

S-phase progression in mammalian cells: modelling the influence of nuclear organization

Alex Shaw · Pedro Olivares-Chauvet ·
Apolinar Maya-Mendoza · Dean A. Jackson

Published online: 13 February 2010
© Springer Science+Business Media B.V. 2010

Abstract The control of DNA replication is of fundamental importance as cell proliferation demands that identical copies of the genetic material are passed to the two daughter cells that form during mitosis. These genetic copies are generated in the preceding S phase, where the entire DNA complement of the mother cell must be copied exactly once. As part of this process, it is known that different regions of mammalian genomes are replicated at specific times of a temporally defined replication programme. The key feature of this programme is that active genes in euchromatin are replicated before inactive ones in heterochromatin. This separation of S phase into periods where different classes of chromatin are duplicated is important in maintaining changes in gene expression that define individual cell types. Recent attempts to understand the structure of the S-phase timing programme have focused on the use of genome-wide strategies that inevitably use DNA isolated from large cell populations for analysis. However, this approach provides a composite view of events that occur within a population without knowledge of the cell-to-cell variability across the

population. In this review, we attempt to combine information generated using genome-wide and single cell strategies in order to develop a coherent molecular understanding of S-phase progression. During this integration, we have explored how available information can be introduced into a modelling environment that best describes S-phase progression in mammalian cells.

Keywords DNA replication · S-phase programme · replication origins · DNA foci · replicon clusters · nuclear organization · modelling S phase

Abbreviations

BAC	Bacterial artificial chromosome
BrdU	5-bromo-2'-deoxyuridine
CDK	cyclin dependent kinase
CTCF	CCCTC-binding factor
FISH	Fluorescence in situ hybridization
ES	Embryonic stem cell

Introduction

The sheer complexity of the replication process is evident from the size of the genome in human cells—proliferating human cells have a diploid ($2n=46$) genome of roughly 6×10^9 bp DNA. Inevitably, this demands that synthesis proceeds from numerous points—about 40,000 are used in each cell cycle—

Responsible Editors: Marie-Noelle Prioleau and Dean Jackson

A. Shaw · P. Olivares-Chauvet · A. Maya-Mendoza ·
D. A. Jackson (✉)
Faculty of Life Sciences, University of Manchester,
MIB, 131 Princess Street,
Manchester M1 7DN, UK
e-mail: dean.jackson@manchester.ac.uk

that are scattered throughout the genome (Berezney et al. 2000; DePamphilis et al. 2006). Such synthetic initiation points, termed origins of DNA synthesis, are of fundamental importance in defining the efficacy of the replication process as they provide targets for binding of the replication machinery and facilitate replication licensing, which ensures that DNA is replicated once and only once during each cell division cycle (Blow and Dutta 2005).

Controlling the activation of DNA synthesis is a key decision point in the eukaryotic cell cycle and understanding how origins of DNA synthesis are first established on chromatin and then selected for activation is of fundamental importance in replication control (Mendez and Stillman 2003; Gilbert 2004). Across eukaryotes the synthetic machinery is highly conserved. However, with the evolution of organismal complexity and associated increases in genome size it is clear that higher eukaryotes face a substantial challenge in order to ensure that their genomes are replicated intact. With this in mind, it is not surprising that the replication of mammalian genomes takes many times longer than primitive eukaryotes, which provide our most tractable model systems; under optimal growth condition budding yeast replicate their genomes in ~1 h whereas human cells require ~10 h. This ~10-fold difference in the duration of S phase is not directly linked to the size of replication units, which on average only vary by ~3-fold. Indeed, the key difference is that in mammalian cells only about 10–15% of the genome is associated with replicons that are engaged in synthesis at any point during S phase. Like unicellular yeasts (Raghuraman et al. 2001), mammalian cells replicate specific regions of the genome at defined times of S phase (Goen and Cedar 2003; Aladjem 2007) so that active genes within open or dynamic euchromatin tend to be replicated early in S phase while more condensed heterochromatin replicates later.

It is reasonable to assume that a temporally structured S phase is likely to be of biological importance. Notably, cell differentiation in higher eukaryotes correlates with cell type specific patterns of replication timing (Hiratani et al. 2008), which in broad terms reflect changes in gene expression. This link between replication timing and gene expression may be of fundamental importance in maintaining patterns of expression as there is some evidence that histones with different post-translational modifica-

tions are deposited on DNA that is replicated during early or late S phase (Zhang et al. 2002; Lande-Diner et al. 2009). Moreover, the fact that early and late firing replication origins can be regulated by a molecular switch that involves the differential activation of potential origins based on their interaction with specific cyclin-CDK complexes (Donaldson 2005; Katsuno et al. 2009) implies that there is a biological imperative to maintain the timing programme.

The S-phase programme

S phase in mammalian cells is structured in time so that euchromatin is replicated before heterochromatin. At admittedly low resolution, this separation is clear from the discrete labelling of chromosomal bands visualized on metaphase chromosomes after prior labelling of DNA—using precursor analogues such as bromo-deoxyuridine (BrdU)—during different intervals of the preceding S phase (Holmquist 1987; Drouin et al. 1994). This type of cytological labelling suggests that the replication of R-band chromatin during early S phase is essentially completed before the synthesis in the chromosomal G-bands can begin. Moreover, the idea that R- and G-band synthesis occurs at discrete times of S phase is reinforced by the observation that in some cell types perturbation of the precursor pools reveals a distinct ‘3C-pause’ which appears to represent the time of switching from early to late synthesis (Drouin et al. 1994; Strehl et al. 1997). This broad-scale timing of chromatin domains of the size of chromosomal bands (i.e. ~10 Mbp) was confirmed in the seminal experiments of Carter et al. (Woodfine et al. 2004), who used hybridization of DNA isolated from G1 and S phase human lymphoblast cells to map the timing of replication based on DNA content. This early study provided a low resolution (~1 Mbp) genome-wide map of the human timing programme.

Recent advances in microarray design and analysis (TimEX) and deep sequencing (TimEX-seq) approaches have confirmed the basic conclusions of this seminal study (Desprat et al. 2009). These high resolution studies provide precise estimates of copy number variation based on the use of Gaussian convolution (noise filtering) to integrate massive numbers of highly redundant measurements. The

following points summarise the key findings of this high resolution analysis:

1. Replication proceeds with a clear temporal programme, with regions of the genome of ~1 Mbp being assigned to replication timing domains with a time resolution of 1-2 h. Importantly, replication domains of this size are unlikely to represent single replicons as forks of ~500 kbp would take at least 5 h to complete synthesis.
2. Replication domains in early and mid/late S phase are distinct and in both cases synthesis initiates within large zones that contain a high density of potential initiation sites. Many long replicons (>250 kbp) link the early and mid/late replicating regions. These 'temporal transition regions' couple the early and mid/late replicating domains and represent ~5% of the genome where replication origins are highly dispersed.
3. Replication timing of a gene locus correlates with the level of gene activity within the locus. Regions that are replicated very early in S phase tend to contain genes with very high levels of transcription whereas regions with genes that have only low levels of transcription tend to be replicated during mid-S phase. Regions of the genome that are expressed late in S phase are remote from transcribed genes.

These general conclusions have been confirmed and extended in many recent studies that have explored both the general timing programme (Farkash-Amar et al. 2008; Hiratani et al. 2008; Desprat et al. 2009) and specific locations where synthesis can begin (Cadoret et al. 2008; Sequeira-Mendes et al. 2009). For a much better appreciation of this information the reader is referred to excellent articles elsewhere in this volume (Farkash-Amar and Simon 2010; Pope et al. 2010; Cadoret and Prioleau 2010).

While the segmentation of mammalian genomes into early and late replicating domains appears robust the time resolution of these experiments is often poor and in some cases large regions of the genome appear to engage synthesis over many hours (Jeon et al. 2005; Karnani et al. 2007). The mechanisms of origin selection clearly influence the efficiency with which different regions of the genome are replicated at precise times of S phase. In this respect, it is important to recognise that the activation of synthesis at individual potential origins is stochastic. Potential

origins will each have different probabilities of firing during each cell cycle, presumably as a consequence of the local chromatin environment. As a result, in individual cells, the majority of potential origins are not used during a particular S phase and most potential origins are replicated passively by forks that emanate from adjacent replication units. Hence, individual potential origins will only provide initiation sites for synthesis in a minority of cells and any specific locus will be replicated at a time that relates to its position relative to the nearest active origin (Hamlin et al. 2010). The analysis of replication intermediates using both 2D-gels (Mesner et al. 2006) and DNA fibres (Lebofsky et al. 2006; Desprat et al. 2009) confirms this stochastic view of origin firing.

Analysis of DNA replication using single cell approaches

Genome-wide approaches undoubtedly provide valuable insight into the distribution of replication origins and the replication timing programme (Farkash-Amar and Simon 2010; Cadoret and Prioleau 2010 in this volume). However, the variable efficiency of origin activation raises obvious questions about mechanisms of origin firing that cannot be addressed using genome-wide strategies, which use large numbers of cells and so are unable to detect any cell-to-cell variability. Hence, it is of clear value to interpret the genome-wide data in the context of complementary studies performed on individual cells.

It has been recognised since the seminal studies of Nakamura et al. (1986) that replication in mammalian cells takes place within specialized nuclear domains where many active replisomes are clustered together. Many subsequent studies (see Jackson 1995 and Zink 2006 for reviews) have used a wide range of modified replication precursor analogues (BrdU, biotin-dUTP and Cy3-dUTP are frequently used examples) to confirm that mammalian cells perform replication at discrete replication sites, which contain groups of polymerase complexes within synthetic factories (Hozak et al. 1994; Leonhardt et al. 2000). The replication machinery within individual factories performs synthesis of small groups of contiguous active replicons, which are replicated together and activated at similar times (Jackson and Pombo 1998; Ma et al. 1998). These replicon clusters can be

visualized as ‘DNA foci’ that contain ~1 Mbp of DNA (Cremer and Cremer 2001). Importantly, these functional targets for DNA synthesis have been shown to represent stable structural units of sub-chromosomal organization. As S phase proceeds, the structure of active centres of DNA synthesis changes according to a predictable programme (Fig. 1), which reflects the disposition of different chromatin classes within the nucleus (Cremer and Cremer 2001; Goetze et al. 2007).

Mapping S-phase progression at the level of DNA foci

A highly structured replication programme (Fig. 1) implies that specific regions of the genome are selected for synthesis at predictable times. This could of course reflect the stochastic activation of different classes of potential replication origins, based perhaps on their interaction within different cyclin/CDK complexes (Katsuno et al. 2009). However, studies in single cells have suggested that the organization of DNA foci contributes to S-phase progression. Analysis of the time of replication of DNA foci in different cell cycles has shown that the same DNA foci are activated with high efficiency (>90%) at the onset of S phase (Jackson and Pombo 1998; Ma et al. 1998). This implies that a robust mechanism regulates the selection of replicon clusters that are targets for synthesis as cells enter S phase. In addition, as S phase proceeds any newly activated replication sites appear to lie next to sites that were engaged in synthesis during the previous period of S phase (Manders et al. 1992). This suggests that the spatial architecture of chromatin foci might be a key determinant of S-phase progression with the sequential activation of foci occurring following a ‘next-in-line’ principle (Manders et al. 1992). This has been confirmed using an analysis of mid/late replication factories in living cells (Sporbert et al. 2002), where analysis of the simplified patterns of active sites allowed the spatial relationship of foci to be mapped at high resolution.

Are structure-function links defined by DNA foci?

If the structure of DNA foci plays a significant role in defining the architecture of the replication programme it is important to understand how individual foci are

defined. In fact, very little is known about the structure of foci and the molecular principles that might allow stable structures to be established and maintained. There is evidence that chromatin foci are maintained by epigenetic chromatin states. For example, the analysis of sub-chromosomal regions with interspersed gene islands (gene-rich regions) and gene deserts (regions with very few active genes) shows that the two chromatin classes are separated into discrete foci with chromatin that does not mix (Fig. 2; Shopland et al. 2006; Goetze et al. 2007). If such specific examples define a general feature of chromatin organization, it is not unreasonable to suggest that chromatin status might dictate the replication timing of foci with different chromatin epi-states. Supporting this model, replication timing in yeast and notably the transition from euchromatin to heterochromatin replication is defined by the acetylation status of histone in the two chromatin compartments (Vogelauer et al. 2002).

While mechanisms that link the structure of DNA foci and their replication timing are a matter for speculation, our understanding of potential links is clearly hindered by deficiencies in our knowledge of foci structure. It is known, for example, that global chromatin loops in mammalian cells correlate with replicon size (Buongiorno-Nardelli et al. 1982; Courbet et al. 2008), perhaps to provide a memory of replicon structure that is transmitted for one cell generation to the next. But how such loops relate to function and the structure of the template within DNA foci is unclear. Historically, numerous studies have described the behaviour of genomic elements such as nuclear scaffold and matrix attachment regions, locus control regions and domains insulators (reviewed in West and Fraser 2005) that together define the architecture of chromatin domains in mammalian cells. More recently, the insulator protein CTCF has emerged as a good candidate to define boundary elements that punctuate the genome to form higher-order chromatin domains (Phillips and Corces 2009). Intriguingly, sites of CTCF binding have also been shown to be sites of cohesin accumulation, suggesting that they might assume special structural properties that contribute to architecture of chromatin loops (Parelho et al. 2008; Hadjur et al. 2009). In addition, hotspots of CTCF binding have been shown to establish unique features in the local chromatin environment (Fu et al. 2008; Zhang et al. 2008), which might contribute to the formation of

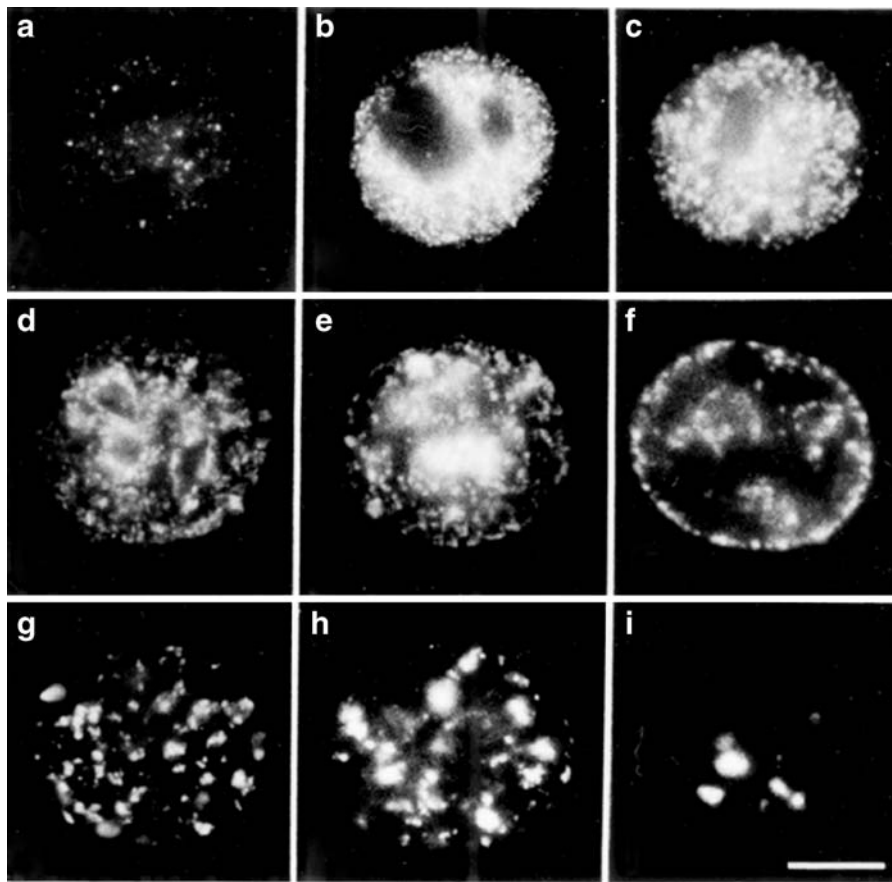


Fig. 1 The spatial distribution of active replication sites during S phase. During S phase, different classes of chromatin are replicated at different times. Chromatin that contains the majority of transcribed genes, within chromosomal R bands, is replicated over the first ~4 h of S phase. During this period, active sites of DNA synthesis are in discrete foci dispersed throughout the nuclear interior (**a–c**). At mid-S phase, replication begins to switch to more inert chromatin and patterns of replication foci that reflect the peripheral location of hetero-

entropy-driven higher-order chromatin conformations (St-Jean et al. 2008).

Single molecule analysis on DNA Fibres

While the analysis of replication foci in situ provides some molecular insight to support the genome-wide studies, the analysis of DNA foci within nuclei is also limited in scope by resolution. The low resolution information within foci in situ can, however, reveal additional high resolution information if the DNA that they contain is analyzed after preparation of spread DNA fibres. Labelled DNA fibres can then be used to

chromatin are seen (**d–f**). Finally, heterochromatic blocks of late replicating chromatin are duplicated within the nuclear interior (**g–i**). Images shown are replication sites labelled in permeabilized HeLa cells—using biotin-dUTP—that were fixed and indirectly immuno-labelled under conditions that preserve nuclear organization. The bar is 5 microns. For more details see Hozak et al (1994). Reproduced with permission from the Company of Biologists

develop detailed information about fork rates and the distribution of active replicons and how individual replicons are activated in different cell cycles (Jackson and Pombo 1998; Takebayashi et al. 2001). Most importantly, DNA fibres prepared from cells that were labelled with different replication precursor analogues during consecutive cell cycles provided compelling evidence that structurally stable replicon clusters generate DNA foci that represent both structural and functional sub-chromosomal units (reviewed in Maya-Mendoza et al. 2009).

In recent years, the analysis of DNA fibres prepared from cells that have been labelled with a range of replication precursors has revealed funda-

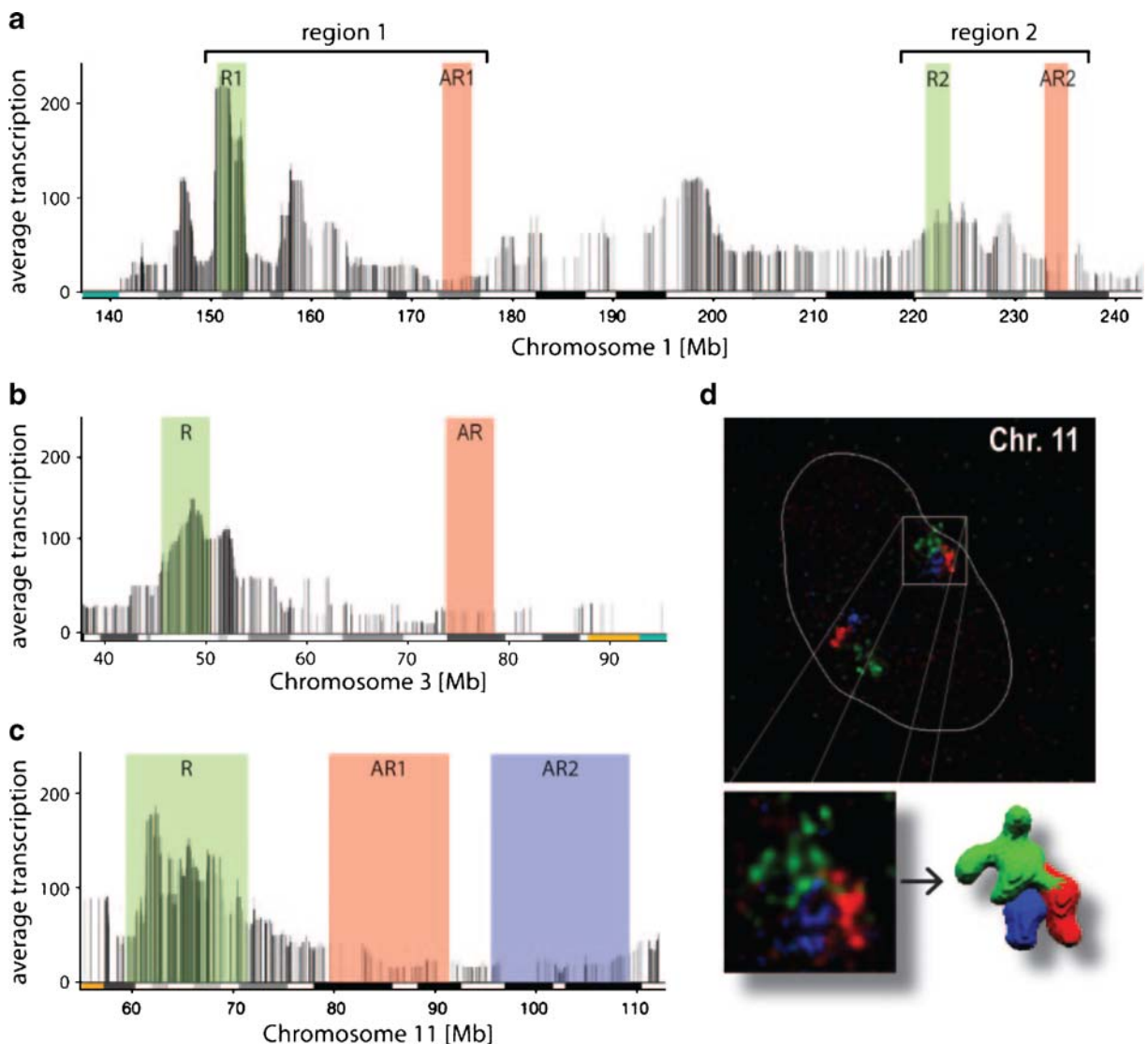


Fig. 2 DNA foci are structural units of higher-order chromatin folding. Mammalian genomes are folded into chromatin domains that assume a variety of chromatin environments as a result of local patterns of gene expression. Genomic regions that are rich in active genes—known as gene ridges (*R*) or gene islands—are separated by gene-poor domains—known as anti-ridges (*AR*) or gene deserts. The local architecture of three chromosomal loci with interspersed ridges and anti-ridges are shown above (**a–c**). To test the nuclear distribution of the different chromatin compartments, the three regions (one ridge and two anti-ridges) highlighted in (**c**) were visualized in situ using FISH (fluorescent in situ hybridization) to probe the target loci (**d**). The FISH probes for the three target loci were prepared from contiguous arrays of BAC clones, which spanned the regions shown. The three pools of BAC clones

were differentially labelled prior to hybridization. Visualization of the labelled probes shows that three regions under analysis are constrained within discrete local domains. Notably, while the BAC pools cover ~10 Mbp of DNA for each region, in all cases the fluorescent signal was concentrated locally in domains of ~500 nm. Based on their number, each of these domains contains roughly 1 Mbp of DNA. The gene-rich and gene-poor compartments are self-contained (i.e. discrete) and the chromatin environment within the compartments defines the volume occupied (gene-rich compartments are more open) and the spatial architecture of the domains within the each chromosome territory. Images taken from Goetze et al. (2007) and published with permission of the American Society of Microbiology

mental information about the structure of eukaryotic replicons and the replication programme (reviewed by Tuduri et al. 2010). Studies evaluating the activation of potential origins across specific chromosomal regions have been especially informative (Lebofsky et al. 2006; Conti et al. 2007). A key feature of these studies has been the recognition that potential sites of initiation of DNA synthesis are typically distributed throughout ~10 kbp chromatin domains. However, pre-initiation complexes that are selected from these zones to activate synthesis are recognised inefficiently, so that in individual cells only about 1/3rd support initiation during a particular cell cycle. Activated origins appear to be selected at random so that different combinations of active origins are seen in different cells (Lebofsky et al. 2006). Moreover, adjacent active origins, which are typically separated by roughly 100 kbp of DNA in mammalian cells, are often seen to be activated at similar times and in most cases synthesis proceeds with forks that grow at very similar rates (Lebofsky et al. 2006; Conti et al. 2007).

Simulating S-phase progression in mammalian cells

As so little is understood about the molecular principles that regulate S-phase progression we wanted to assess if *in silico* simulations could be developed to model features of nuclear organization that contribute to the chromatin environment and drive the S phase programme. To do this, we have attempted to incorporate information described above that is derived from both genome-wide and single cell studies. In considering possible mechanisms, one might begin by suggesting two extreme scenarios. In the first, the activation of potential origins might be fundamentally stochastic, so that initiation is driven by random choice with the proviso that the chromatin environment modulates choice so that different regions of the genome will be replicated preferentially at different times. Euchromatin is known to engage synthesis before heterochromatin and it is possible to argue that subtle differences in chromatin structure might contribute to replication timing within these chromatin compartments. In the second, the chromatin environment defines the sites that are selected for initiation of synthesis at the onset of S phase but thereafter replication spreads from these primary

initiation sites so that the downstream replication programme is defined by the activation of genetically adjacent chromatin domains along chromosomes. This scenario represents origin activation driven by a next-in-line model of S-phase progression. Of course, as these extremes are not mutually exclusive the molecular mechanism of progression *in vivo* might involve a mixture of stochastic and genetically coupled activation events.

Modelling the chromatin environment

Published models to describe eukaryotic DNA replication have focussed predominantly on stochastic models of origin activation. Most attention has focussed on organisms with simple replication programmes (Lygeros et al. 2008; Herrick et al. 2002; Rhind 2006) and only recently have the models been used to explore aspects of replication in the S phase of somatic mammalian cells (Goldar et al. 2009; Ge and Blow 2009). A comprehensive analysis of these published models is presented elsewhere in this volume (Rhind et al. 2010; Hyrien and Goldar 2010).

In mammalian cells, local chromatin environments play a major role in S-phase progression. Hence, any viable model of S phase must incorporate parameters related to the orderly synthesis of the major chromatin compartments and evaluate established features of organization related to the mechanisms involved. In particular, any model of the mammalian S phase must incorporate replicon clusters (within DNA foci) as the basic targets for DNA synthesis and evaluate how replication spreads between these structures. Here, in order to simulate the activation of replicon clusters, we have taken data for the distribution of replicons within replicon clusters from Jackson and Pombo (1998); primary data sets were used to model the profile of inter-origin separations within clusters and used in combination with the published distribution of active replicons/cluster. The distribution of replicons within replicon clusters that are replicated at different times of S phase have a similar average structure (Maya-Mendoza et al. 2007), despite differences in their spatial organization and nuclear distribution (Shopland et al. 2006; Goetze et al. 2007; see Fig. 2).

Hence, in considering the different features that define the chromatin environment we propose that a biologically informative simulation of the mammalian

S phase should incorporate the following conditions during modelling:

1. DNA in chromosomal R- and G-bands is replicated preferentially at defined times of S phase, with synthesis of R-band chromatin in early S phase and G-band chromatin in mid/late S phase. The differential probability of origin activation will be determined by expression of appropriate cyclin-CDK complexes (Katsuno et al. 2009).
2. Throughout S phase, replicons are activated in small groups within functional replicon clusters (Jackson and Pombo 1998).
3. Clusters that are active during consecutive intervals of S phase are defined predominantly by chromosome structure.
4. Mammalian S phase is regulated by a mechanism that restricts the absolute level of synthesis, so that only 10–15% of the genome is engaged in synthesis at any time. The mechanism that drives this ‘replication rheostat’ is unknown.
5. Replication of different chromatin classes occurs at different rates (Takebayashi et al. 2001).

A basic modelling framework

Using these experimentally defined conditions, model development has obvious potential to inform our understanding of mechanisms that drive the spread of replication throughout mammalian genomes. In the analysis that follows, models were implemented in Matlab and tested by fitting to the natural duration of S phase in order to assess biological efficacy; mammalian S phase takes ~10 h to complete and for the purpose of simulation we have restricted initiation to the first 8 h of this period. Using human chromosome 6 to build our model (Goldar et al. 2009), we first estimated the number of replicon clusters required to complete S phase (Fig. 3) using cluster architectures from Jackson and Pombo (1998) and variable fork rates from Takebayashi et al. (2001). In this simple form of simulation, all clusters have an equal probability of being activated. Hence, the simulation defines the absolute number of activation events required to complete synthesis and estimates the probability of cluster firing during defined intervals in order to perform synthesis in the desired time. As expected, as potential origins are consumed

the probability that remaining origins will be activated increases so that replication completes on schedule (Fig. 3). The pattern of activation seen in this profile reflects the structure of replicon clusters—the lengths of replicons within clusters dictates the timing when their synthesis can complete and this is linked to activation of new clusters. The decline in initiation events towards the middle of S phase is a consequence of the reduced rate of fork elongation at that time.

The profile of cluster activation seen in Fig. 3 can also be represented to show the absolute levels of DNA synthesis as S phase proceeds. This readout is used to map the success of S phase in the simulations shown in Fig. 4. In these simulations we incorporate key features of the chromatin environment into the model using R- and G-band coordinates taken from the UCSC Table Browser with the March 2006 genome assembly (Karolchik et al. 2003). These chromosome banding patterns were applied to the simulation and the probabilities of activation within euchromatin and heterochromatin adjusted to mimic the effect of chromatin environment on the activation of potential origins. We also used expression data from Katsuno et al (2009) to simulate the effect on differential activation of R- and G-band replicon clusters given that the G-band clusters are activated by increasing Cyclin A-CDK1 expression towards late S phase—the availability of Cyclin A-CDK1 was modelled to rise starting at 2 h after the onset of S phase and reach a peak 4 h later (Fig. 4b). During this compound simulation, R band replication was activated at the onset of S phase and proceeded as before until the increasing expression of Cyclin A-CDK1 allowed origin activation within G-band clusters.

Simulations were developed to identify optimal probabilities of origin activation as defined by the amount of deviation from the average synthetic quota (defined by the replication rheostat) required to complete synthesis within 10 hours. The simulation incorporates variable activation probabilities of G-band clusters as Cyclin A-CDK1 expression increases between 2–6 h of S phase. With a sigmoidal expression profile, the minimum variation from the DNA quota per minute was found to occur with a maximum G band cluster firing efficiency of 0.0045/cluster/min (Fig. 4b). A slightly higher maximal probability of 0.0054/cluster/min was seen when the increase in expression was linear.

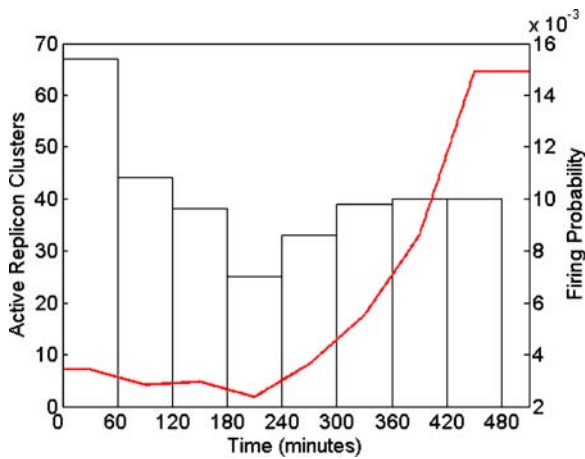


Fig. 3 Calculation of replicon cluster firing efficiencies. This simulation describes the architecture of replicon clusters and the probability of cluster firing during replication of human chromosome 6. This chromosome contains 171 Mbp of DNA, so simple calculations allow us to determine the fraction of the chromosome that must be replicated within each hour window of S phase when initiation can occur for 8 h. Using cluster architectures from Jackson and Pombo (1998), this simulation calculates the number of replicon clusters that must be activated to ensure the required amount of DNA synthesis within each hour of S phase (*bars*). We used the published cluster architectures and replicon lengths, which approximated to a normal distribution ($\mu=140.6238$ kbp, $\sigma=58.8192$). Cluster architectures were generated independently for each simulation by random sampling of the experimentally derived data sets (average values of 5,000 independent simulations are shown). Replication of individual clusters was programmed to proceed at constant rate and variable fork rates across S phase were smoothed to prevent discontinuities in the simulation. Then, as the distribution of replicon clusters defines the number of active clusters that will be needed to complete synthesis in the required time the necessary firing probabilities can be calculated (*red line*). In this example, the average profile of cluster architectures requires that 326 foci are activated to replicate the 171 Mbp chromosome (most foci contain 250–1,000 kbp of DNA). It is assumed that all unreplicated clusters have an equal probability of activation. Hence, for each time point of S phase the simulation uses the absolute number of synthetic units to estimate the probability/cluster/minute that is required to complete synthesis on schedule. At the onset of S phase, 67 clusters are activated to engage the required level of synthesis, with a probability of $67/326 \times 60 = 0.034$ clusters/min. During the 7th hour of the simulation 40 of the remaining 80 clusters are activated so the probability of activation increases to $40/80 \times 60 = 0.083$ clusters/min.

The firing efficiency profiles generated using the optimal conditions (with adjustments after 6 h to compensate for the dwindling pool size) are shown in Fig. 4c. These outputs include the effects of cluster banding on the existing model framework and approach a realistic biological representation of

chromosome structure in vivo. As shown in Fig. 4c, the different patterns of increasing cyclin expression had only a slight effect on S-phase progression; the sigmoidal profile was used in later models.

Spatial architectures of replication foci

So far our analysis has simulated the effects of replicon clustering within DNA foci, variable fork rates throughout S phase and the differential activation of potential origins during early and mid/late S phase based on their chromatin environment. To add molecular complexity to the simulations, we next evaluated how models might be affected by different mechanisms of S-phase progression (see Fig. 5). This aspect of the modelling is designed to assess how next-in-line and stochastic models of cluster activation influence S-phase progression. Simulations were performed using the conditions developed in Fig. 4 to test which parameters give the best fit to the established S phase duration (Fig. 6). In this analysis, different modelling environments were compared using an end-time where 95% of DNA was replicated; this limits the effect of rare events that can lead to very long end-times. To simulate the effect of a next-in-line mechanism of origin activation different multiplier values (between 1 and 5000) were incorporated into the model. This feature alters the probability with which replicon clusters are selected for activation based on changes in the chromatin environment that arise during replication of neighbouring clusters. A low resolution scan of the parameter space, comparing a range of maximum firing efficiencies for the sigmoidal curve (between 0.0001 and 0.0083/cluster/min), highlights a number of regions of biological interest (Fig. 6). In this phase plot, each of these areas of interest indicates the impact of different parameter sets and thus different mechanisms that are driving the progress of S phase (Fig. 7).

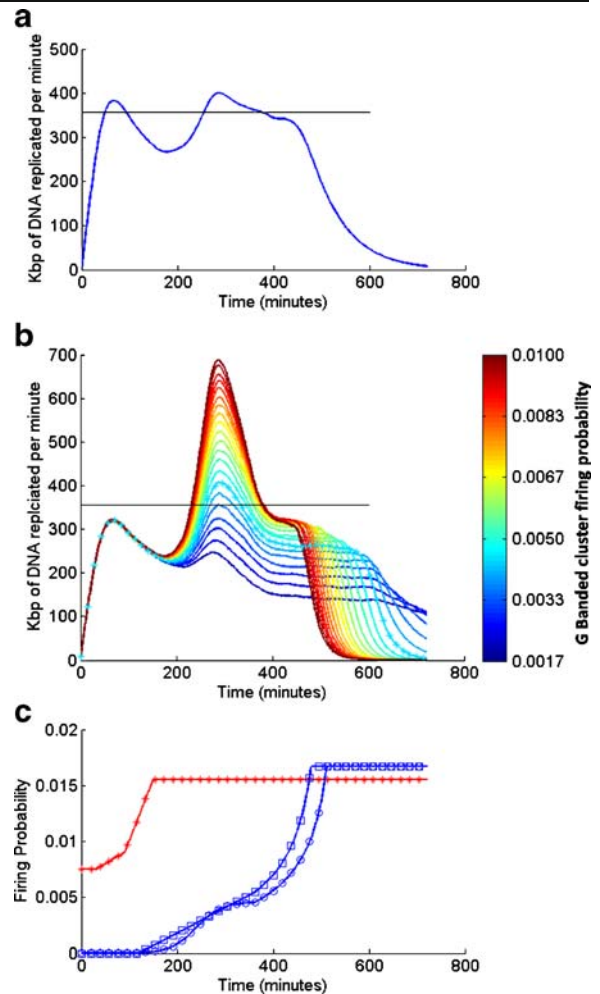
The following conclusions were drawn from simulations that test three alternative models of S-phase progression:

1. Origin selection is stochastic

A null hypothesis that ignores any relationship between DNA foci would simply alter the probabilities of origin activation towards late S phase, based

Fig. 4 Modelling DNA replication across S phase. The distribution of cluster firing probabilities generated in Fig. 3 was used to simulate an averaged profile of DNA replication throughout S phase (a). A linear representation of human chromosome 6 was created and divided into replicon clusters using distribution data from Jackson and Pombo (1998) and distribution of firing probabilities applied. Each cluster has the potential to be activated during time steps of 1 min. Once activated, DNA within each replicon of a cluster replicates according to the specific fork speed relevant to the particular period of S phase and terminates on meeting a neighbouring fork. The blue line shows the progress of synthesis (DNA replicated in kbp/min averaged over 5,000 simulations) and black line the quota of DNA synthesis required to complete S phase on schedule. A modified version of the simulation shown in (a) was generated to accommodate the effect of different chromatin environments in chromosomal R- and G-bands (b). Using the R/G band configuration across human chromosome 6, probabilities of cluster activation were calculated first in R-bands and then in G-bands using a range of different potential maximum values as Cyclin A-CDK1 concentrations increased. Data shown were generated by modelling a sigmoidal increase in Cyclin A-CDK1 expression between 2–6 h of S phase. A linear increase in expression was also tested (not shown). The range of maximum firing probabilities is shown in the accompanying colour-bar. Once the maximum value is reached at 6 h, the probability is adjusted to account for the decreasing pool of unreplicated clusters. Each coloured plot of DNA output therefore refers to the DNA replicated (kbp/min) under different maximum G Band firing probabilities. Output is averaged over 1,000 simulations for each parameter and firing probabilities are measured per cluster per minute. The optimal probability of 0.0045, giving the closest adherence to the DNA quota, is highlighted (cyan stars). Rates of synthesis within different clusters throughout S phase can be transformed to monitor overall levels of synthesis as S phase proceeds (c). Firing probabilities were generated as before with R-bands firing (red line) during early S phase followed by the optimized firing of G-band clusters (blue/cyan lines). The optimal G-band cluster firing probabilities were the maximum values giving the closest fit to the DNA replication quota. With linear increase in Cyclin-CDK expression (blue line with squares), the probability of cluster firing within G-bands peaked with optimal probability of 0.0054/cluster/min at 6 h. With sigmoidal increase (blue line with circles), an optimal probability of 0.0045/cluster/min at 6 h was seen

on expression of activating cyclin-CDK complexes (Fig. 4). In this case, the maximum probability of G-band firing defines the behaviour of the model. A maximum probability of 0.004/cluster/min was therefore tested as a case study (Fig. 6, position a). This parameter set gives an average variation from quota of 74.14 kbp/min and completes 95% of DNA replication within 8.4 h, with absolute completion by 10.8 h. The standard deviation at absolute completion was 67.0 min. This mechanism therefore provides a stable and timely completion of S phase. However, this model does generate a high level of ab initio cluster



activation of 41.5%. Additionally, whilst the ratio of single sided firing events to dual sided is 2.2, this is a consequence of the high levels of ab initio firing. Importantly, the distribution of origin firing in this case is skewed very late into S phase and predicts a level of very late synthesis that is not seen experimentally. Predictably, increasing the maximum probability of activation results in a shift in activation but also leads synthesis to complete at unrealistically early times. Allowing a small effect of fork elongation on cluster firing probabilities (with a maximum probability of 0.0033/cluster/min and a $\times 2$ increase in cluster firing if forks are encroaching—Fig. 6, position b) reduces the length of S phase even though the maximum probability of activation is reduced. The variation of completion times is seen to rise slightly however, showing that this limited spatial effect has little beneficial consequence on the behaviour of the system.

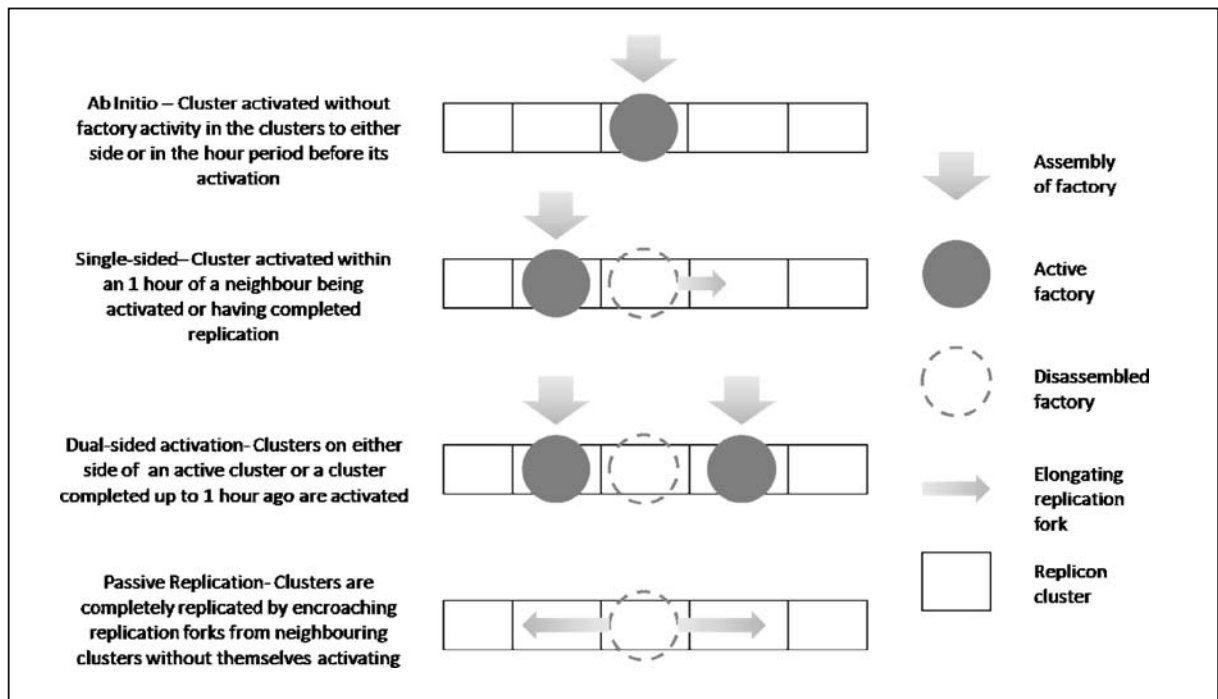


Fig. 5 Mechanisms of S phase propagation

2. Encroaching forks drive cluster firing

Next-in-line models of S-phase progression predict that the spread of encroaching forks is the driving factor that increases local firing probabilities. To simulate this, the model was set with a low maximum value of 0.0008/cluster/min for origin firing efficiency and a high multiplier value of 1,500, so that clusters with encroaching forks have a high probability of engaging synthesis (Fig. 6, position c). With these settings, 95% of DNA replication is completed in 8.3 h and total completion occurs within 11.2 h on average. The distribution of these completion times is more varied than in model (1), with a standard deviation of 73.7 min for the absolute completion times and 21.6 min for the 95% completion times. However, while the next-in-line conditions produce more variable end-times the dynamics of cluster firing give a better fit to biological profiles of origin activation (Goldar et al 2009) and yield a reduced rate (17%) of ab initio activation events.

3. Hybrid-driven cluster firing

A final possibility is that alterations in G band cluster firing efficiencies are driven by a mixture of

the mechanisms explored in (1) and (2). This was simulated in the model through a multiplier value for fork encroachment of ten and a maximum firing probability of 0.0022/cluster/min (Fig. 6, position d). The combination of factors still gives a 95% completion time of 8.4 h with absolute completion in 11.0 h. The variation of the completion times lies between that of the two alternative models, as does the rate of ab initio activations at 31.0%. With a ratio of single activation events to dual activation of 1.86, these conditions allow a significant increase in activation by encroaching forks relative to the stochastic model. However, the spatial effects are not strong enough to drive a high ratio of dual cluster activation events, as is seen at higher levels of spatial activation by fork encroachment.

To explore how changes in the chromatin environment might influence the switching of synthesis between neighbouring replicon clusters, we performed simulation that incorporated sub-optimal fork elongation rates in order to mimic possible fork stalling, which might occur as synthesis switches from one replication cluster to the next. Variable probability settings in the range 1–50% were used to simulate different extents of fork failure. From these simulations, it is evident that the ‘fork elongation’

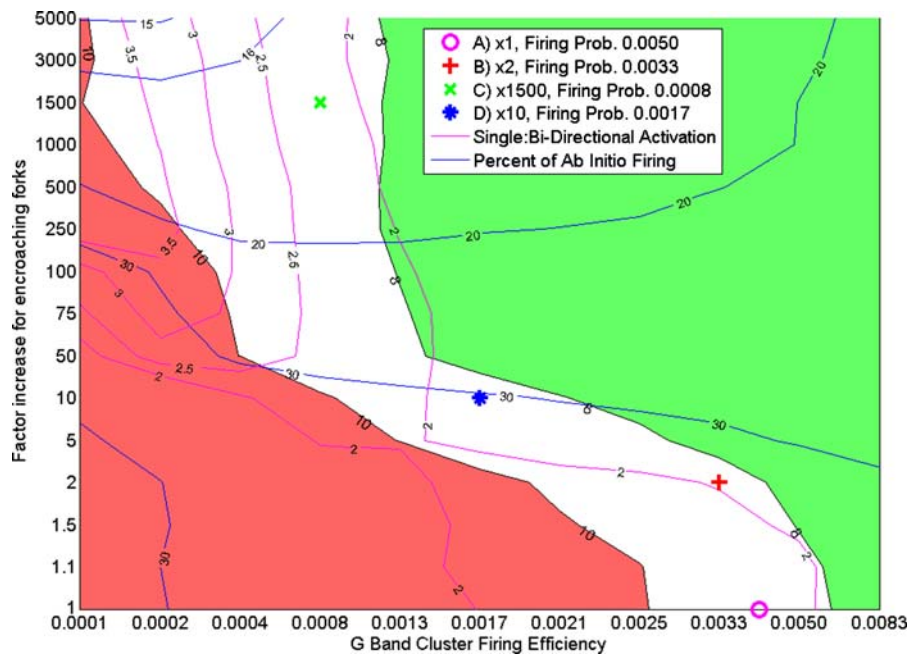


Fig. 6 A phase diagram of the explored parameter space. Using the method demonstrated in Fig. 4b, different firing probabilities were tested against a range of values to model spatial activation of DNA foci, using models described in Fig. 5. As synthesis within active clusters completes the extending forks growing out from the flanking replicons begin to interact with chromatin of neighbouring clusters. Here, we test how this influences the probability of activation within the adjacent cluster—the extent of this increase was modelled over a range of probabilities from $\times 1$ (no change) to $\times 5,000$ (highly probable). Given these parameter sets, an approximated phase space is created, which displays a number of key results: *Black contours* indicate completion times for replicating 95% of DNA. The red area indicates parameters giving a 95%

completion time over 10 h and the green area indicates parameter settings giving a 95% completion time of less than 8 h. The *white area* therefore represents a set of biologically relevant parameters within which S phase would complete on schedule. To assist interpretation, additional features of interest have been imposed over the analysis: (1) *magenta contours* indicate the ratio of single/dual activation events as described in Fig. 5—contours are labelled intermittently along their length (numbered 2, 2.5, 3, 3.5); (2) *blue contours* indicate percentage of ab initio firing events—contours are labelled intermittently along their length (numbered 15, 16, 20, 30). Biologically interesting positions (a–d) are indicated by coloured icons on the figure and discussed as case studies in the text

model is most susceptible to fork failure. Even so, a 6% probability is required to drive completion of 95% of DNA replication beyond 9 h and a 16% probability of failure is required to prolong S phase beyond 10 h. The ‘hybrid model’ is less sensitive to fork failure and completes 95% of DNA replication within 9 h even with a 15% chance of fork failure. In this case, S phase completes within 10 h as long as the probability of fork failure does not exceed 34%. Interestingly, increased levels of fork stalling also drives the hybrid model to generate a higher ratio of single/dual coupled activation events, while the spatially driven model maintains a constant ratio.

Predictions of replication timing profile generated by the final model were tested for biological efficacy by comparison with timing profiles generated using

TimEX-seq protocols from human ES cells (Desprat et al. 2009). The replication timing data for chromosome 6 was segmented into 100 time windows and a cumulative frequency profile showing the progress of DNA synthesis was generated (Fig. 7c). The whole chromosome profile was then segmented into R- and G-band regions using the recognised coordinates (see above) to generate separate timing profiles for the two major chromatin compartments. Comparison of the our S phase simulation with the TimEX-seq profiles (Fig. 7) shows that the replication timing data generated from human ES cells map closely to the data generated by our in silico simulations. Similarities were most evident at the level of total synthesis, where in both cases the accumulation of replicated DNA was essentially linear. However, the individual

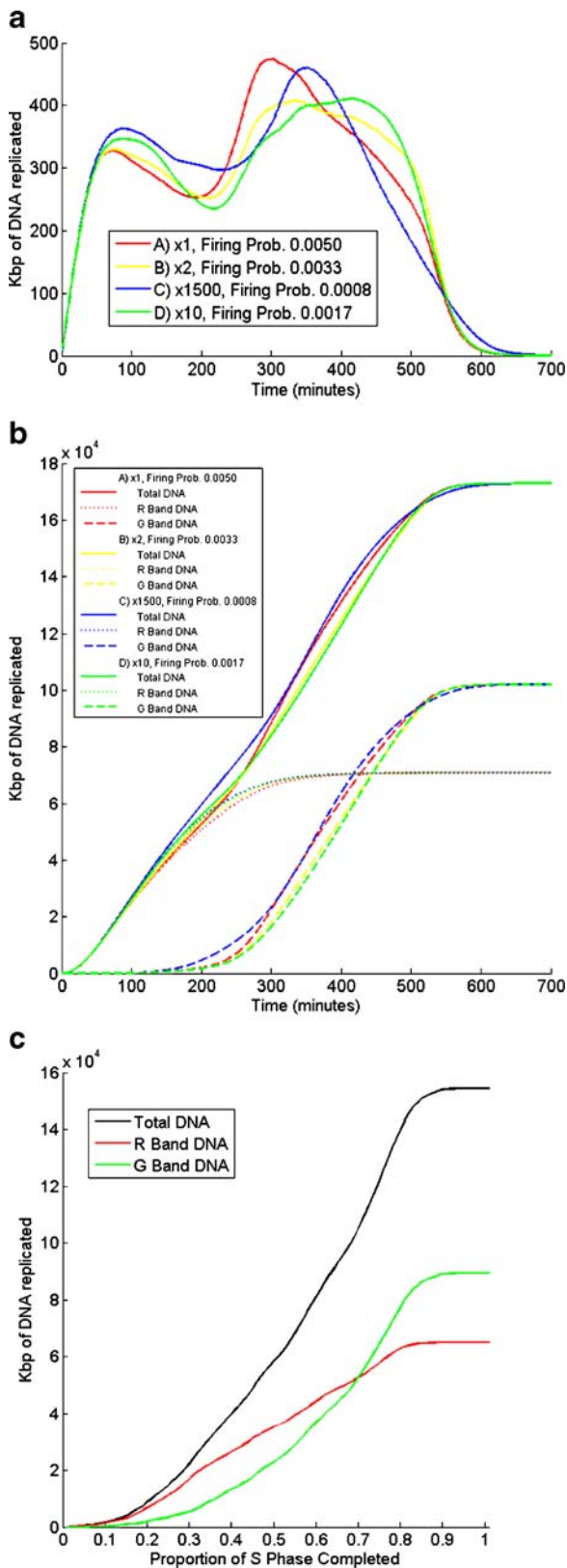


Fig. 7 Testing the models—comparison with genome-wide replication timing data. Four sets of simulations (**a**, **b**) were performed using the parameter sets highlighted in Fig. 6. For each, a model was created as described in Fig. 4, using one of a range of firing probabilities for G band clusters. For each firing probability, spatial effects were then tested based on the activation of clusters by encroaching forks (Fig. 5). Different plots (coloured lines) indicate parameters used in each set of simulations (see keys). Amplification factors (xn) define the adjusted firing probability that was applied when a cluster is activated by encroaching replication forks. Maximum firing probability refers to the probability of firing of a G-band cluster at the 6-h time point, based on the optimal concentration of activating cyclin-CDK complexes at that time. For each set of simulations (averages of 1,000 independent simulations are shown), the amount of DNA replicated (kbp/min) at each time point is determined (**a**) and converted into a cumulative replication profile (**b**), which shows the progress of synthesis. For each case study, solid lines indicate the total DNA replicated and broken lines display DNA synthesis within chromosomal R-bands (dotted lines) and G-bands (dashed lines). **c** The *in silico* simulations shown (**b**) were tested against experimentally derived profiles using the TimEX-seq data set from human ES cells (Desprat et al. 2009). The replication profile for chromosome 6 was generated by segmenting the published TimEX-seq data into 100 time intervals. This data set showing the amount of synthesis at different points throughout S phase was converted into a cumulative frequency plot of genome duplication across the sample. Plots showing S-phase progression were generated for the entire chromosome (Total DNA) and individually for chromosomal R- and G-bands, as shown

profiles for replication of R- and G-band DNA show significant discordance. This was particularly evident during mid-S phase, when the TimEX-seq data showed a higher level of G-band replication and prolonged R-band synthesis. Based on these profiles, the basic assumption that synthesis of R- and G-band DNA occurs during mutually distinct periods of S phase appears to be flawed. Hence, while the preference to engage synthesis in R-bands before G-bands is clear, the data do not suggest that an obligatory mechanism ensures that the cytologically defined chromosomal bands are replicated in a strict temporal order.

Conclusions and perspective

It has been known for many years that sites of initiation of DNA synthesis in mammalian cells are closely linked to local levels of RNA synthesis and that in general terms synthesis in gene-rich chromosomal R-bands occurs early in S phase and G-band

synthesis occurs later. Hence, the synthesis of mammalian genomes is thought to follow a temporal programme, which could be of fundamental biological importance if distinct chromatin states are specifically reproduced at defined times of S phase.

In this review, we set out to assess how different experimental approaches have been used to inform our understanding of DNA synthesis in mammalian cells and then assimilate ideas from different sources into a model of S-phase progression. Simulations were then used as an *in silico* approach to test alternative models of S phase. We tested a number of basic features related to genome architecture and local chromatin environments and then focussed on alternative mechanisms that might allow synthesis to propagate throughout the mammalian genome. In particular, we assessed how replication might spread between replication domains that contain ~1 Mbp of DNA. Specifically, we evaluated the behaviour of models of S phase that were based on both the stochastic activation of replication domains and the sequential activation of genetically linked DNA foci, according to the ‘next-in-line’ hypothesis of S-phase progression (Manders et al. 1992; Sporbert et al. 2002). As an alternative to these extremes, we considered a hybrid model, which incorporates a combination of S phase propagation using the next-in-line principle together with a level of external or *ab initio* activation events that are not influenced by encroaching forks from neighbouring replicons. Such initiation events might arise with different probabilities at different times of S phase, for example in response to changes in expression of specific cyclin-CDK complexes as S phase proceeds (Katsuno et al. 2009). The hybrid model incorporates a spatial component and temporal features related to changes in the chromatin environment. This model also accommodates a variable probability of origin activation so that the probability of clusters firing within G-bands remains low, but is enhanced by the presence of encroaching replication forks. Interestingly, we find that while this hybrid model is less reliant on fork elongation than the basic fork encroachment model, it shares some of the spatial dynamic benefits whilst being less susceptible to fork stalling. The fitness of this model is thus at least partially reliant on the probability that forks progress from one cluster to the next and appears to provide the best representation of the system *in vivo*.

In testing a range of alternative models, we have defined a parameter space that is likely to describe the biologically relevant mechanisms of S-phase progression in mammalian cells (region of interest highlighted in Fig. 6). Under the optimal parameter settings, comparison with experimental data shows that the model provides an excellent representation of replication for human chromosome 6 during the mammalian S phase (Fig. 7). However, we note a significant discrepancy between experimental data (Desprat et al. 2009) and our simulations of replication timing for designated R- and G-bands. This failure of the model implies that the chromosome-wide timing and order of R- and G-band replication is not defined with high precision. In particular, it is notable that while early cytological studies described a clear temporal separation in R- and G-band replication (Drouin et al. 1994) genome-wide analysis of the timing programme has shown that R-bands replicate before G-bands but that replication of the cytologically defined DNA compartments occurs throughout S phase (Desprat et al. 2009). Many features of the replication process might contribute to this observation. In particular, while genome-wide studies give a composite view of synthesis within huge cell populations it is clear that potential origins are used inefficiently so that the time of replication of specific chromosomal regions must reflect their location relative to the nearest active origin. While regions of the genome that have a high-density of active genes provide hot-spots for initiation of DNA replication—these will likely correlate with active regions at the onset of S phase – regions with lower levels of transcriptional activity provide weak targets for initiation and appear to replicate inefficiently, so that many potential origins are not used in most cells.

Based on our analysis, it is clear that the temporal restriction of R- and G-band replication to specific periods of S phase is an over-simplification that must be re-evaluated if we are to develop biologically robust models of S-phase progression. Specifically, it will become necessary to move away from the low resolution cytological chromosomal banding patterns, which generally incorporate chromosomal subdomains of 5–20 Mbp, and towards high-resolution patterns of chromatin epi-states that better reflect local patterns of gene expression. Such improvements in resolution should provide a better insight into the molecular mechanisms that drive the mammalian S phase so that synthesis is performed with the efficacy

required to ensure the preservation of genome integrity.

Acknowledgements The authors are indebted to BBSRC (AS, DJ), the Wellcome Trust (AM-M) and CONACyT (National Council for Science and Technology, Mexico; PO-C) for support.

References

- Aladjem MI (2007) Replication in context: dynamic regulation of DNA replication patterns in metazoans. *Nat Rev Genet* 8:588–600
- Berezney R, Dubey DD, Huberman JA (2000) Heterogeneity of eukaryotic replicons, replicon clusters and replication foci. *Chromosoma* 108:471–484
- Blow JJ, Dutta A (2005) Preventing re-replication of chromosomal DNA. *Nat Rev Mol Cell Biol* 6:476–486
- Buongiorno-Nardelli M, Micheli G, Carri MT, Marilley M (1982) A relationship between replicon size and supercoiled loop domains in the eukaryotic genome. *Nature* 298:100–102
- Cadoret J-C, Prioleau M-N (2010) Genome-wide approaches to determining origin distribution. *Chromosome Res* (in press)
- Cadoret J-C, Meisch F, Hassan-Zadeh V et al (2008) Genome-wide studies highlight indirect links between human replication origins and gene regulation. *Proc Natl Acad Sci USA* 105:15837–15842
- Conti C, Sacca B, Herrick J, Lalou C, Pommier Y, Bensimon A (2007) Replication fork velocities at adjacent replication origins are co-ordinately modified during replication in human cells. *Mol Biol Cell* 18:3059–3067
- Courbet S, Gay S, Arnoult N et al (2008) Replication fork movement sets chromatin loop size and origin choice in mammalian cells. *Nature* 455:557–560
- Cremer T, Cremer C (2001) Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat Rev Genet* 2:292–301
- DePamphilis ML, Blow JJ, Ghosh S et al (2006) Regulating the licensing of DNA replication origins in metazoan. *Curr Opin Cell Biol* 18:231–239
- Desprat R, Thierry-Mieg D, Lailler N et al (2009) Predictable dynamic programme of timing of DNA replication in human cells. *Genome Res* 19:2288–2299
- Donaldson AD (2005) Shaping time: chromatin structure and the DNA replication programme. *Trends Genet* 21:444–449
- Drouin R, Holmquist GP, Richer CL (1994) High-resolution replication bands compared with morphological G-bands and R-bands. *Advances Hum Genet* 22:47–115
- Farkash-Amar S, Simon I (2010) Genome-wide analysis of the replication programme in mammals. *Chromosome Res* (in press)
- Farkash-Amar S, Lipson D, Polten A et al (2008) Global organization of replication time zones of the mouse genome. *Genome Res* 18:1562–1570
- Fu Y, Sinha M, Peterson CL, Weng Z (2008) The insulator binding protein CTCF positions 20 nucleosomes around its binding sites across the human genome. *PLoS Genet* 4:e1000138+
- Ge XG, Blow JJ (2009) A model for DNA replication showing how dormant origins safeguard against replication fork failure. *EMBO Rep* 10:406–412
- Gilbert DM (2004) In search of the holy replicator. *Nat Rev Mol Cell Biol* 5:848–854
- Goen A, Cedar H (2003) Replicating by the clock. *Nat Rev Mol Cell Biol* 4:25–32
- Goetze S, Mateos-Langerak J, Gierman HJ et al (2007) The three-dimensional structure of human interphase chromosomes is related to the transcriptome map. *Mol Cell Biol* 27:4475–4487
- Goldar A, Marsolier-Kergoat M, Hyrien O (2009) Universal temporal profile of replication origin activation in eukaryotes. *PLoS ONE* 4:1–7
- Hadjur S, Williams LM, Ryan NK et al (2009) Cohesins form chromosomal cis-interactions at the developmentally regulated IFNG locus. *Nature* 460:410–413
- Hamlin JL, Mesner LD, Dijkwel PA (2010) A winding road to origin discovery. *Chromosome Res* (in press)
- Herrick J, Jun S, Bechhoefer J, Bensimon A (2002) Kinetic model of DNA replication in eukaryotic organisms. *J Mol Biol* 320:741–750
- Hiratani I, Ryba T, Itoh M et al (2008) Global reorganization of replication domains during embryonic stem cell differentiation. *PLoS Biol* 6:2220–2236
- Holmquist GP (1987) Role of replication time in the control of tissue-specific gene expression. *Am J Hum Genet* 40:151–173
- Hozak P, Jackson DA, Cook PR (1994) Replication factories and nuclear bodies: the ultrastructural characterization of replication sites during the cell cycle. *J Cell Sci* 107:2191–2202
- Hyrien O, Goldar A (2010) Mathematical modelling of eukaryotic DNA replication. *Chromosome Res* (in press)
- Jackson DA (1995) Nuclear organization: uniting replication foci, chromatin domains and chromosome structure. *BioEssays* 17:587–591
- Jackson DA, Pombo A (1998) Replicon clusters are stable units of chromosome structure: evidence that nuclear organization contributes to the efficient activation and propagation of S phase in human cells. *J Cell Biol* 140:1285–1295
- Jeon Y, Bekiranov S, Karnani N et al (2005) Temporal profiles of replication of human chromosomes. *Proc Natl Acad Sci USA* 102:6419–6424
- Karnani N, Taylor C, Malhotra A, Dutta A (2007) Pan-S replication patterns and chromosomal domains defined by genome-tiling arrays of ENCODE genomic areas. *Genome Res* 17:865–876
- Karolchik D, Baertsch R, Diekhans M et al (2003) UCSC genome browser database. *Nucl Acids Res* 31:51–54
- Katsuno Y, Suzuki A, Sugimura K et al (2009) Cyclin A-CDK1 regulates the origin firing programme in mammalian cells. *Proc Natl Acad Sci USA* 106:3184–3189
- Lande-Diner L, Zhang JM, Cedar H (2009) Shifts in replication timing actively affect histone acetylation during nucleosome reassembly. *Mol Cell* 34:767–774
- Lebofsky R, Heilig R, Sonnleitner M, Weissenbach J, Bensimon A (2006) DNA replication origin interference increases the

- spacing between initiation events in human cells. *Mol Biol Cell* 17:5337–5345
- Leonhardt H, Rahn HP, Weinzierl P et al (2000) Dynamics of DNA replication factories in living cells. *J Cell Biol* 149:271–279
- Lygeros J, Koutroumpas K, Dimopoulos S et al (2008) Stochastic hybrid modeling of DNA replication across a complete genome. *Proc Natl Acad Sci USA* 105:12295–12300
- Ma H, Samarabandu J, Devdhar RS et al (1998) Temporal dynamics of DNA replication sites in mammalian cells. *J Cell Biol* 143:1415–1425
- Manders EM, Stap J, Brakenhoff GP, van Driel R, Aten JA (1992) Dynamics of 3-dimensional replication patterns during the S-phase, analyzed by double labelling of DNA and confocal microscopy. *J Cell Sci* 103:857–862
- Maya-Mendoza A, Petermann E, Gillespie DA, Caldecott KW, Jackson DA (2007) Chk1 regulates the density of active replication origins during the vertebrate S phase. *EMBO J* 26:2719–2731
- Maya-Mendoza A, Tang CW, Pombo A, Jackson DA (2009) Mechanisms regulating S phase progression in mammalian cells. *Front Biosci* 14:4199–4213
- Mendez J, Stillman B (2003) Perpetuating the double helix: molecular machines at eukaryotic DNA replication origins. *BioEssays* 25:1158–1167
- Mesner LD, Crawford EL, Hamlin JL (2006) Isolating apparently pure libraries of replication origins from complex genomes. *Mol Cell* 21:719–726
- Nakamura H, Morita T, Sato C (1986) Structural organizations of replication domains during DNA synthetic phase in the mammalian nucleus. *Exp Cell Res* 165:291–297
- Parelho V, Hadjur S, Spivakov M et al (2008) Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell* 132:422–433
- Phillips JE, Corces VG (2009) CTCF: master weaver of the genome. *Cell* 137:1194–1211
- Pope BD, Hiratani I, Gilbert DM (2010) Domain-wide regulation of DNA replication timing during mammalian development. *Chromosome Res* (in press)
- Raghuraman MK, Winzeler EA, Collingwood D et al (2001) Replication dynamics of the yeast genome. *Science* 294:115–121
- Rhind N (2006) DNA replication timing: random thoughts about origin firing. *Nat Cell Biol* 12:1313–1316
- Rhind N, Yang SC-H, Bechhoefer J (2010) Reconciling stochastic origin firing with defined replication timing. *Chromosome Res* (in press)
- Sequeira-Mendes J, Diaz-Uriarte R, Apedaile A et al (2009) Transcription initiation activity sets replication origin efficiency in mammalian cells. *PLoS Genet* 5:e1000446
- Shopland LS, Lynch CR, Peterson KA et al (2006) Folding and organization of a contiguous chromosome region according to the gene distribution pattern in primary genomic sequence. *J Cell Biol* 174:27–38
- Sporbert A, Gahl A, Ankerhold R, Leonhardt H, Cardoso MC (2002) DNA polymerase clamp shows little turnover at established replication sites but sequential de novo assembly at adjacent origin clusters. *Mol Cell* 10:1355–1365
- St-Jean P, Vaillant C, Audit B, Arneodo A (2008) Spontaneous emergence of sequence-dependent rosette like folding of chromatin fiber. *Phys Rev E* 77:e061923
- Strehl S, Lasalle J, Lalande M (1997) High-resolution analysis of DNA replication domain organization across an R/G-band boundary. *Mol Cell Biol* 17:6157–6166
- Takebayashi SI, Manders EM, Kimura H, Taguchi H, Okumura K (2001) Mapping sites where replication initiates in mammalian cells using DNA fibres. *Exp Cell Res* 271:263–268
- Tuduri S, Tourrière H, Pasero P (2010) Defining replication origin efficiency using DNA fiber assays. *Chromosome Res* (in press)
- Vogelauer M, Rubbi L, Lucas I, Brewer BJ, Grunstein M (2002) Histone acetylation regulates the time of replication origin firing. *Mol Cell* 10:1223–1233
- West AG, Fraser P (2005) Remote control of gene transcription. *Hum Mol Genet* 14:R101–R111
- Woodfine K, Fiegler H, Beare DM et al (2004) Replication timing of the human genome. *Hum Mol Genet* 13:191–202
- Zhang JM, Xu F, Hashimshony T, Keshet N, Cedar H (2002) Establishment of transcriptional competence in early and late S phase. *Nature* 420:198–202
- Zhang N, Kuznetsov SG, Sharan SK, Li K, Rao PH, Pati D (2008) A handcuff model for the cohesin complex. *J Cell Biol* 183:1019–1031
- Zink D (2006) The temporal programme of DNA replication: new insights into old questions. *Chromosoma* 115:273–287