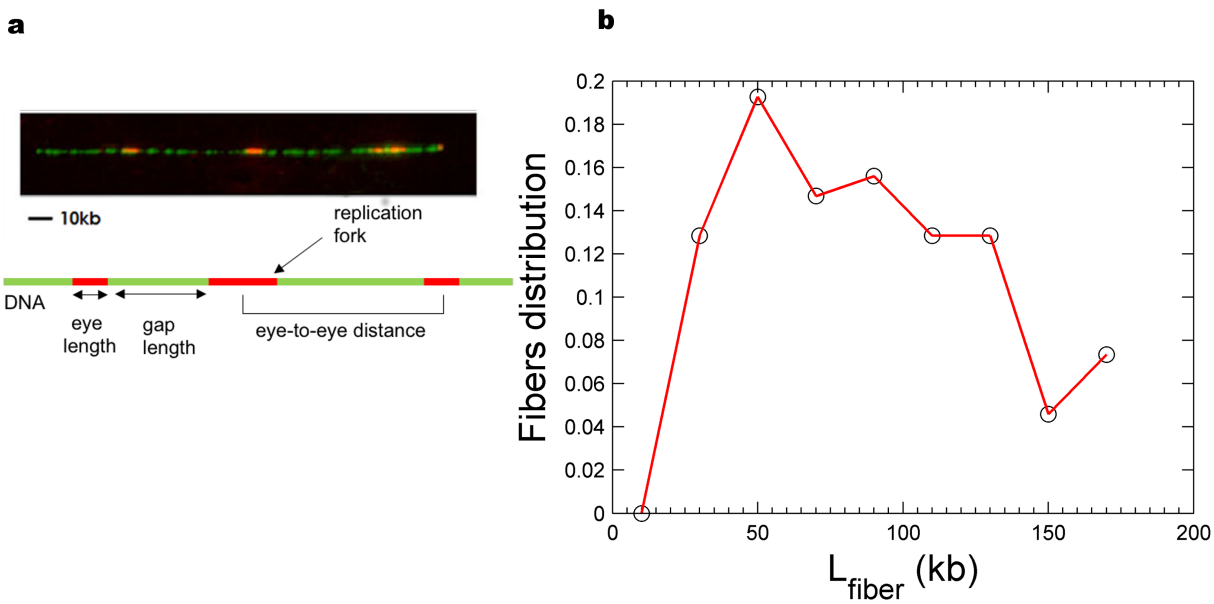

Fig 1. Characteristics of combed DNA molecules. a. Example of combed DNA molecule. The top panel is a fluorescence microscopy of a representative, stretched DNA fiber (green) containing replication eyes (red). The bottom panel is a schematic illustration of measured parameters in replication studies using DNA combing. **b**. Molecular length distribution (global replicated fraction of 8%) of combed DNA fibre. The black open circles are the experimentally measured and the red curve is the simulated cut molecular length distributions, respectively.

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 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 fibre, 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

Results

Finding the best integrative model of unperturbed S phase

Our previous model [35] failed to simultaneously reproduce the eye-to-eye distance distribution and the $I(f)$ of the same control experiment (Fig 2). This discrepancy could be explained if initiation events have a strong tendency to cluster [25, 27]. 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 their finite length [27]. Chk1 action has been proposed to regulate origins clusters [33]. However, Chk1 inhibition by UCN-01 did not result in the broader eye-to-eye distribution predicted by random origin firing (Fig 2 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 MM5) (S1 Appendix). MM1 corresponds to a mean field hypothesis of origin firing : all potential origins have a constant firing probability

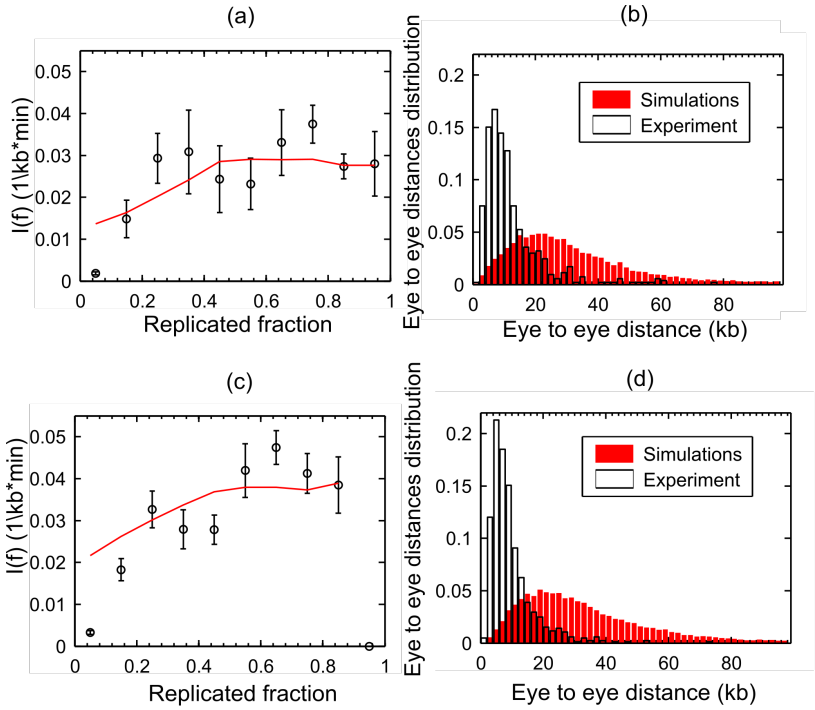

Fig 2. Chk1 does not control origin clustering. The black symbols are experimental data and the red curves are simulations. (a) and (c) Fitting of $I(f)$ data extracted from raw data published in [35] as described in material and methods for control and Chk1 inhibition experiments respectively. The discrepancy in values between the extracted data and those published in [35] are due to difference in thresholding and the lack of smoothing of the extracted data in this work. (b) and (d) Discrepancy between experimental and simulated distributions of eye-to-eye distances in control and Chk1 inhibition experiments, respectively.

P_{out} [38, 46]. MM2 corresponds to MM1 but assuming that replication forks can have a variable speed [47, 48]. MM3 corresponds to MM1 with a local perturbation, whereby the proximity of forks facilitates origin firing [30, 49] over a distance d downstream of an active fork where the probability of origin firing is P_{local}. In MM4 origin firing does not follow the mean field hypothesis but assumes that the genome can be segmented into regions of high and low probabilities of origin firing [31, 49] as accepted for most eukaryotes [8, 43, 50–54]. In this scenario, the probability of origin firing of potential origins located 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, MM5 combines the specific features of MM3 and MM4 into a single model. Furthermore, to verify if the localized nature of potential origins [43, 44] 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 except for MM2.

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. MM5 with localized potential origins (Fig 3) provided the best fit to the experimental data (S1 Appendix, Figure 9). The increase in concordance between MM5 and the data occurs at the expense of increasing the number of parameters, which is justifiable on

statistical grounds (S1 Appendix,TABLE 2) and the predictive ability of MM5 is verified (S2 Appendix).

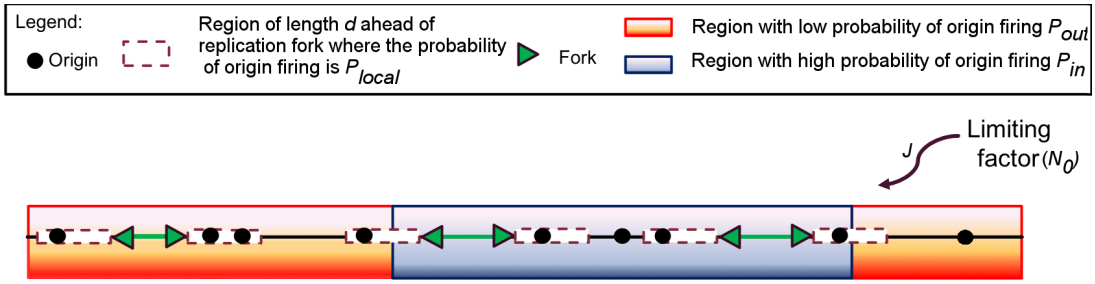

Fig 3. Schematic representation of MM5. Potential replication origins located in a fraction θ of the genome (not necessary contiguous) have a probability of firing P_{in} higher than probability of firing P_{out} of potential origins located in the complementary genome fraction $1-\theta$. The firing of a potential origins requires 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.

We used MM5 to analyse unchallenged, checkpoint inhibited and Chk1 over expressed S phase (S3 Appendix). In all cases MM5 was able to model concomitantly $I(f)$ and eye-to-eye distance distribution (Fig 4). In conclusion, while MM5 does not include all the possible mechanisms involved in DNA replication process and its regulation, it can adequately predict the spatio-temporal dynamics of DNA replication and its regulation by checkpoint mechanisms using a limited number of processes.

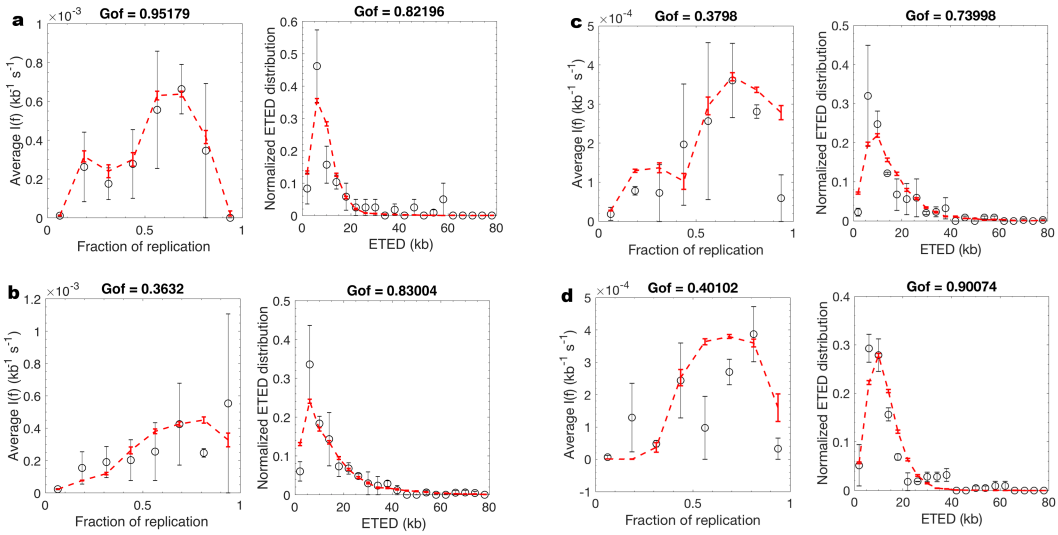

Fig 4. MM5 captures the essential processes necessary to model the regulation of DNA replication by Chk1. **a and b**. Unchallenged (8% global replication fraction) and Chk1 inhibited samples (22% global replication fraction) corresponding to the same experiment and harvested at same time. **c and d**. Unchallenged (46% global replication fraction) and Chk1 over expressed samples (22% global replication fraction) corresponding to the same experiment and harvested at same time. The black open circles are experimental data and the dashed red lines are the fit obtained by MM5 model

Retrieving the dynamics of an unchallenged S phase using the MM5 model

MM5 faithfully reproduced the temporal and spatial program of DNA replication from unperturbed S phase samples with global replicated fractions of 8%, 19% and 53% (S1 Appendix, Figure 9; S3 Appendix, Figures 1 and 2). The fitted values of parameters changed as S phase progressed (Fig 5). However, only changes in J, θ, P_{out} and d were

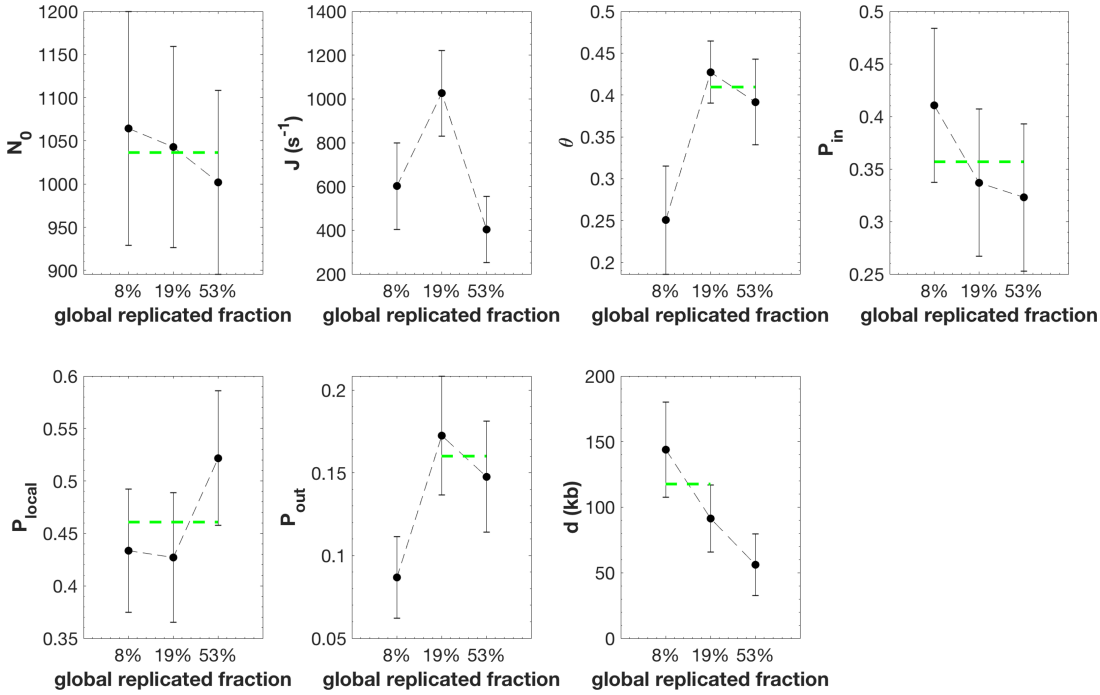

Fig 5. Inferred model parameters by fitting unchallenged S phase data as global replicated fraction increases. The black circles are the averaged value of the parameter over 100 independent fitting processes and the error bars are standard-deviations. The green dashed line is the mean value among consecutive parameters which differences are not statistically significant (S3 Appendix Figure 3).

statistically significant (S3 Appendix Figure 3). In particular we found that J increased from 8% to 19% replication and then dropped 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 [49]. 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 [33, 35].

Modeling DNA replication under Chk1 inhibition and over expression

To decipher the regulation of origin firing by Chk1, we examined if the MM5 model could also reproduce the replication program observed when the intra-S phase checkpoint was perturbed by the specific Chk1 inhibitor UCN-01 or by Chk1 over expression. We analyzed combed fibres from a replicated sample in the presence of UCN-01 (replicated fraction 22%) and in Chk1 over expression condition (replicated fraction 22%) that had spent the same interval of time in S phase as the control sample (global replicated fraction of 8% for UCN-01 and 46% in presence of Chk1 over expression). The MM5 model reproduced the experimental observations very well (S3 Appendix, Figure 4 and Figure 7, $GoF_{global} = 0.85$ for UCN-01 and $GoF_{global} = 0.65$ for Chk1 over expression). The two parameters J and θ, 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) (Fig 6 and S3 Appendix Figure 5). P_{out} was higher in the UCN-01 treated sample than in the control samples with the same harvesting time but unchanged once comparing similar replicated fraction. In the same manner, J, and θ, were significantly lower in the Chk1 over expressed sample than in the control sample with the same incubation time (Fig 6 b and S3 Appendix Figure 8). However, P_{out} and the other parameters were unchanged compared to control samples.

MM5 belongs to the general family of KJMA models that probabilistically describes the state of a nucleating and growing system [55]. In this framework, probabilities describing the nucleation are analogous to the probabilities of origin firing [56] and their values only depend on the parameters that describe the state of the system that in our case only the global fraction of replicated DNA. Hence, It seems natural that for two samples with the same replication fraction the values of probabilities P_{out}, P_{in} and P_{local} remain unchanged.

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, MM5 detected that Chk1 inhibition by UCN-01 increased origin firing [34, 35, 57–60]. However, upon Chk1 over expression (i) the fraction θ of the genome decreases, (ii) P_{out} is insensitive to Chk1 over expression and (ii) the import/activation rate of the limiting factor, J, is decreased, while the starting number of factors, N_0, is unaffected. As was expected, MM5 detected that Chk1 over expression reduced the number of fired origins [35].

In conclusion , the level of 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 [34, 57, 58, 61] at the start of S phase, and (iii) by controlling the import/activation rate of limiting firing factors [34]. However, no significant differences in the strength of origin regulation by nearby forks (P_{local}) was observed after Chk1 inhibition or over expression, suggesting that this local action is not mediated by Chk1 [33, 39].

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

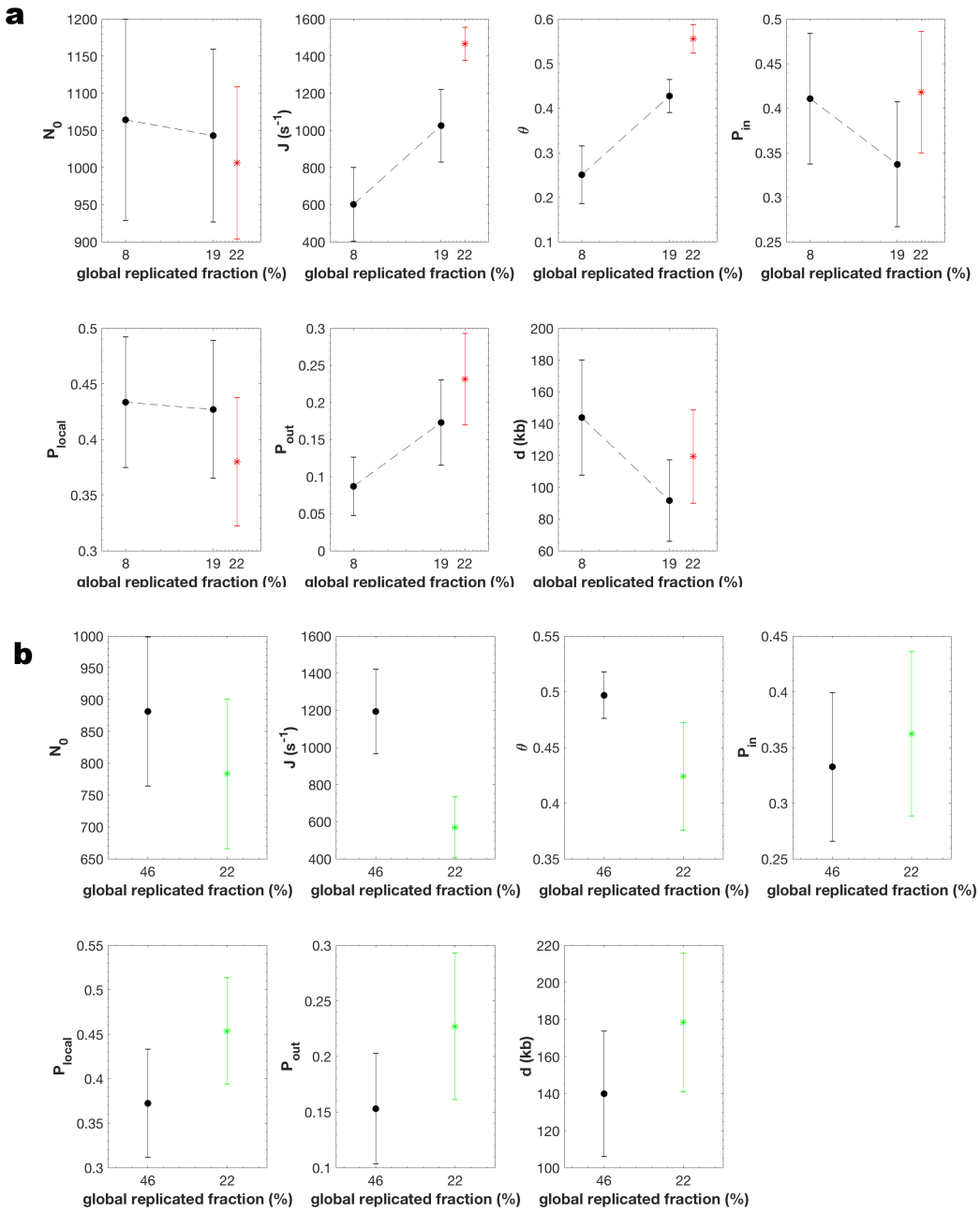

Fig 6. J, θ, and the P_{out} are the only parameters that change when comparing unchallenged, (a) Chk1 inhibited and (b)Chk1 over-expressed S phase The black circle is the averaged value of the parameter over 100 independent fitting processes of unchallenged S phase and the error bars are standard-deviations. The red star (**a**) is the averaged value of the parameter over 100 independent fitting processes of Chk1 inhibited sample and the error bars represent the standard-deviations. The green star (**b**) is the averaged value of the parameter over 100 independent fitting processes of Chk1 over-expressed sample and the error bars represent the standard-deviations.

distributions of replication tracks as analyzed by DNA combing at different stages of S phase. We found that model MM5 with discrete potential origins best reproduced the experimental data with a minimal number of adjustable parameters. This model combines five assumptions [29, 31, 35, 38, 43, 44, 46, 49, 62, 63]: 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 [49, 64] ; 5) the probability of origin firing is heterogeneous along the genome [31, 43].

We used MM5 to model DNA combing data from *Xenopus* egg extracts in presence or absence intra-S checkpoint inhibition and activation. In all 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 [34, 59, 65] and decreased after Chk1 over expression. Importantly, this model assumes a heterogeneous probability of origin firing and suggests that Chk1 exerts a global origin inhibitory action during unperturbed S phase [35] by following two possible mechanisms: (i) the first path corresponds to the regulation of the number of available replication limiting factors by Chk1 protein and (ii) the second path corresponds to the ability of Chk1 protein to reduce the capacity of potential origins to fire outside domains with high probability of origin firing. The strength of the second path decreases from the beginning of S phase to reach its minimal value after the first quarter of S phase. On the other hand, the constancy of the initial number of limiting factors N_0 in the presence or absence of UCN-01 or Chk1 over expression suggests that Chk1 does not actively control origins or the available number of replication limiting factors before S phase actually starts [36, 66, 67]. Interestingly, a better statistical match between the model and the data was obtained by assuming that the rate of DNA synthesis is variable downstream of replication forks. Indeed, the downstream of a replication forks the rate of DNA synthesis depends on the speed of replication fork and the frequency of firing of closeby potential replication origins [55]. Our analysis suggests that this variability cannot be mapped to a model with variable fork speed, but it is compatible with an increased probability of origin firing in the neighbourhood of an active replication fork. These observations indicate that MM5 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 [35]. Therefore, the earliest origins must be immune to Chk1 inhibition while later potential origins are strongly inhibited. The comparisons among the modelling of Chk1 inhibition, over expression 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 [68], *Drosophila* [60] and *Xenopus* [69] 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* [70], 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 [36]. 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 idea that the rate of DNA synthesis depends on the S-phase time and position of replication forks. During our modelling based on statistical ground we showed that the domino-like view of DNA replication progression [49, 64] better describes the measured quantities from combed DNA molecules than the hypothesis of variable fork velocity [71]. 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 [30]. While this could in principle reflect an origin firing exclusion zone ahead of forks [23, 49], our model did not allow for a negative P_{local} but the fact that the discrete distribution of potential replication origins better describes the experimental data than the continuous distribution confirm the necessity of the existence of origin firing exclusion zones between two converging replication forks. 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) [35, 39]. However, we find that the range, d, and the strength, P_{local}, of origin stimulation by nearby forks, were both insensitive to checkpoint inhibition or activation (Fig 6 a and b). Other potential mechanisms such as propagation of a supercoiling wave ahead of forks may better explain this insensitivity to Chk1 inhibition [72].

Heterogeneous probability of origin firing

In this model, the origin firing process in *Xenopus* egg extracts is not reliably 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 low and 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 [61]. 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 [32]. 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 (Fig 5 and Fig 6), 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 (Fig 6 a) or have spent the same time interval in S phase in a Chk1 over-expressed condition (Fig 6 b), we noticed that the probability of origin firing in the Chk1-immune domains (P_{in}) did not change upon Chk1 inhibition or over expression. This further suggests that these domains actually escape the regulation of origin firing by Chk1 that rules the rest of the genome. It is an interesting observation that in Chk1-immune regions where the probability of origin firing is high, the temporal difference between two firing events would be smaller than in other regions of the genome. This leads to an observed synchrony of origin firing and therefore to an effective observed clustering of replication eyes on a single DNA fibre.

Conclusion

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 [54] encourages us also to assume the hypothesis that in *Xenopus* early embryos the position of early replication domains are conserved from one nucleus to another leading not to a flat but structured mean replication timing pattern similar to other eukaryotic systems [6, 8, 51].

Supporting information

S1 Appendix. Choosing the best model. In this appendix we describe in detail why MM5 is statistically the best model to describe the data.

S2 Appendix. Testing the predictive capacity of MM5. In this appendix we test the predictive capacity of the fitting procedure coupled to MM5.

S3 Appendix. Fitting the experimental data with MM5. In this appendix we extract the values of free parameters from experimental data for unchalanged, Chk1 inhibited and Chk1 over-expressed S phase and discuss their variations.

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

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

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

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

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

5. 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;2(8). doi:10.1371/journal.pone.0000722.

6. Baker A, Audit B, Chen CL, Moindrot B, Leleu A, Guilbaud G, et al. Replication Fork Polarity Gradients Revealed by Megabase-Sized U-Shaped Replication Timing Domains in Human Cell Lines. PLoS Comput Biol. 2012;8(4). doi:10.1371/journal.pcbi.1002443.

7. Audit B, Baker A, Chen CL, Rappailles A, Guilbaud G, Julienne H, et al. Multiscale Analysis of Genome-Wide Replication Timing Profiles Using a Wavelet-Based Signal-Processing Algorithm. Nat Protoc. 2013;8(1):98–110. doi:10.1038/nprot.2012.145.

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

9. 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;18(3):231–239. doi:10.1016/j.ceb.2006.04.001.

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

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

12. 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;30(23):4805–4814. doi:10.1038/emboj.2011.404.

13. 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;21(24):2055–2063. doi:10.1016/j.cub.2011.11.038.

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

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

16. 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;20(7):1457–1462.

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

18. 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;60(7-8-9):297–304. doi:10.1387/ijdb.160139oh.

19. 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.

20. 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;12(12):4511–4520. doi:10.1002/j.1460-2075.1993.tb06140.x.

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

22. 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.

23. 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;296(3):769–786. doi:10.1006/jmbi.2000.3500.

24. 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;300(5):1133–1142. doi:10.1006/jmbi.2000.3930.

25. 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;152(1):15–26.

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

27. 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;279(27):28071–28081. doi:10.1074/jbc.M401574200.

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

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

30. 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;3(2):223–229.

31. 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;10(3):722. doi:10.1002/msb.134859.

32. 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;36(17):5623–5634. doi:10.1093/nar/gkn533.

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

34. 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;57(3):492–505. doi:10.1016/j.molcel.2014.12.003.

35. Platel M, Goldar A, Wiggins JM, Barbosa P, Libeau P, Priam P, et al. Tight Chk1 Levels Control Replication Cluster Activation in Xenopus. PLoS One. 2015;10(6). doi:10.1371/journal.pone.0129090.

36. Forey R, Poveda A, Sharma S, Barthe A, Padioleau I, Renard C, et al. Mec1 Is Activated at the Onset of Normal S Phase by Low-dNTP Pools Impeding DNA Replication. Molecular Cell. 2020;doi:10.1016/j.molcel.2020.02.021.

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

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

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

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

41. 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;188(2):209–221. doi:10.1083/jcb.200911037.

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

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

44. 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;7:e35192. doi:10.7554/eLife.35192.

45. 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;277(36):33049–33057. doi:10.1074/jbc.M204438200.

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

47. Conti C, Saccà B, Herrick J, Lalou C, Pommier Y, Bensimon A. Replication Fork Velocities at Adjacent Replication Origins Are Coordinately Modified during DNA Replication in Human Cells. Mol Biol Cell. 2007;18(8):3059–3067. doi:10.1091/mbc.E06-08-0689.

48. Yousefi R, Rowicka M. Stochasticity of Replication Forks' Speeds Plays a Key Role in the Dynamics of DNA Replication. PLOS Computational Biology. 2019;15(12):e1007519. doi:10.1371/journal.pcbi.1007519.

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

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

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

52. 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;25(12):1886–1892. doi:10.1101/gr.195305.115.

53. Petryk N, Kahli M, d'Aubenton-Carafa Y, Jaszczyszyn Y, Shen Y, Silvain M, et al. Replication Landscape of the Human Genome. Nat Commun. 2016;7(1):10208. doi:10.1038/ncomms10208.

54. 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;27(8):1406–1416. doi:10.1101/gr.218602.116.

55. Jun S, Bechhoefer J. Nucleation and Growth in One Dimension. II. Application to DNA Replication Kinetics. Phys Rev E. 2005;71(1):011909. doi:10.1103/PhysRevE.71.011909.

56. Jun S, Zhang H, Bechhoefer J. Nucleation and Growth in One Dimension. I. The Generalized Kolmogorov-Johnson-Mehl-Avrami Model. Phys Rev E. 2005;71(1):011908. doi:10.1103/PhysRevE.71.011908.

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

58. 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;218(9):2865–2875. doi:10.1083/jcb.201902085.

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

60. 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;38(4):399–412. doi:10.1016/j.devcel.2016.07.023.

61. 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;26(11):2719–2731. doi:10.1038/sj.emboj.7601714.

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

63. 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;6(1). doi:10.1038/msb.2010.8.

64. Guilbaud G, Rappailles A, Baker A, Chen CL, Arneodo A, Goldar A, et al. Evidence for Sequential and Increasing Activation of Replication Origins along Replication Timing Gradients in the Human Genome. PLOS Computational Biology. 2011;7(12):e1002322. doi:10.1371/journal.pcbi.1002322.

65. 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;27(16):5806–5818. doi:10.1128/MCB.02278-06.

66. 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;16(18):2327–2332. doi:10.1101/gad.1013502.

67. 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;158(5):863–872. doi:10.1083/jcb.200204127.

68. Moiseeva TN, Yin Y, Calderon MJ, Qian C, Schamus-Haynes S, Sugitani N, et al. An ATR and CHK1 Kinase Signaling Mechanism That Limits Origin Firing during Unperturbed DNA Replication. PNAS. 2019;116(27):13374–13383. doi:10.1073/pnas.1903418116.

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

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

71. Ciardo D, Goldar A, Marheineke K. On the Interplay of the DNA Replication Program and the Intra-S Phase Checkpoint Pathway. Genes. 2019;10(2):94. doi:10.3390/genes10020094.

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

Appendix 1

1 Different models

To model experimental observations a series of nested models (Table 1) were compared with experimental data. Below are the fits of each model to experimental sample with

Table 1. The nested models and their variables

Model	N_0	J	P_{out}	P_{in}	P_{local}	θ	d	v
MM1	√	√	√					
MM2	√	√	√					√
MM3	√	√	√		√		√	
MM4	√	√	√	√		√		
MM5	√	√	√	√	√	√	√	

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 ($GoF = 1 - \frac{1}{N}\sum_{i=1}^{N}\frac{\left(y_{fit}^{i}-y_{exp}^{i}\right)^2}{\left(y_{exp}^{i}-\langle y_{exp}\rangle\right)^2}$, where $\langle y_{exp}\rangle$ represents average value of the experimental data set and N is the number of data points). 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} (Table 2), the model that best described the whole data set is the MM5 with localized distribution of potential origins: its GoF_{global} value is closest to one. However, MM5 also has the highest number of fitting variables (7) compared to other models (MM1 has 3 fitting variables, MM2 has 4 fitting variables,MM3 and MM4 have 5 fitting variables), and facilitating fit to the data.

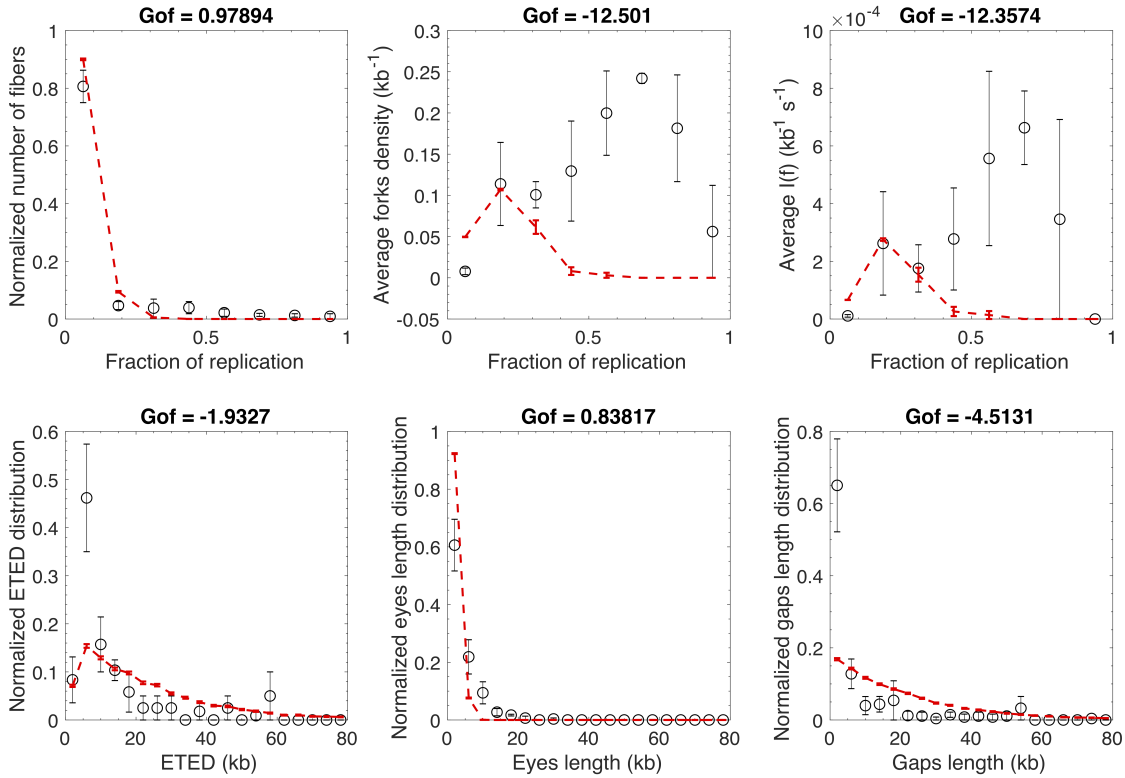

Fig 1. Modeling experimental data with MM1 model in the case where the potential origins are continuously distributed along the genome. Open circles are experimental data and the red dashed line is the fit.

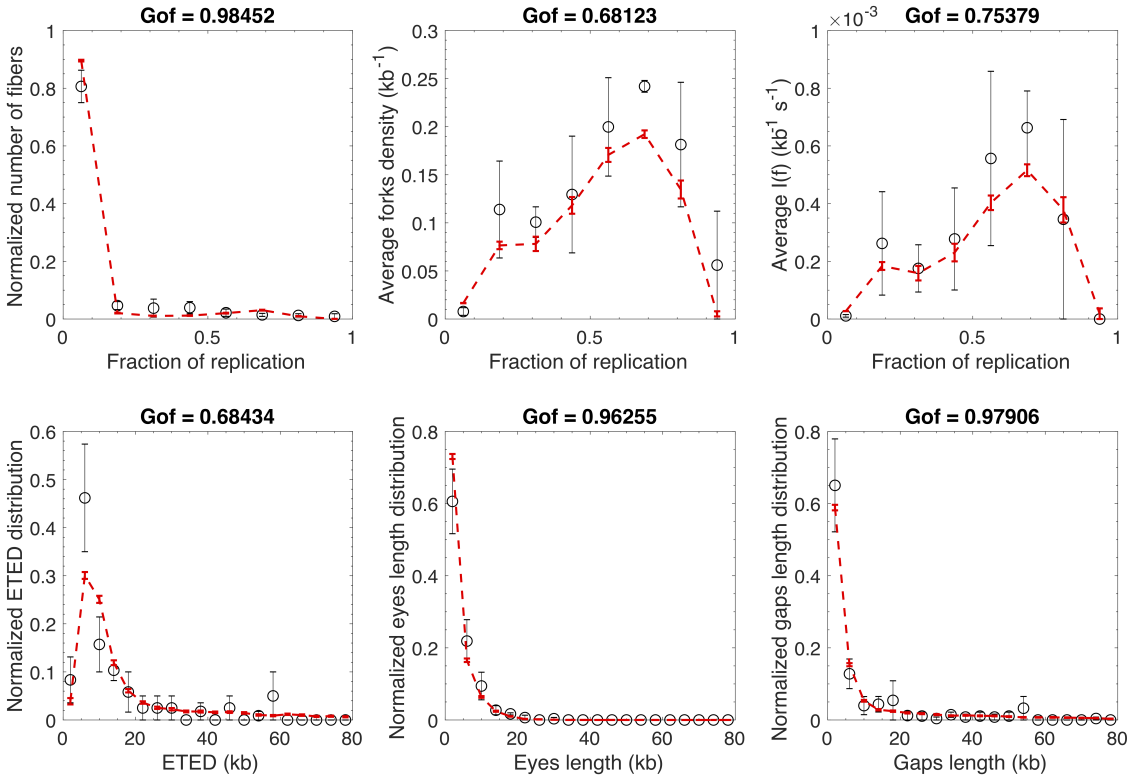

Fig 2. Modeling experimental data with MM3 model in the case where the potential origins are continuously distributed along the genome. Open circles are experimental data and the red dashed line is the fit.

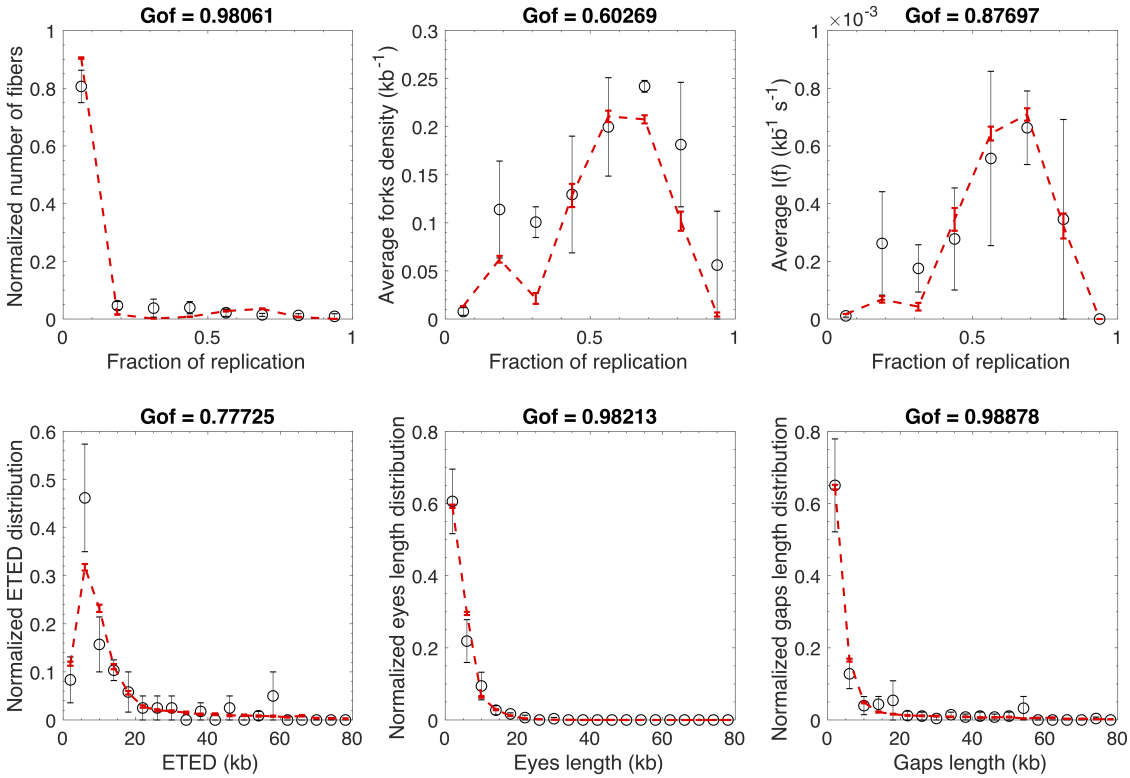

Fig 3. Modeling experimental data with MM4 model in the case where the potential origins are continuously distributed along the genome. Open circles are experimental data and the red dashed line is the fit.

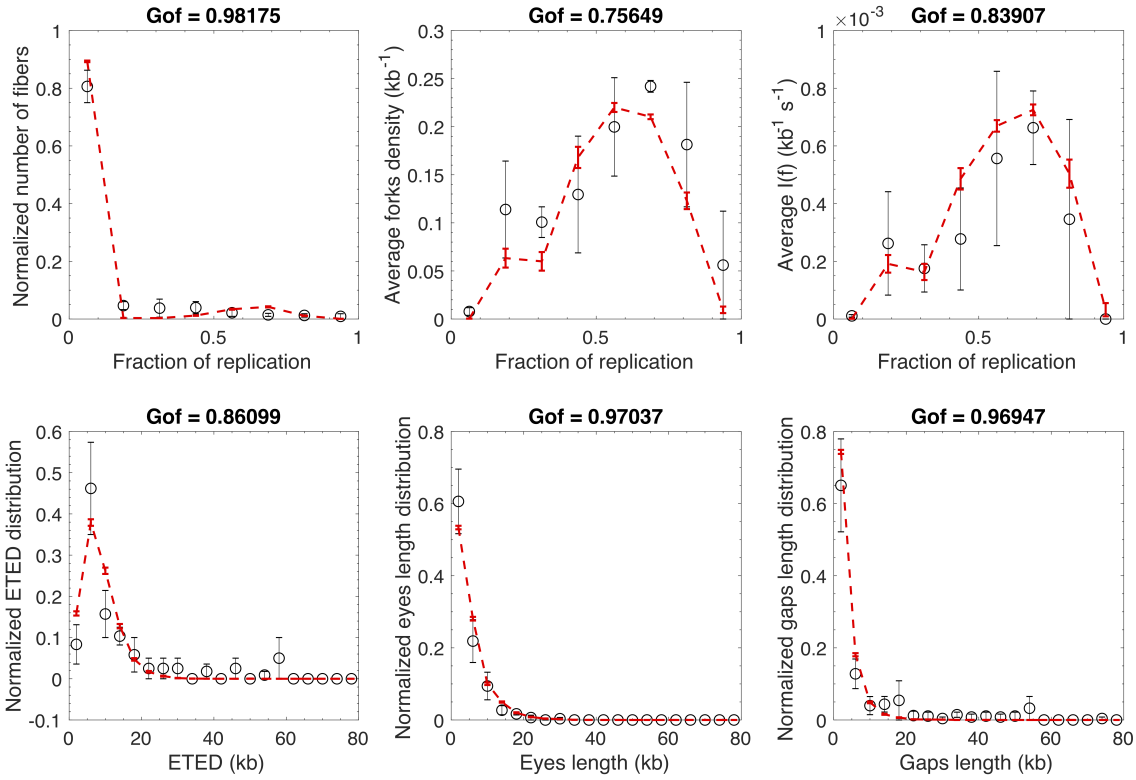

Fig 4. Modeling experimental data with MM5 model in the case where the potential origins are continuously distributed along the genome. Open circles are experimental data and the red dashed line is the fit.

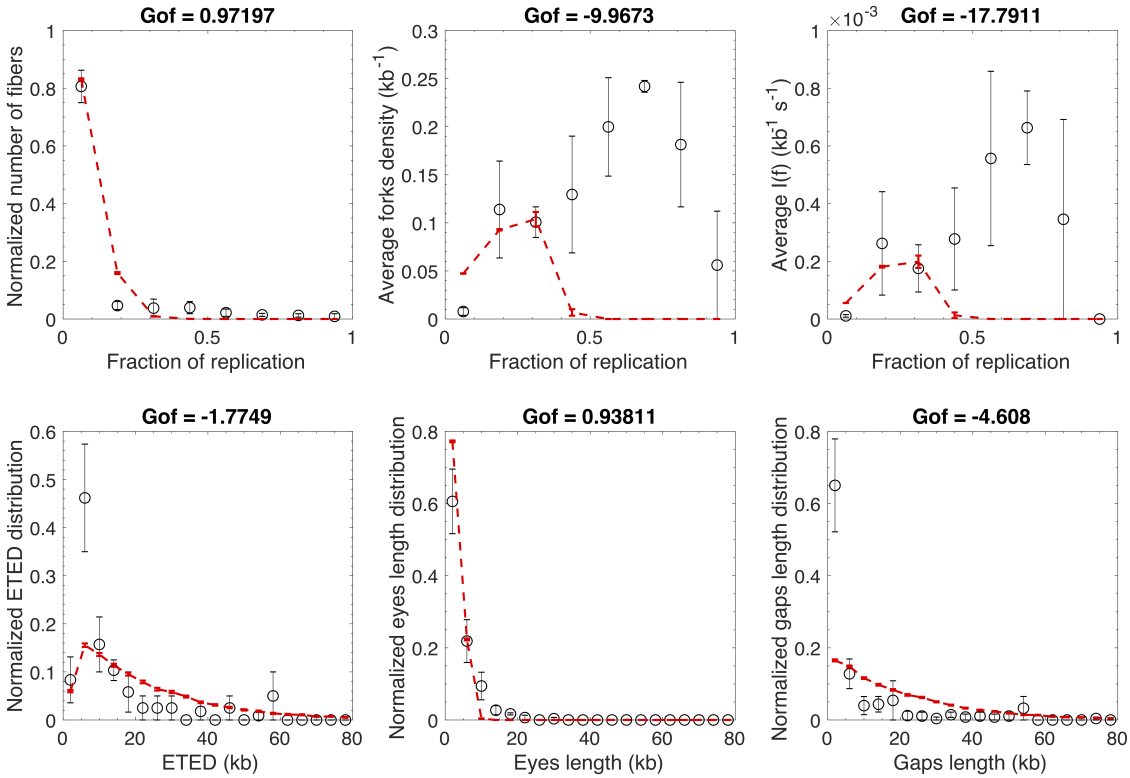

Fig 5. Modeling experimental data with MM1 model in the case where the potential origins form a discrete set along the genome. Open circles are experimental data and the red dashed line is the fit.

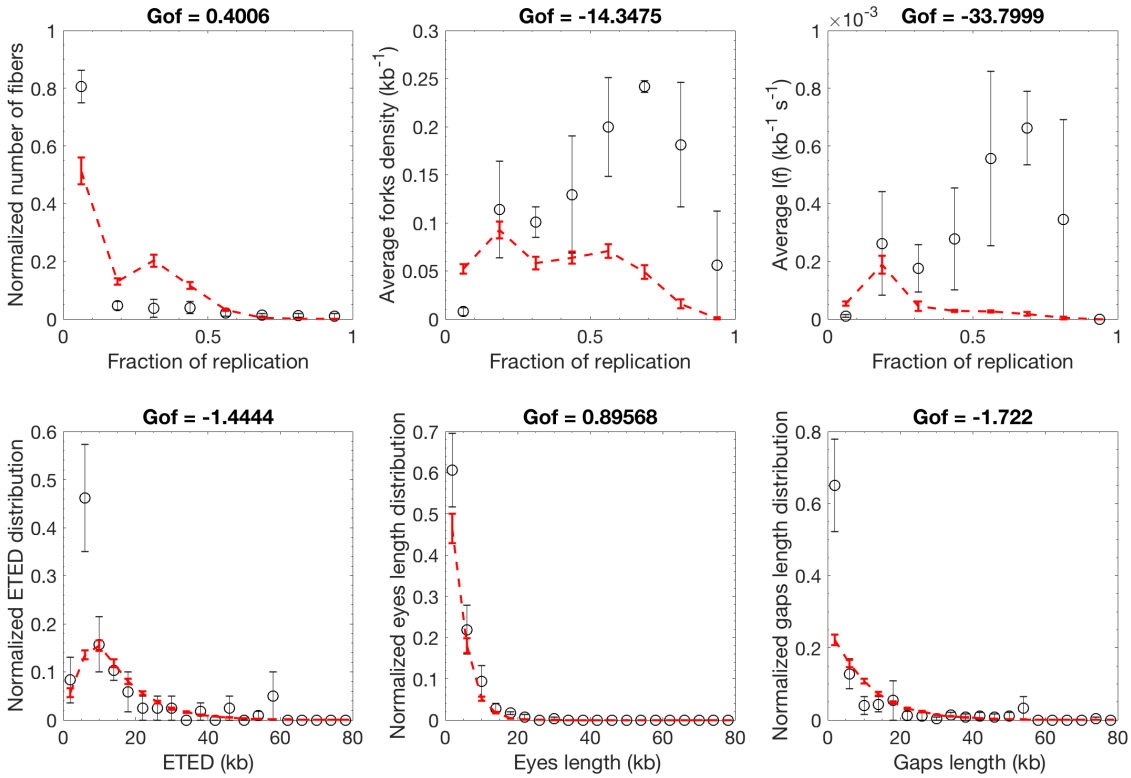

Fig 6. Modeling experimental data with MM2 model in the case where the potential origins form a discrete set along the genome. Open circles are experimental data and the red dashed line is the fit.

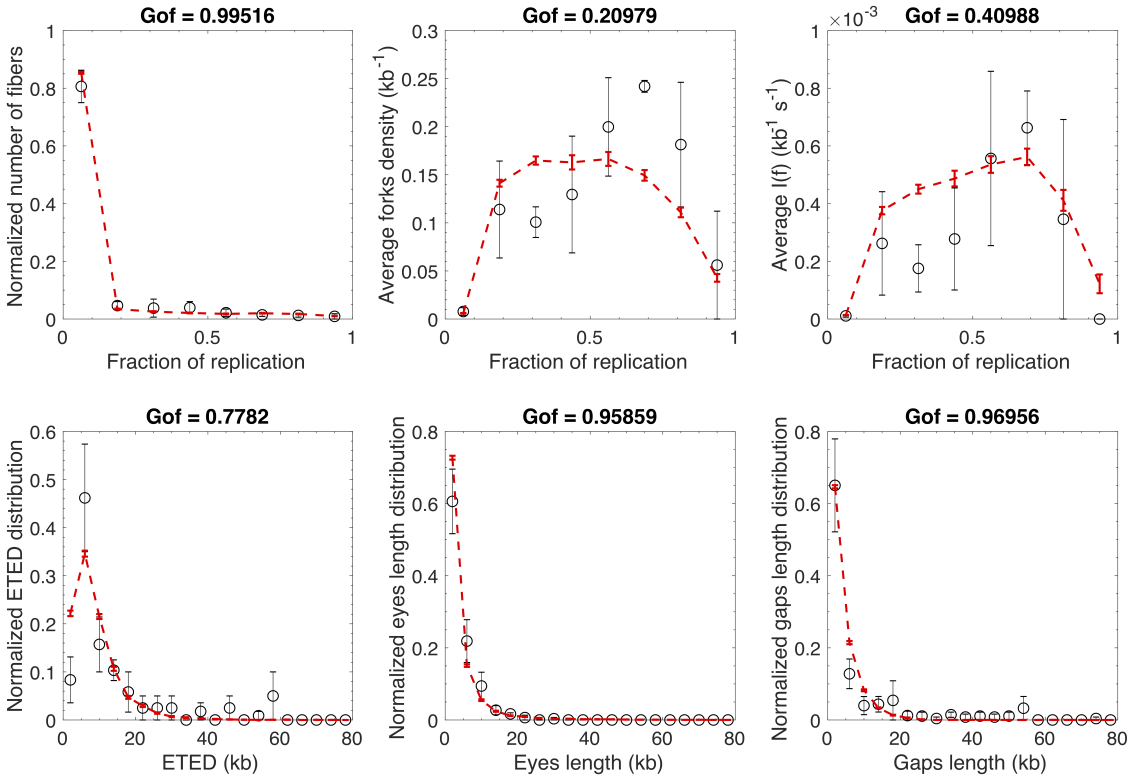

Fig 7. Modeling experimental data with MM2 model in the case where the potential origins form a discrete set along the genome. Open circles are experimental data and the red dashed line is the fit.

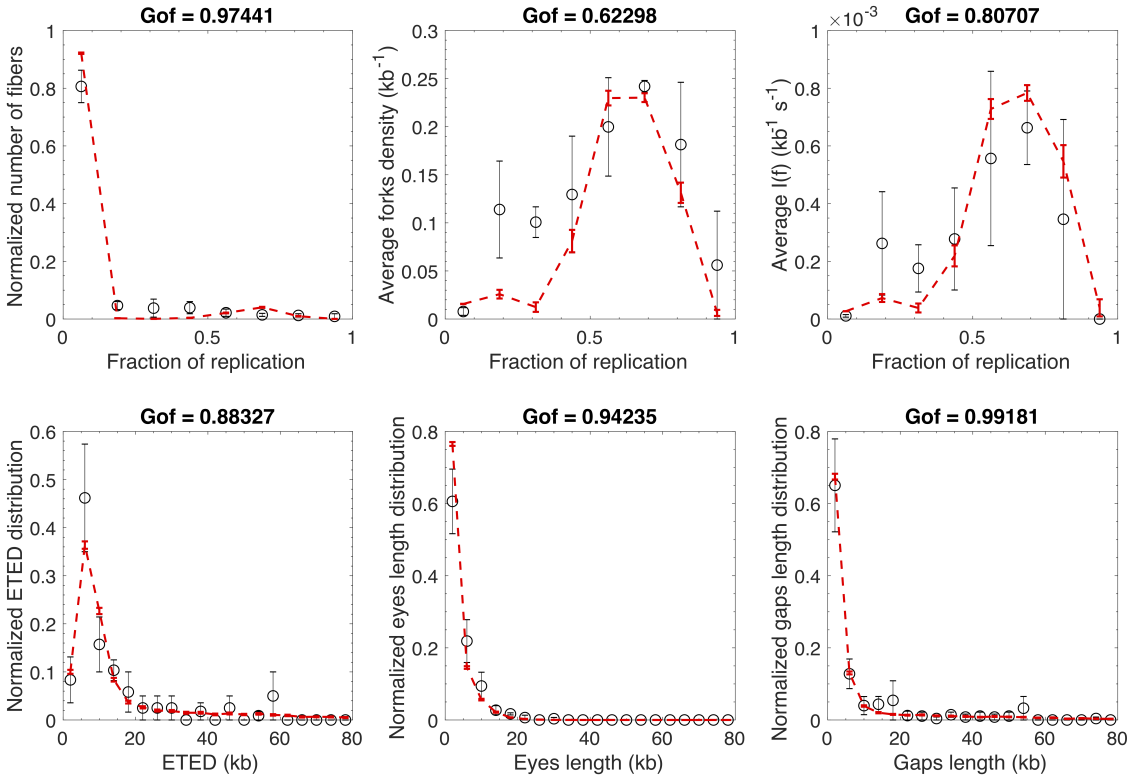

Fig 8. Modeling experimental data with MM4 model in the case where the potential origins form a discrete set along the genome. Open circles are experimental data and the red dashed line is the fit.

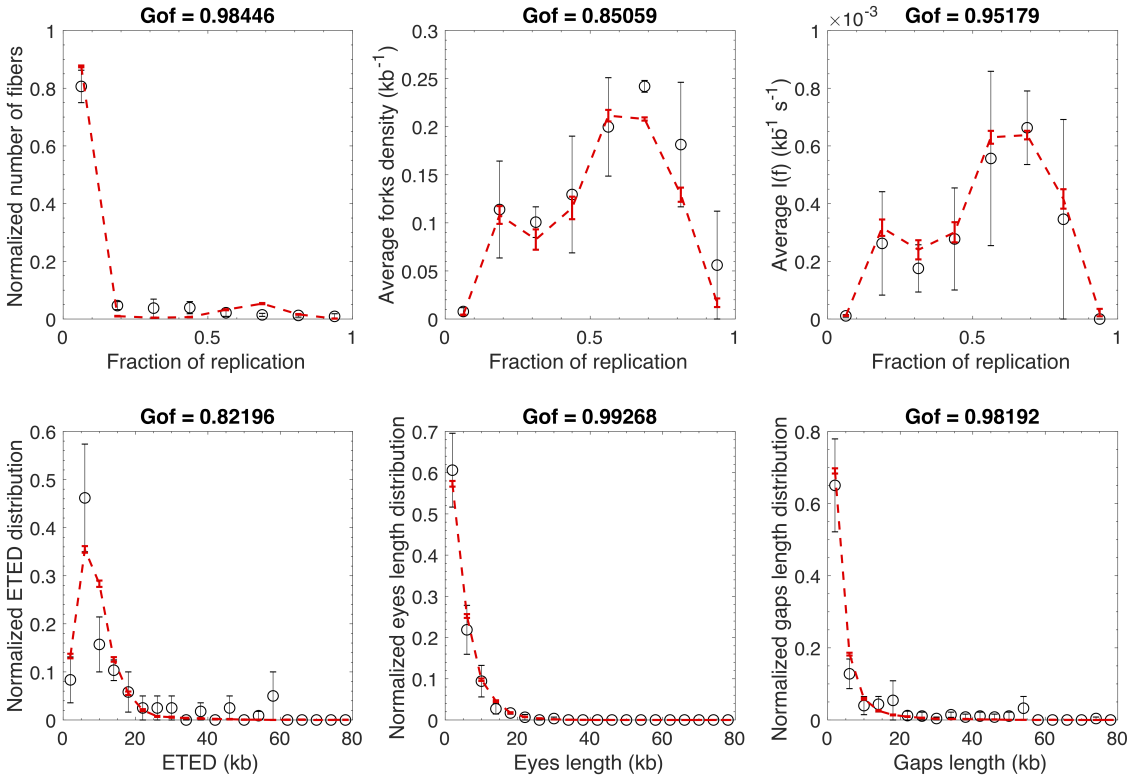

Fig 9. Modeling experimental data with MM5 model in the case where the potential origins form a discrete set along the genome. Open circles are experimental data and the red dashed line is the fit.

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

	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		-8.34		0.53
MM3	0.85	0.72	0.08	0.10
MM4	0.87	0.88	0.08	0.09
MM5	0.90	0.92	0.08	0.05

2 Models comparison

To address whether the better data fit with MM5 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 [1] and the Akaike's criterion (ΔAIC) that is based on information theory [2]. We can objectively reject MM1 and MM2 as they did not reproduce in a satisfactory manner the averaged fork density, $I(f)$ and eye-to-eye distances distributions (Figures 1, 5 and 6). MM3 and MM4 satisfactorily reproduced all measured quantities (Figures 2 , 3, 7 and 8) but with lower GoF_{global} value than the MM5 models (Table 2). The discrete MM5 model has higher GoF_{global} value than the continuous one, whereas the continuous MM3 and MM4 models were better than or equal to their discrete version, respectively (Table 2). To choose the best model, we compared the discrete MM5 model, continuous MM3, MM4 and MM5 corresponding to fits with highest GoF_{global} values (Table 2). Comparing the discrete MM5 with the continuous MM3, MM4 and MM5 models led in all cases to $F > 1$ with p-values $p < 10^{-6}$ and negative ΔAIC values (Table 3). To verify if the increase in model's complexity does always leads to $F > 1$ and negative ΔAIC values, we fitted the experimental data with the MM6 model as described in Appendix 2 (Figure 10, $GoF_{global} = 0.87$ and $(y_{exp} - y_{fit})^2 = 0.07$) that has 10 adjustable parameters.

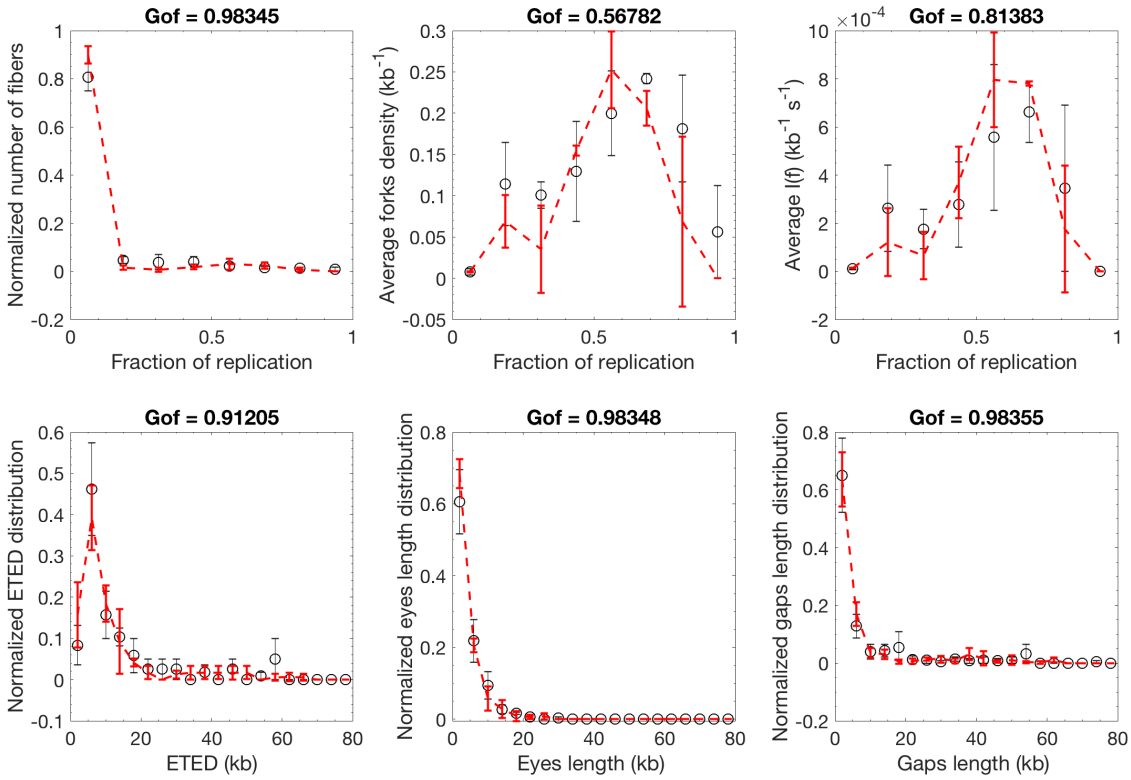

Fig 10. Modeling experimental data with MM6 model in the case where the potential origins form a discrete set along the genome. Open circles are experimental data and the red dashed line is the fit.

In this case $F < 1$ and $\Delta AIC > 0$ (Table 3), implying that MM6 is overfitting the

data. The discrete MM5 model is therefore the best model and the observed increase in GoF_{global} is not the consequence of increased complexity of the model.

Table 3. Values of F-test, the associated p-value (p) and the ΔAIC when the discrete MM5 model is compared with continuous MM3, MM4 and MM5 model.

model	F	p	ΔAIC
Continuous MM3	19.3	1.5×10^{-7}	-30.2
Continuous MM4	16.9	8.3×10^{-7}	-26.6
Continuous MM5	∞	Not defined	-31.1
MM6	-5.3	1	26.3

References

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

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

Appendix 2

The real DNA replication process is far more complex than any of the considered models. To explore how accurately MM5 can map a more complex process, we built, based on replication process in other eukaryotes [1–7] and our previous model [8], a more elaborate model (MM6) to generate *in silico* data with 8%, 19% and 53% global replicated fractions.

1 The MM6 model used to generate the *in silico* data

In MM6, 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 course 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 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 1 of materials and methods, 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.

Table 1. Values of MM5's parameters. These values are chosen arbitrarily.

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

2 Fitting the *in silico* data by MM5 model

By independently fitting the simulated profiles of each global replicated fraction, we implicitly assume that samples could originate from separated experiments, hence MM5 parameters values are possibly different for each global replicated fraction. This allows us to accurately reproduce observations from each sample (Figures 1, 2 and 3).

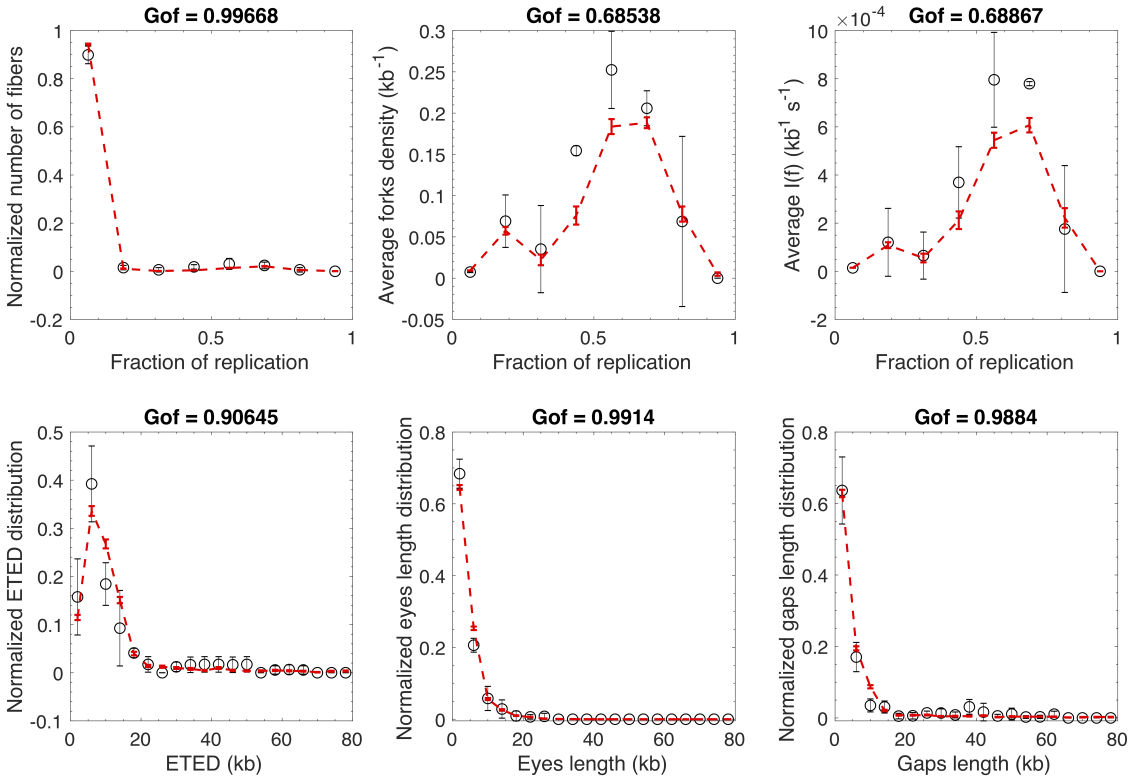

Fig 1. Modeling 8% global replicated fraction simulated data with discrete MM5 model. Open circles are simulated data and the red dashed line is the fit. $GoF_{global} = 0.96$

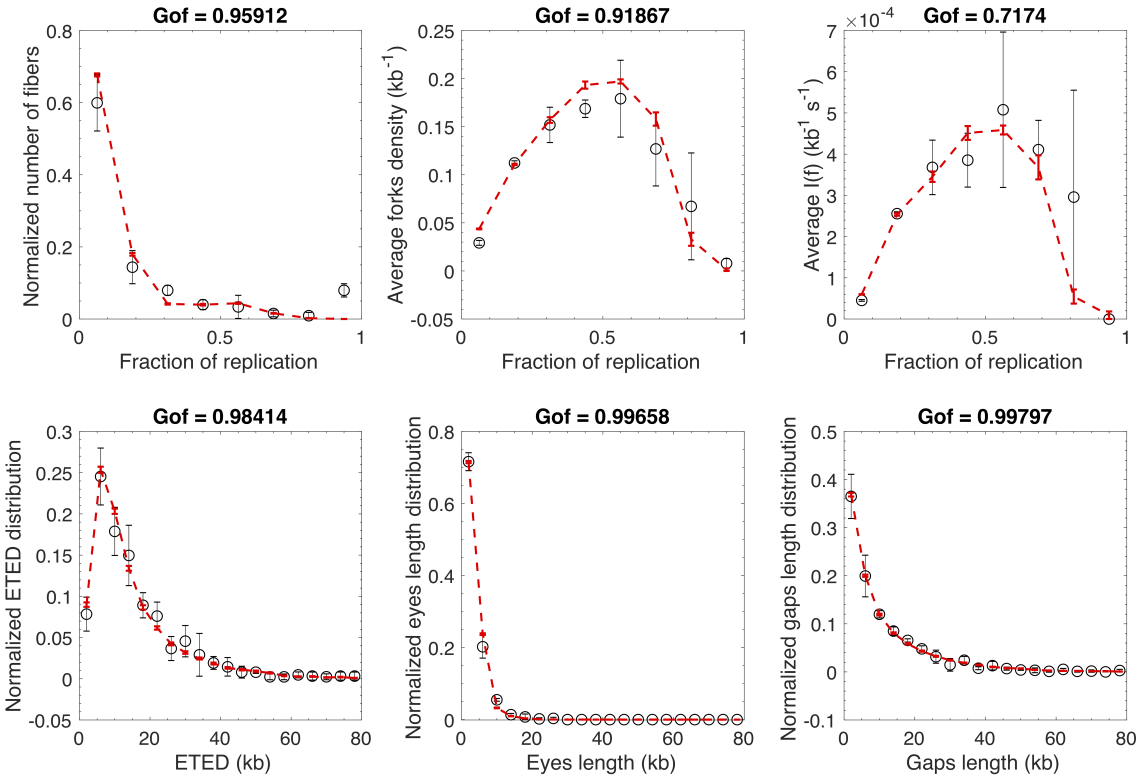

Fig 2. Modeling 19% global replicated fraction simulated data with discrete MM5 model. Open circles are simulated data and the red dashed line is the fit. $GoF_{global} = 0.97$

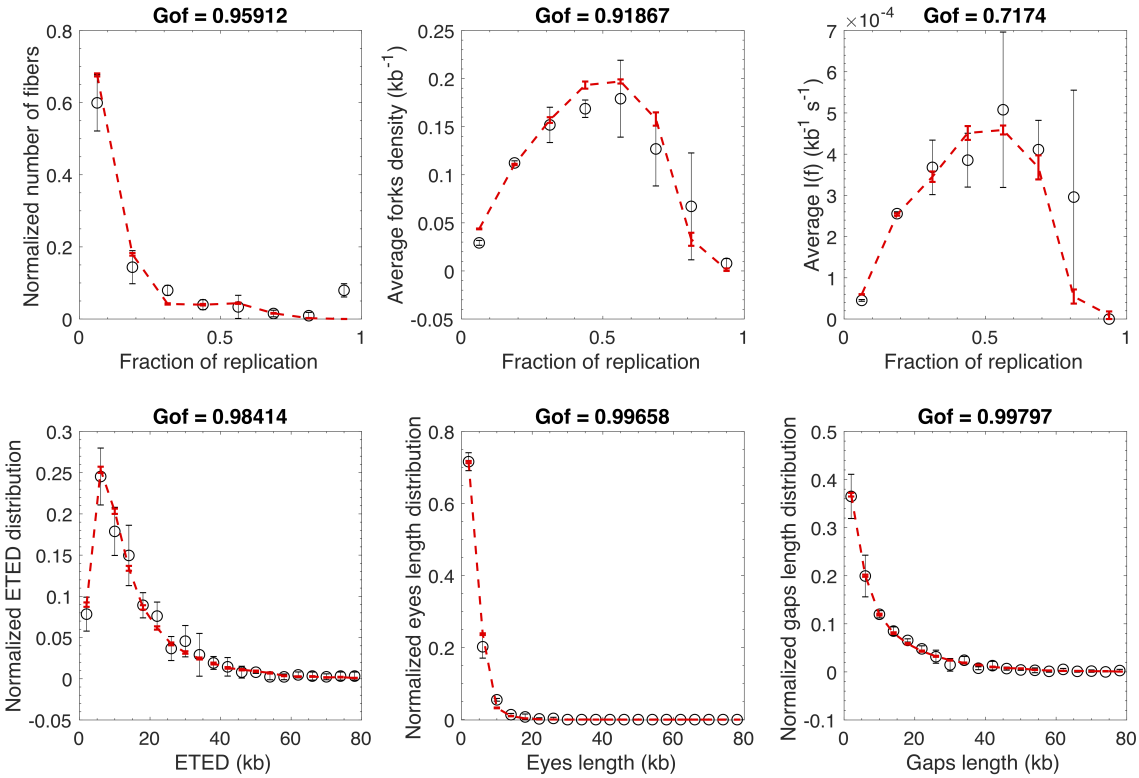

Fig 3. Modeling 53% global replicated fraction simulated data with discrete MM5 model. Open circles are simulated data and the red dashed line is the fit. $GoF_{global} = 0.82$

3 Reduction of MM6 to MM5

In the MM6 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$P_{in} = \frac{1}{2}(P_0 + P_{dom})$, 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 $P^{in}_{local} = \frac{1}{2}(P_0 + P_{dom} + P_{fork})$. Outside these domains, the global probability of origin firing is modulated by the probability of origin inhibition $P_{out} = \frac{1}{2} P_0 (1 - P_{inhib})$. 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 $P^{out}_{local} = \frac{1}{2}(P_0 + P_{fork})[1 + P_{inhib}(P_{deinhib} - 1)]$. Local probabilities of origin firing only influence origins over a distance d_{fork} downstream of a fork. The MM5 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 MM6 model, we combined the parameters of the model to retrieve the parameters of MM5 (TABLE 2). The values of these

Table 2. Reducing MM6 to MM5.

MM5	equivalence with MM6
N_0	N_0
J (s^{-1})	J
θ	$\frac{N_{dom} l_{dom}}{L}$
P_{in}	$\frac{1}{2}(P_0 + P_{dom})$
P_{local}	$\frac{1}{2}(P_0 + P_{fork})[1 + (1 - \theta) P_{inhib}(P_{deinhib} - 1)] + \theta P_{dom}$
P_{out}	$\frac{1}{2} P_0 (1 - P_{inhib})$
d (kb)	d

parameters can be compared directly to parameters of MM5 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 (Figure 4 and 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. Figure 5 shows that the differences of predicted parameters values among the 3 considered samples were not statistically significant, as was expected. All $t < 1$ and $\chi^2 < 1$ (Figure 5), 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}.

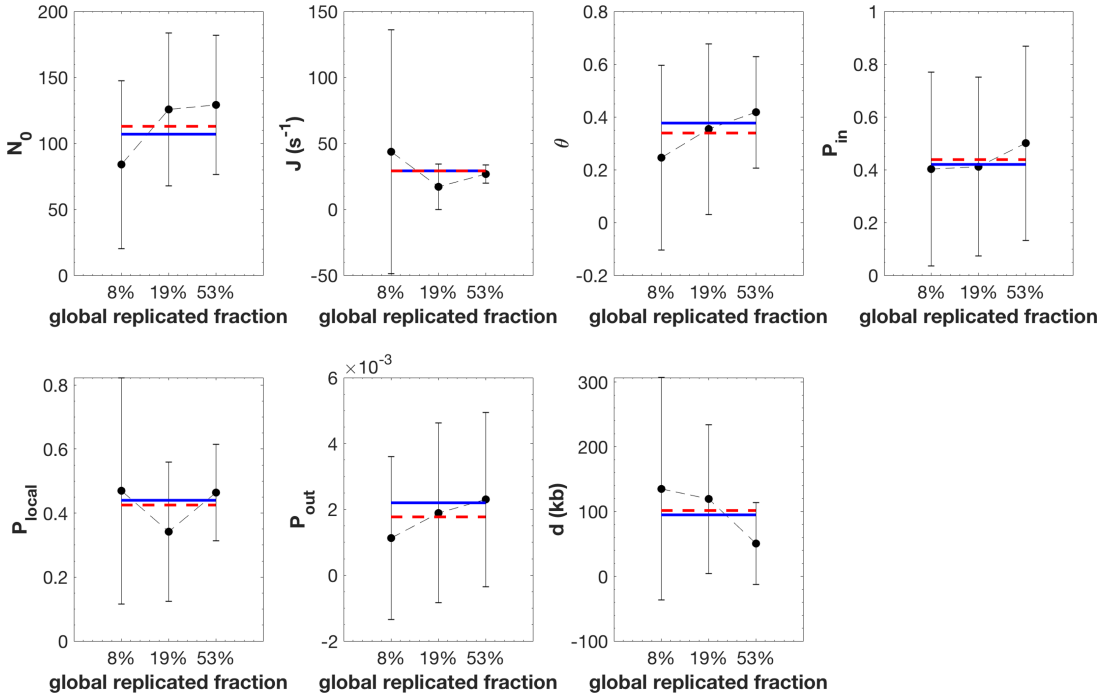

Fig 4. The fitting strategy infers accurately the expected values for the reduced MM6 free parameters. The black circles correspond to the averaged value of the parameter over 100 independent fits and the error bars are the standard-deviations. The solid blue line is the expected value of the parameter as obtained in TABLE 3. The red dashed line is the mean value of the parameter obtained by averaging the parameter inferred values over the 3 samples.

Table 3. Comparison between the expected and inferred values of MM5 parameters.

MM5	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 $(t < 1)$

The ability of the fitting procedure i) to circumscribe the values of MM5 model parameters close to the expected ones (TABLE 2) and ii) to retrieve the constancy of these parameter's values as the global degree of replication increases (Figure 5) demonstrates the adequacy of our fitting strategy to recover the dynamic of DNA replication during S phase in the framework of MM5 model by setting the null hypothesis as : the values of MM5 parameters do not change as S phase progresses. Therefore, rejection of this hypothesis for a considered parameter means its variation during S phase.

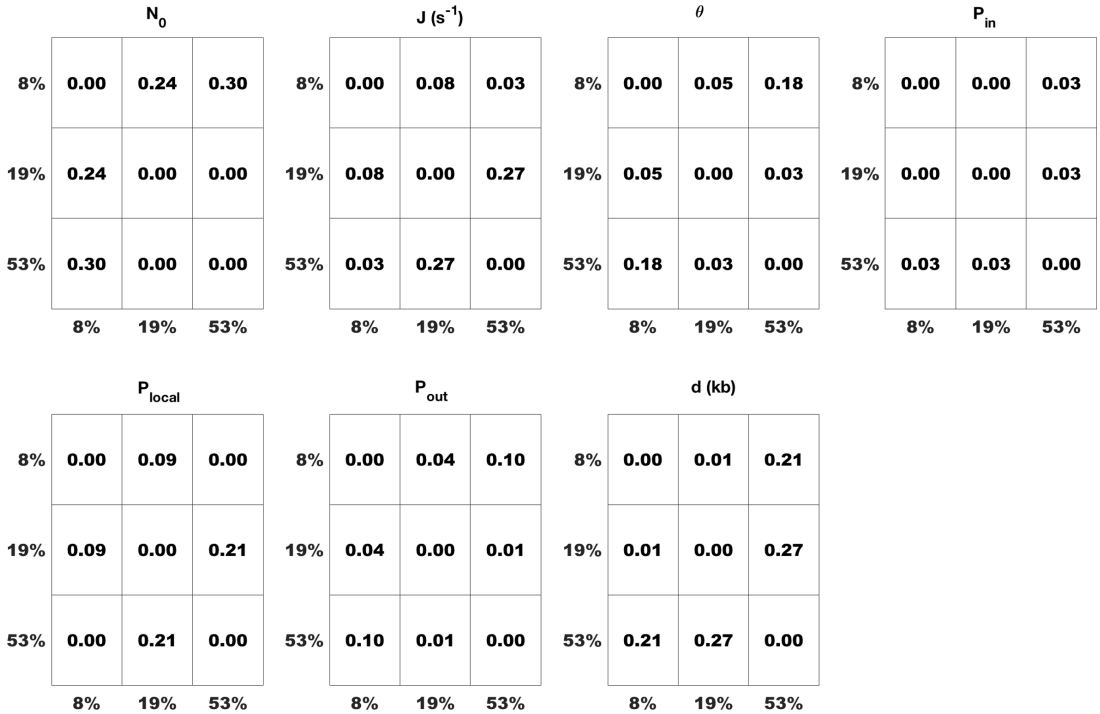

Fig 5. The values of each MM5 model parameter were compared pair-wise between samples with different global replicated fraction. The statistical significance of their difference was assessed by χ^2 test and represented as a binary heat map where not statistically significant differences are coloured in white and statistically significant difference are coloured in blue. The number in each box is the χ^2 coefficient.

In conclusion, any variation in parameter value detected by MM5 when analysing samples at different time points independently can be considered as statistically significant. Therefore, MM5 can adequately predict complex DNA replication dynamics using a limited number of processes.

References

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

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

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

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

5. 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;25(12):1886–1892. doi:10.1101/gr.195305.115.

6. Petryk N, Kahli M, d'Aubenton-Carafa Y, Jaszczyszyn Y, Shen Y, Silvain M, et al. Replication Landscape of the Human Genome. Nat Commun. 2016;7(1):10208. doi:10.1038/ncomms10208.

7. 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;27(8):1406–1416. doi:10.1101/gr.218602.116.

8. Platel M, Goldar A, Wiggins JM, Barbosa P, Libeau P, Priam P, et al. Tight Chk1 Levels Control Replication Cluster Activation in Xenopus. PLoS One. 2015;10(6). doi:10.1371/journal.pone.0129090.

Appendix 3

1 Fitting the experimental profiles by MM5 model : Unchallenged S phase

We fitted independently the measured profiles for each global replicated fraction by discrete MM5 model. The fits of observations from 8% global replicated fraction are presented in Appendix 1, Figure 9 and those of 19% and 53% are presented in Figures 1 and 2 respectively. In TABLE 1 we give the value of the fitted parameters. The reliability of observed differences among inferred MM5 parameters are assessed statistically by using χ^2 coefficient as defined in Appendix 2 (Figure 3).

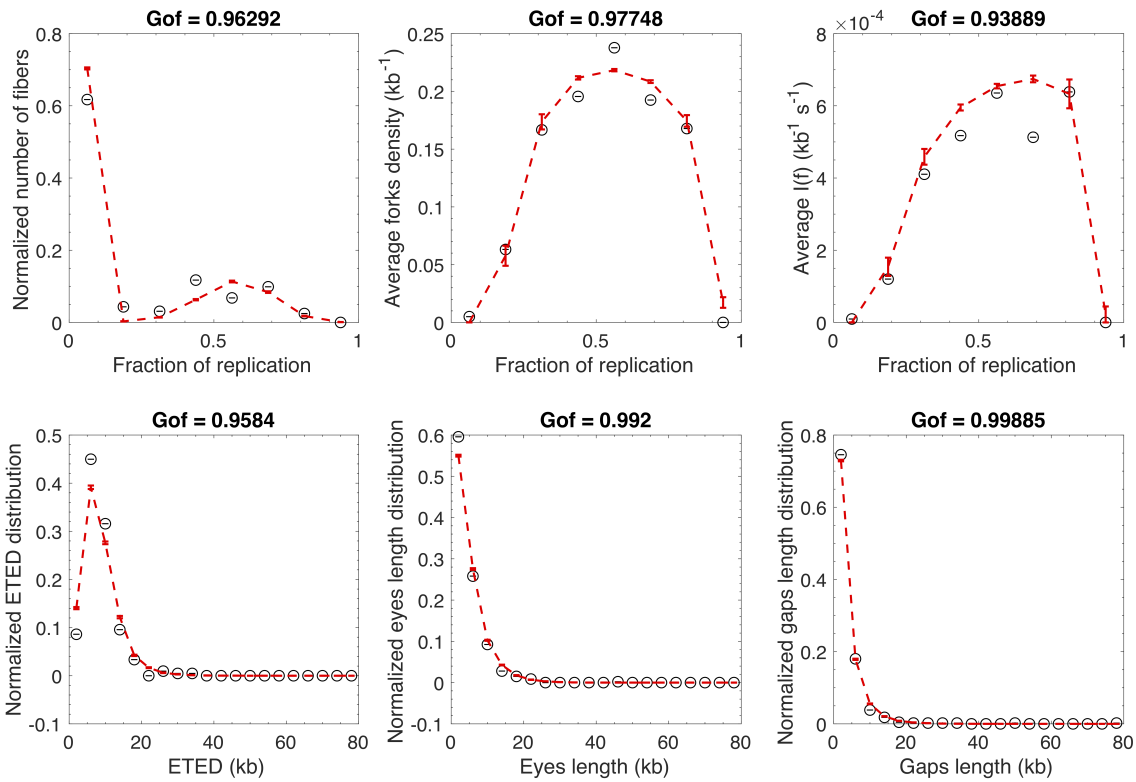

Fig 1. Modeling measured sample with 19% global replicated fraction with the discrete MM5 model. Open circles are simulated data and the red dashed line is the fit. $GoF_{global} = 0.96$

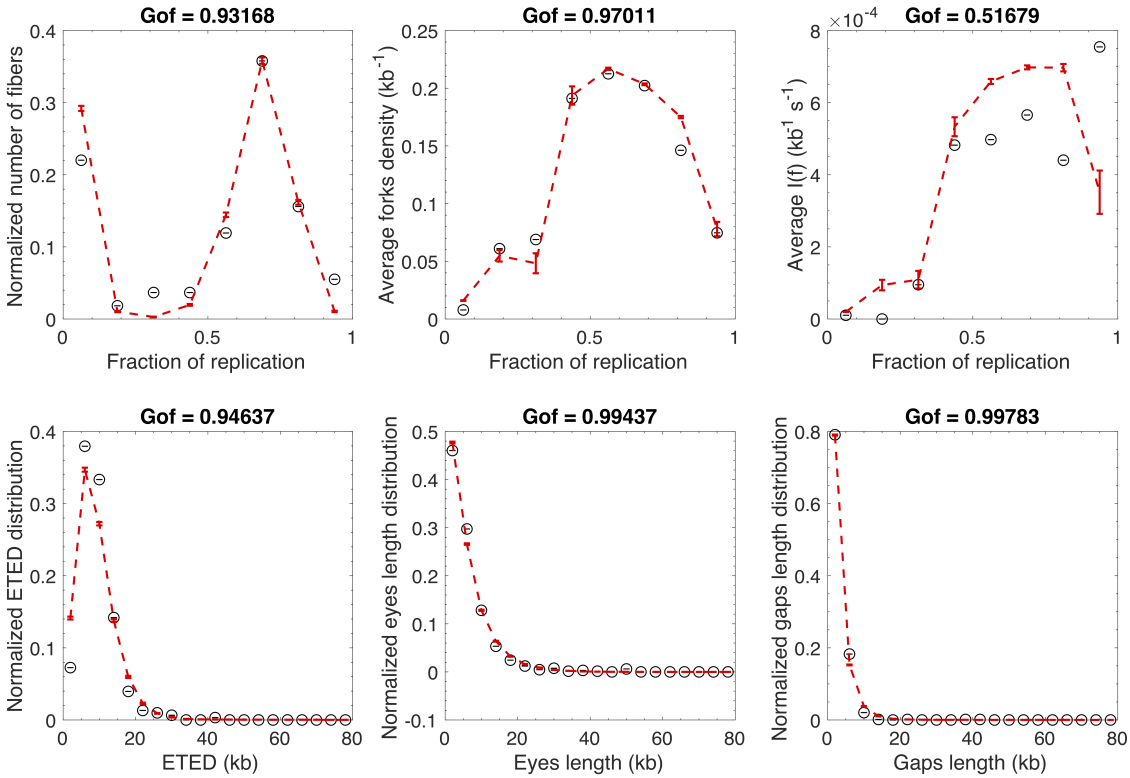

Fig 2. Modeling measured sample with 53% global replicated fraction with the discrete MM5 model. Open circles are simulated data and the red dashed line is the fit. $GoF_{global} = 0.90$

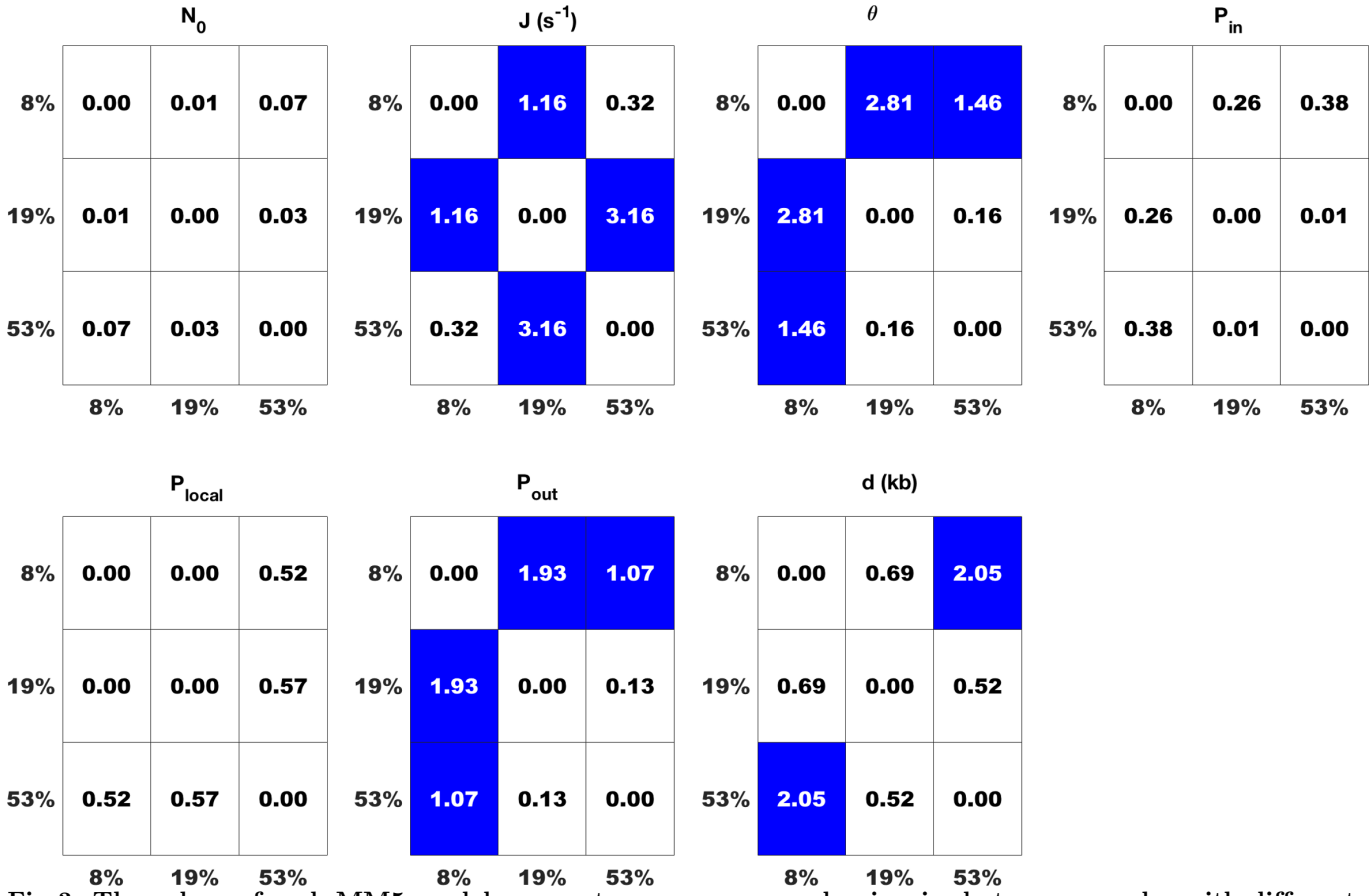

Fig 3. The values of each MM5 model parameter were compared pair-wise between samples with different global replicated fraction. The statistical significance of their difference was assessed by χ^2 test and represented as a binary heat map where the white colour represents no statistically significant difference and the blue colour represents statistically significant difference. The number in each box is the χ^2 coefficient.

2 Fitting the experimental profiles by MM5 model : Chk1 inhibited S phase

We fitted with the discrete MM5 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 Figure 4 and the obtained parameters values are given in TABLE 1. The reliability of observed differences among inferred MM5 parameters between controls and Chk1 inhibited sample are assessed statistically by using χ^2 coefficient as defined in Appendix 2 (Figure 5).

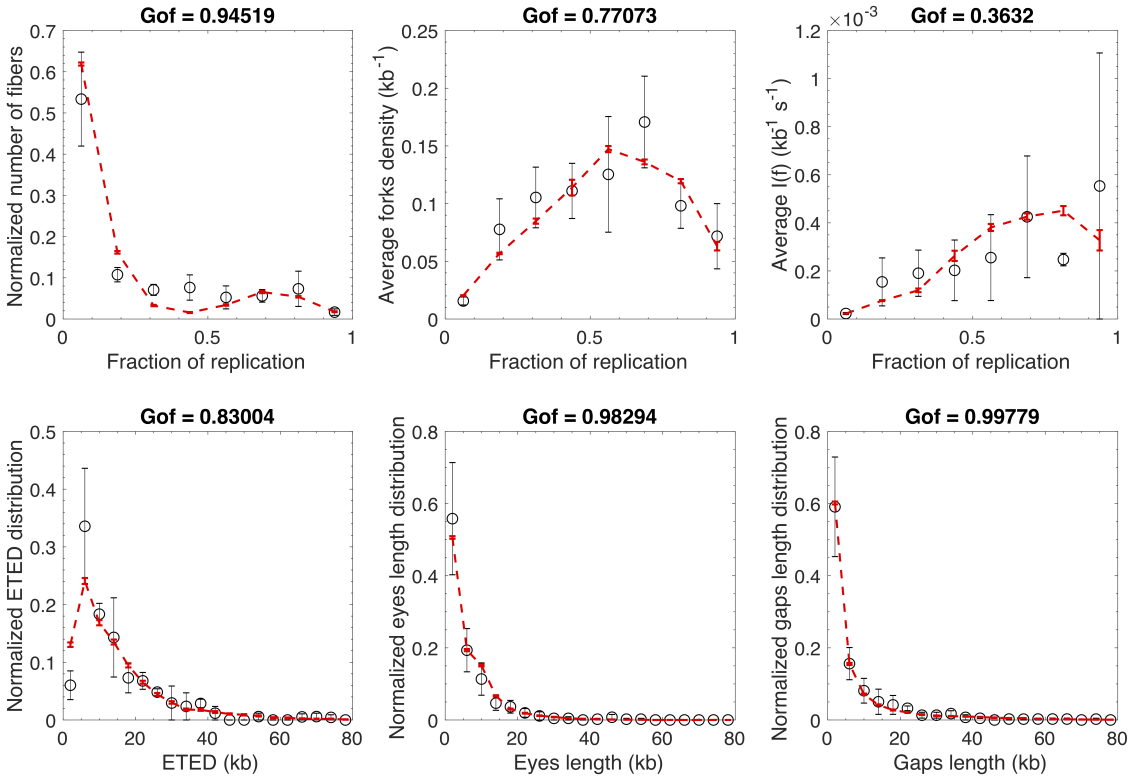

Fig 4. Modeling a measured sample with 22% global replicated fraction in presence of UCN-01with discrete MM5 model. Open circles are simulated data and the red dashed line is the fit. $GoF_{global} = 0.85$

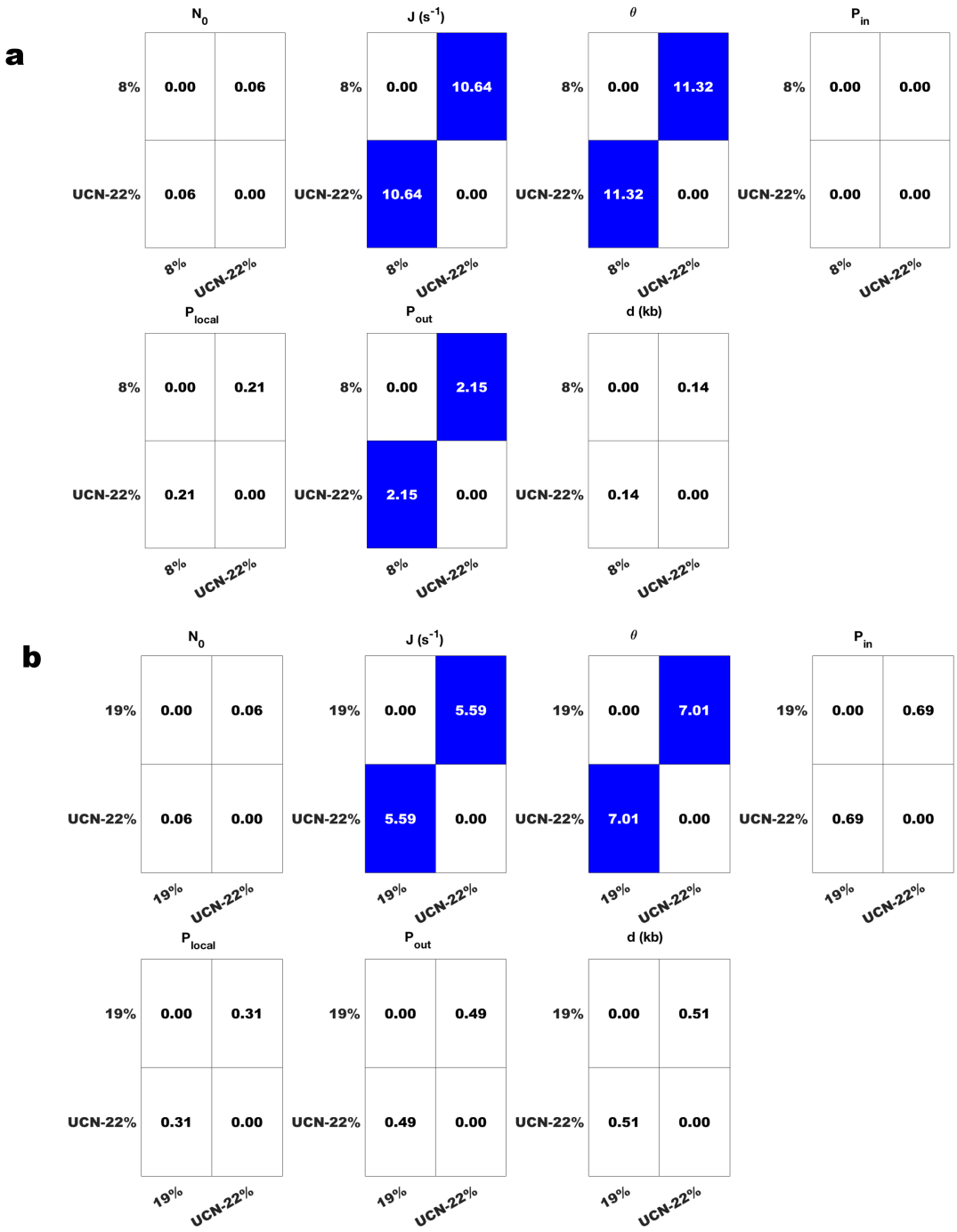

Fig 5. a. Comparing samples that have spent the same time interval in S phase. b. Comparing samples that have similar global replication fractions. The values of each MM5 model parameter were compared pair-wise between samples with different global replicated fraction. The statistical significance of 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.

3 Fitting the experimental profiles by MM5 model : Chk1 over-expressed S phase

We fitted with the discrete MM5 model a Chk1 over-expressed sample with the same incubation time than the control sample with 46% global replicated fraction. The global replicated fraction of the of the Chk1 over-expressed sample was 22%. The fits are presented in Figures 6 and 7 and the obtained parameters values are given in TABLE 1. The reliability of observed differences among inferred MM5 parameters between controls and Chk1 over-expressed sample are assessed statistically by using χ^2 coefficient as defined in Appendix 2 (Figure 8).

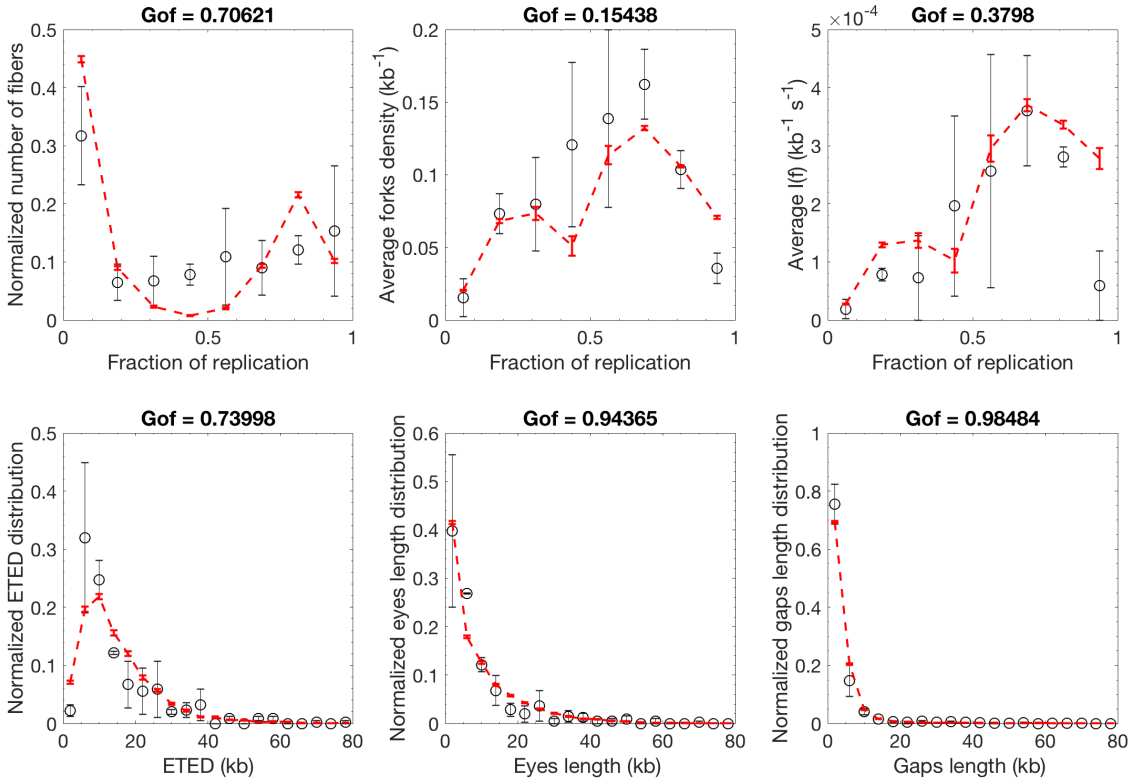

Fig 6. Modeling a measured sample with 46% global replicated fraction with discrete MM5 model. Open circles are simulated data and the red dashed line is the fit. $GoF_{global} = 0.65$

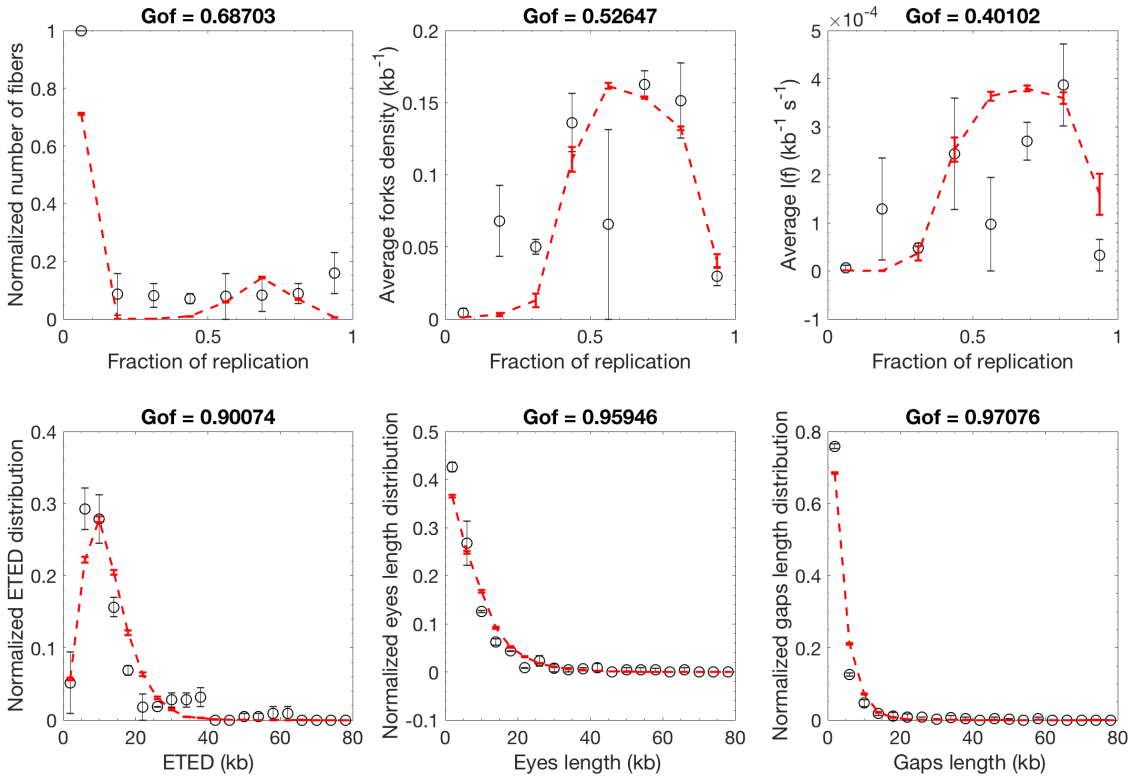

Fig 7. Modeling a measured sample with 22% global replicated fraction where Chk1 is over-expressed with discrete MM5 model. Open circles are simulated data and the red dashed line is the fit. $GoF_{global} = 0.74$

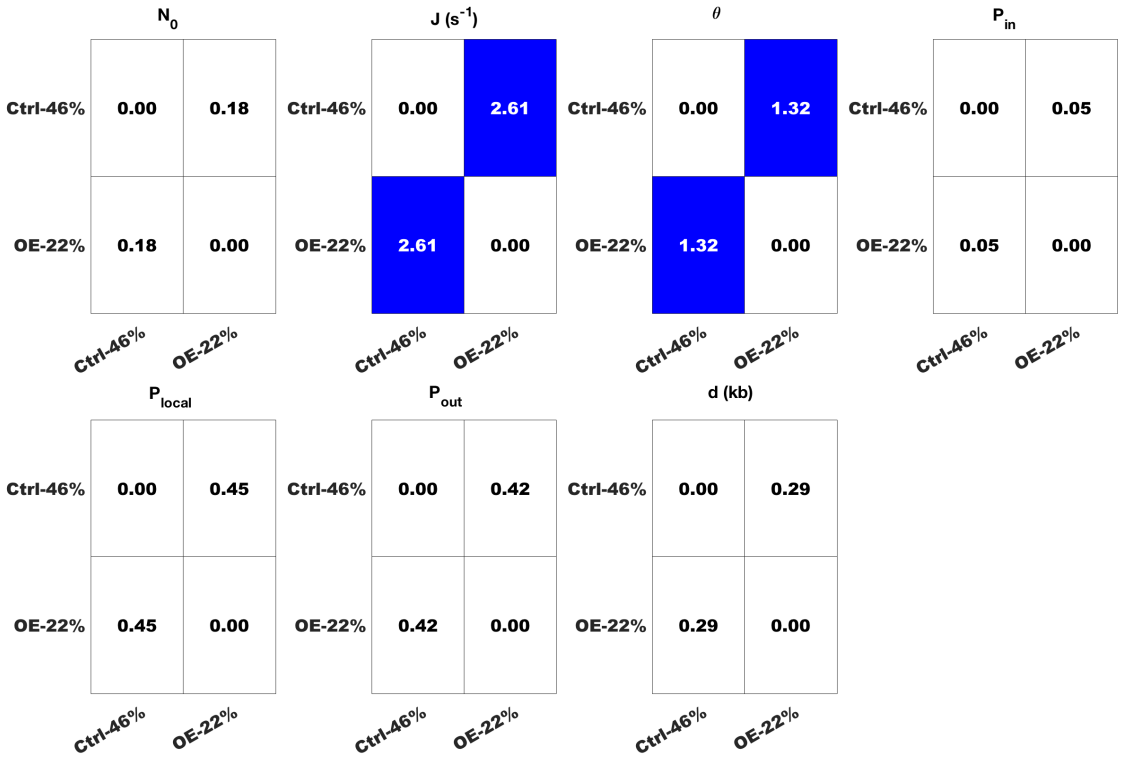

Fig 8. Comparing samples that have spent the same time interval in S phase. The values of each MM5 model parameter were compared pair-wise between samples with different global replicated fraction. The statistical significance of their difference was assessed by χ^2 test and represented as a binary heat 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.

Table 1. Values and the corresponding errors of MM5 parameters for the best fit of each sample and each condition.

MM5	unchallenged: 8%	unchallenged: 19%	unchallenged: 53%	UCN-01: 22%	unchallenged: 46%	over-expressed: 22%
N_0	1064 ± 135	1043 ± 116	1002 ± 106	1006 ±102	880±58.5	784±58.7
J (s^{-1})	601 ± 198	1026 ± 196	404 ± 151	1467 ± 89	1190±114	569±82.1
θ	0.25 ± 0.06	0.43 ± 0.04	0.39 ± 0.05	0.56 ± 0.032	0.50±0.01	0.42±0.02
P_{in}	0.41 ± 0.07	0.34 ± 0.07	0.32 ± 0.07	0.42 ± 0.07	0.33±0.03	0.36±0.04
P_{local}	0.43 ± 0.06	0.43 ± 0.06	0.52 ± 0.06	0.38 ± 0.06	0.37±0.03	0.45±0.03
P_{out}	0.09± 0.02	0.17 ± 0.04	0.15 ± 0.03	0.23 ± 0.04	0.15±0.02	0.23±0.03
d (kb)	143.8 ± 36.3	91.5 ± 25.6	56.1 ± 23.6	119.3 ± 29.3	139.9±16.9	178.4±18.7