# PERSPECTIVES

# In search of the holy replicator

*David M. Gilbert*

Abstract | After 40 years of searching for the eukaryotic replicator sequence, it is time to abandon the concept of 'the' replicator as a single genetic entity. Here I propose a 'relaxed replicon model' in which a positive initiator–replicator interaction is facilitated by a combination of several complex features of chromatin. An important question for the future is whether the positions of replication origins are simply a passive result of local chromatin structure or are actively localized to coordinate replication with other chromosomal activities.

In the summer of 1962, François Jacob and Sydney Brenner sketched in the sand of France's La Tranche-sur-Mer a model (FIG. 1a) to explain the regulation of DNA synthesis in bacteria[1]. Their 'replicon model' proposed that replication was regulated by a positive interaction between an initiator protein, which is encoded by a structural gene, and a specific genetic element of recognition, which is known as the replicator. Although originally intended to explain the replication of circular bacterial chromosomes, the replicon model was quickly adopted as the paradigm for eukaryotic chromosomes as well (FIG. 1b). Proteins with the properties of initiators have been identified in species from *Escherichia coli* to man. By contrast, genetically defined replicators have yet to be identified in most eukaryotes and are clearly dispensable for replication in many eukaryotic systems. Even in its shortcomings, though, the model has been useful. Indeed, the replicon model was the driving force behind the discovery of bacterial and budding yeast replicators as well as initiator proteins — from dnaA in prokaryotes[2]

to the origin-recognition complex (ORC) in eukaryotes[3]. Indeed, all cellular chromosomes can be thought of as replicons, or clusters of replicons — the replication of which is regulated, at least in part, by a positive initiator–replicator interaction.

In this historical perspective, I will summarize the salient findings over the course of these four decades (TIMELINE), with a focus on strategies to identify metazoan replicators and the debate over their existence. I propose that we replace the concept of the replicator as a specific DNA sequence with that of a context-dependent element that is recognized by the initiator through several features, which can include DNA sequence. By embracing the complexity of replicators as being akin to that of promoter elements, future experiments can define these various features that influence the initiator–replicator interaction and address the biological significance of specifying replication-origin sites, given that site-specific initiation is dispensable for regulated genome duplication[30].

## The power of prediction
One of the beauties of the replicon model was the conceptually straightforward predictions that it provided. A replicator can confer the ability to replicate onto any DNA segments that are linked to it, and this became relatively simple to test after the development of recombinant-DNA technology in the early 1970s. By linking various segments of *E. coli* plasmids to a non-replicating DNA fragment that contained a selectable marker, Stanley Cohen and colleagues identified sequences that could confer, in *cis*, the ability of these molecules to replicate autonomously[4]. This

straightforward assay (BOX 1) was later used to identify many different replicators, including oriC in *E. coli*[5]. Once a replicator sequence is defined, the initiator can be identified as a protein that binds to the replicator, and the dnaA protein in *E. coli* fulfilled this criterion[2]. Allowing for a few minor adjustments, it is clear that the replicon model is not only a useful paradigm, but that it accurately describes all bacterial replication systems under normal physiological selective pressures. Moreover, the replicator DNA sequence and the initiator protein are highly conserved across different bacterial species[6].
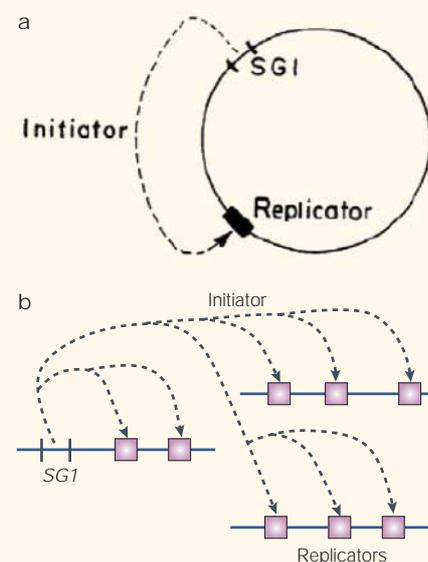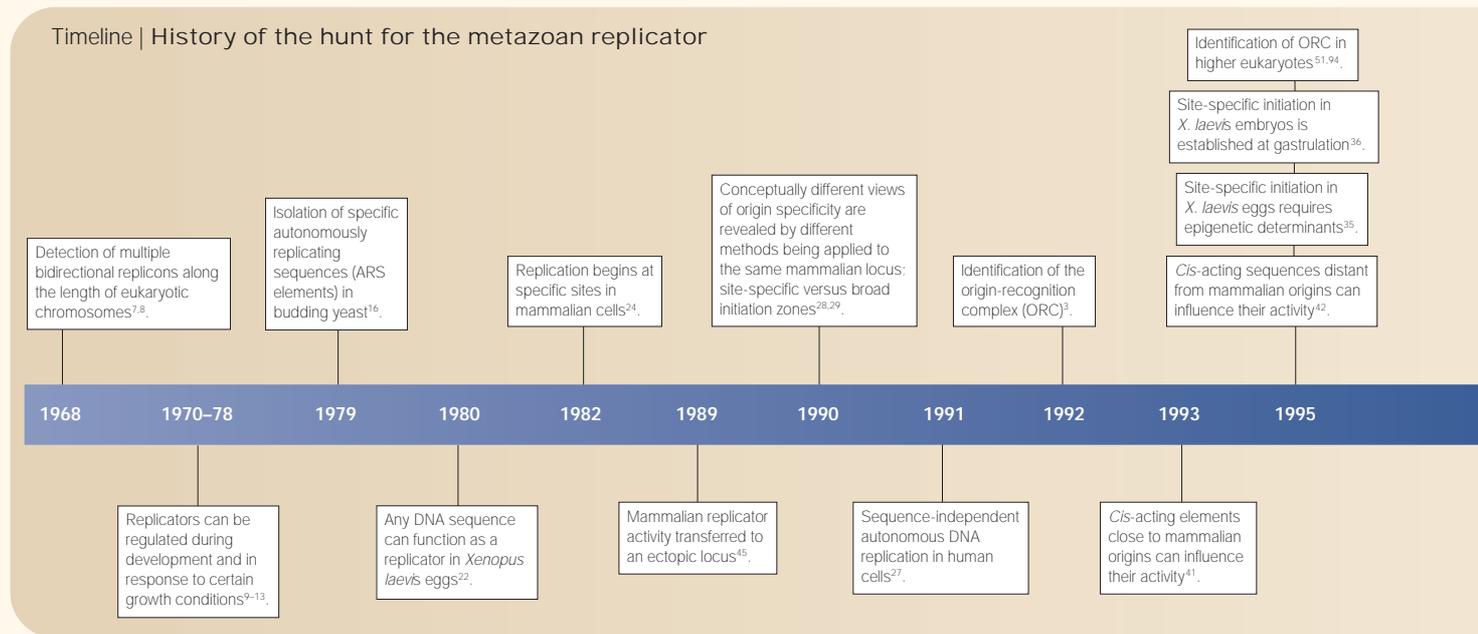
Figure 1 | **The replicon model. a** | Original conception of the replicon model. "The replicon is assumed to be a circular structure carrying two specific genetic determinants. A structural gene (*SG1*) determines the synthesis of a diffusible active element, the initiator. The initiator acts on a replicator, allowing the beginning of the replication which proceeds along the circular structure"[1]. Reproduced with permission from REF. 1 © (1964) Cold Spring Harbor Laboratory Press. **b** | The replicon model in the form that it was originally presumed to apply to eukaryotes.

Timeline | **History of the hunt for the metazoan replicator**

Detection of multiple bidirectional replicons along the length of eukaryotic chromosomes[7,8].

Isolation of specific autonomously replicating sequences (ARS elements) in budding yeast[16].

Replication begins at specific sites in mammalian cells[24].

Conceptually different views of origin specificity are revealed by different methods being applied to the same mammalian locus: site-specific versus broad initiation zones[28,29].

Identification of the origin-recognition complex (ORC)[3].

*Cis*-acting sequences distant from mammalian origins can influence their activity[42].

Site-specific initiation in *X. laevis* eggs requires epigenetic determinants[35].

Site-specific initiation in *X. laevis* embryos is established at gastrulation[36].

Identification of ORC in higher eukaryotes[51,94].

| 1968 | 1970–78 | 1979 | 1980 | 1982 | 1989 | 1990 | 1991 | 1992 | 1993 | 1995 |

Replicators can be regulated during development and in response to certain growth conditions[9–13].

Any DNA sequence can function as a replicator in *Xenopus laevis* eggs[22].

Mammalian replicator activity transferred to an ectopic locus[45].

Sequence-independent autonomous DNA replication in human cells[27].

*Cis*-acting elements close to mammalian origins can influence their activity[41].

## Eukaryotes: a different problem

It did not take long for investigators to appreciate that eukaryotic chromosomes had fundamentally different problems to solve, which could not be accommodated by the replicon model in its simplest form. Whereas bacteria usually have a single circular chromosome with only one replicator per chromosome, eukaryotes have several linear chromosomes that are each many times larger. Early autoradiography studies of DNA synthesis along the lengths of isolated DNA fibres showed that these chromosomes replicate using hundreds to thousands of replication origins[7,8] (TIMELINE). Moreover, there was clearly flexibility in the usage of replication origins at different developmental stages[9–11] and under different growth conditions[12,13]. Also, replication origins were shown to fire asynchronously throughout the DNA-synthesis phase of the cell cycle[14,15]. So, even though the existence of a common eukaryotic replicator sequence was proposed (FIG. 1b), it was already clear that additional layers of regulation would be needed to explain such flexibility.

Given these known complexities, perhaps it was not fair to hold the bacterial replicon model to the standard of eukaryotic replication. But the advent of budding yeast genetics in the 1970s provided the ideal model system in which to use the autonomous-replication assay, which had been so successful in bacterial systems, to identify eukaryotic replicators. The immediate success of this approach[16], which resulted in the cloning of several 'autonomously replicati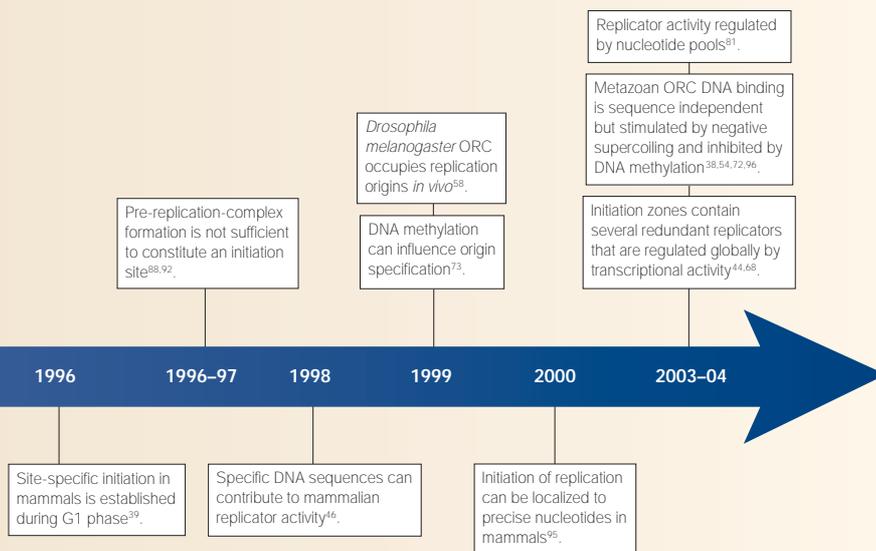ng sequences' (ARS elements), all but confirmed the replicon model as a useful paradigm for eukaryotes. Indeed, ARS elements have since been shown to be sites of replication initiation in their native context[17–19], and to contain a common DNA sequence that is required for replication initiation both on plasmids[20] and within chromosomes[21]. Hence, this proof-of-principle experiment that was carried out 25 years ago set the stage for the hunt for replicators in higher eukaryotic organisms. Ironically, budding yeast has turned out to be an exceptional eukaryote with respect to the importance of a specific consensus replicator sequence.

## The mystery of the metazoa

As they have genomes that are up to 100 times larger in size than those of yeast, mapping replication origins in higher eukaryotes was expected to be difficult. But no one could have predicted the degree of disparity that was uncovered between different systems, and even within individual systems that were studied using different methodologies. In 1980, Richard Harland and Ron Laskey demonstrated that any of several DNA fragments that were microinjected into *Xenopus laevis* eggs, even those containing mostly bacterial DNA, replicated with equal efficiency and in a regulated once-per-cell-cycle fashion[22]. The results were compelling, and they were reproducible[23]. In stark contrast, Nick Heintz and Joyce Hamlin demonstrated shortly thereafter that initiation of replication could be localized to specific sites in mammalian chromosomes[24]. So, the great debate had begun. The *X. laevis* experiments showed that once-per-cell-cycle replication could be accomplished without the need for specific replicator sequences. But mammalian cells clearly did not initiate replication at random sites, and *X. laevis* embryos are a specialized system, which is designed to duplicate the entire genome within 20 minutes. For most, this was sufficient reason to disregard the *X. laevis* embryo results, and the quest for the 'holy grail' of the human replicator shifted into high gear.

## ARS assays: all or nothing

The obvious first strategy of choice was the autonomously replicating sequence (ARS) assay that had worked so well in bacteria and yeast (BOX 1). It is impossible to say how many clever permutations of this approach were tried and failed or found to be non-reproducible[25]. In most cases, very poor transient replication activity[26] quickly resulted in selection for plasmids that had stably integrated into chromosomes[25]. Then, in the late 1980s, Michelle Calos and colleagues demonstrated that the efficiency of ARS assays could be greatly improved by preventing the loss of plasmids from the nucleus using DNA sequences that tether plasmids to segregating chromosomes. The results of these experiments were astonishingly clear. Plasmids could be maintained for many generations, but replication was independent of DNA sequence and was initiated at random sites within these plasmids[27]. The dirty 'random' word had reared its ugly head again, placing a nasty stigma on the ARS assay. For those bent on finding the holy replicator, it was clear that some nuance of mammalian cell culture precluded the ability of ARS assays to reveal physiological replicator activity.

Timeline content:

Replicator activity regulated by nucleotide pools[81].

Metazoan ORC DNA binding is sequence independent but stimulated by negative supercoiling and inhibited by DNA methylation[38,54,72,96].

*Drosophila melanogaster* ORC occupies replication origins *in vivo*[58].

Pre-replication-complex formation is not sufficient to constitute an initiation site[88,92].

DNA methylation can influence origin specification[73].

Initiation zones contain several redundant replicators that are regulated globally by transcriptional activity[44,68].

| 1996 | 1996–97 | 1998 | 1999 | 2000 | 2003–04 |

Site-specific initiation in mammals is established during G1 phase[39].

Specific DNA sequences can contribute to mammalian replicator activity[46].

Initiation of replication can be localized to precise nucleotides in mammals[95].

## Religion, politics and replication origins

By the early 1990s, most investigators had conceded that classic genetic approaches, including those that employed promising viral models (BOX 2), were insufficient for identifying metazoan replicators, and many sought a more complete characterization of replicators in their native context. Finding replication intermediates in a genome of 3 billion basepairs required new technological developments, and several new methods for mapping replication origins emerged. Rather than clarifying things, however, the 'random versus site specific' debate escalated to a whole new level. In 1990, almost simultaneously, two reports in which different methods were used to map initiation sites within the Chinese hamster dihydrofolate reductase (*DHFR*) locus arrived at diametrically opposed conclusions. One study claimed that >80% of initiations within a 27-Kb region emanated from a single site[28], whereas the other claimed that replication initiated at sites distributed virtually at random throughout the very same region[29]. These results shifted the debate from a philosophical comparison of different biological systems to a detailed technical evaluation of different methodological approaches, leaving investigators unable to reach a consensus as to the basic structure of metazoan replication origins nor the best means to evaluate this structure.

Nearly 15 years have passed and the distribution of initiation sites at this and many other metazoan loci has still not been resolved. As more replication origins were mapped, some were localized to very defined sites, whereas others mapped indiscriminately throughout 'initiation zones' of 10–50 Kb (reviewed in REF. 30). Most of these sites have been mapped by only one method, so the extent to which these differences can be accounted for by methodological considerations is not clear. On the other hand, nearly every replication-origin mapping method has been applied to the *DHFR* locus, and there is still no consensus on the pattern of initiation at this locus. Some methods detect specific 'replicator' sites within the non-transcribed region downstream of the *DHFR* locus[31], whereas other methods reveal initiation sites throughout every interrogated sequence within the same region[32]. At present, the best explanation for these discrepancies is that initiation can take place at many sites throughout this 50-Kb region, but that some methods are only capable of detecting a subset of these sites.

In the meantime, people who worked with the *X. laevis* system had all but proven that any DNA sequence can function as a replicator. After the development of cell-free replication systems that were derived from *X. laevis* eggs, many investigators were able to demonstrate the promiscuity of replication in this system, and the apparently random choice of initiation sites within each DNA molecule[33–35]. The same was found to be true within the native chromosomes of early *X. laevis*[36] and *Drosophila melanogaster*[37] embryos. Importantly, these latter studies revealed that replication initiation is eventually focused to specific sites after the blastula stage of development. These results were interpreted along existing party lines. Proponents of site-specific initiation assumed that random site selection was a special adaptation for rapid replication in early embryos, which was mediated either by a specialized embryonic initiator protein or by the high concentration of maternally accumulated initiator proteins that become diluted during development. These possibilities were ruled out by reconstituting ORC-depleted *X. laevis* egg extracts with human ORC, which restored sequence-independent initiation of replication in a concentration-independent manner[38]. The other party declared that developmentally regulated features of chromatin, rather than a specific DNA sequence, dictate origin specification. This interpretation was supported by the fact that *X. laevis* egg extracts are capable of initiating replication at specific sites when provided with chromatin from somatic cells as a template[35], and that the defining epigenetic structure is re-assembled early in G1 phase of each cell cycle[39]. In fact, simply pre-loading transcription factors onto naked DNA templates is sufficient to direct site-specific initiation in pre-blastula *X. laevis* embryos[40].

## DNA sequence: a clear contributor

New hope in the hunt for the holy replicator arrived when Howard Cedar and colleagues demonstrated that a naturally occurring 8-Kb deletion that encompassed a region of replication-initiation activity within the human β-globin locus (*HBB*) eliminated bidirectional replication in this region[41]. Hailed as a triumph that would deliver us from the dark ages, this result inspired further deletion studies that soon revealed a more complex picture. At both the *HBB* and the *DHFR* locus, deletions 25–50 Kb away from the region of replication-initiation activity also completely eliminated its activity[42,43]. Deletions within the large region of activity at the *DHFR* locus actually augmented activity at the remaining sites[44], implying a high degree of redundancy of potential replicators throughout this region.

A critical test of 'replicator' activity is to show that a DNA segment can direct the initiation of replication when placed in different sequence contexts. In the absence of convenient ARS assays, the next best option was to move putative replicators to other locations within cellular chromosomes. The first such study provided evidence that a 16-Kb segment from within the Chinese hamster ovary (CHO) *DHFR* replication-initiation zone could function as a replicator when integrated at random ectopic chromosomal sites[45]. However, this result was not universally accepted because negative-control fragments could not be directly compared at the same chromosomal sites. This limitation

### Box 1 | The autonomously replicating sequence (ARS) assay
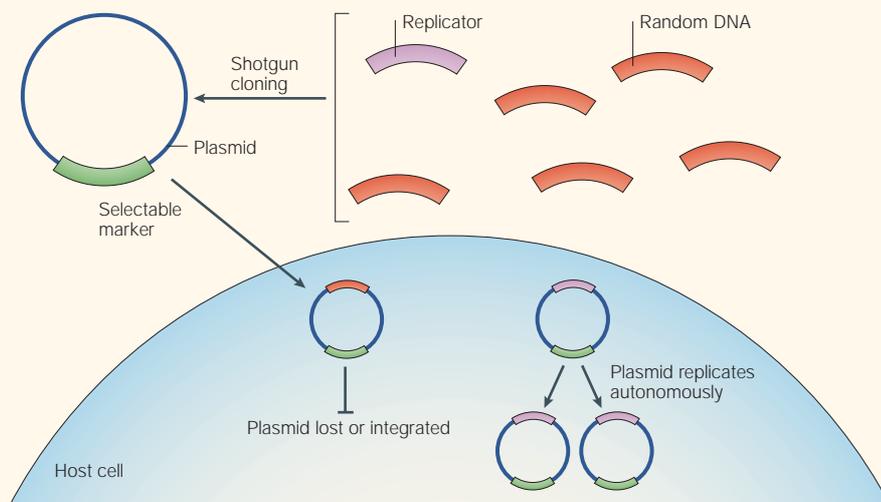


The ARS assay is used to identify replicators in genomes or DNA sequences that can range from those that are as simple as small bacterial plasmids and animal viruses to those that are as complex as budding and fission yeast genomes — the principle is very similar in all cases. The test genome is fragmented and the fragments are cloned into a vector that contains a selectable marker that can function in the appropriate host. The appropriate host is a cell type that is known to contain all the proteins that are necessary to support replication of the desired replicator, including the initiator. The collection of plasmids containing these genome fragments is then introduced into the host in such a way that each cell obtains one or fewer plasmids. Selection for the presence of the plasmid is applied, and the resulting recipient clonal colonies are expanded and evaluated for the ability of the (usually circular) plasmid to be maintained and to replicate as an independent unit, autonomously, within the host cell. In some variations of this assay, replicating plasmids can be detected or enriched without cell-growth selection. Myriad permutations of this assay have been applied to metazoan systems, but have generally failed to identify any specific DNA sequences that confer ARS activity.

was overcome by using microbial recombinases to integrate test fragments into the same site, which provided persuasive evidence that specific DNA sequences are required for replicator activity[46]. More recent studies have revealed conceptually similar results, whether the sequences are derived from highly localized replication origins[47–49] or from the broad *DHFR* initiation zone[50]. In all cases, 'replicator' activity can be transferred to many (but not all) ectopic sites. Importantly, however, these studies have not identified any specific DNA sequences or structural motifs that are shared by different replicators.

Together, these studies show that initiation of replication in mammalian cells requires specific DNA sequences that are both close to and distant from the site of initiation, but no single consensus DNA-sequence motif is either necessary or sufficient for replicator activity. The simplest interpretation of these studies is that there is specificity to the initiator–replicator interaction, but the features that are recognized by the initiator are only indirectly determined by DNA sequences.

### Replicator–initiator interactions

The discovery of a hetero-hexameric ORC in all eukaryotes from budding yeast to man[3,51] provided a whole new strategy to hunt for replicators. Instead of searching for poorly

defined sequences with laborious methods, the sequences to which the initiator binds, logically, should define the replicator. In fact, budding yeast ORC binds to replication-origin DNA sequences substantially better than to random templates[52], and fission yeast ORC has evolved a specialized domain to recognize asymmetric stretches of adenines and thymidines that define its origins of replication[53]. So, it was quite revealing to learn that purified human and *D. melanogaster* ORC have little or no sequence-specific binding activity[38,54]. Interestingly though, *D. melanogaster* ORC has a strong preference for binding to negatively supercoiled DNA[54]. A preference for superhelical tension could explain why both local and distant sequences can affect initiation activity without the requirement for a specific consensus-sequence recognition element.

Chromatin immunoprecipitation (ChIP) experiments can potentially identify ORC-binding sites regardless of their DNA sequence, and recent studies suggest that ORC occupies replication origins in all eukaryotes, although its relative binding affinity compared to other sites is unknown. As expected, budding yeast[55] and fission yeast[56] ORCs show significant enrichment at origins, but some preference for origin occupation is also seen in flies[57,58] and humans[59]. One drawback of most ChIP studies is that they interrogate only a few carefully chosen sites, which does not provide an unbiased view of the sequences to which ORC binds. Hybridization of DNA sequences that co-immunoprecipitate with ORC in complete-genome arrays (ChIP-chip) can provide this information and have already proved successful in budding yeast[60]. Interestingly, not all ARS consensus sequences are identified, which

### Box 2 | The contribution of viral replicators

The small double-stranded DNA viruses (for example, polyoma and papilloma viruses) were the preferred model for studying eukaryotic replication in the 1980s, and because they use the host's machinery for replication, they were particularly promising models for mammalian chromosome replication. In fact, most of our knowledge of the events that occur downstream of replication-origin unwinding comes from studies with simian virus-40 (SV40), in which defined *in vitro* replication systems identified the proteins that are essential for DNA synthesis in human cells[83,84]. Initiation of replication clearly requires the interaction of a virally encoded initiator protein with a defined replicator sequence that could be genetically dissected with classic autonomously replicating sequence (ARS) assays[85]. But these viruses do not replicate once per cell cycle, they supply their own virally encoded initiator, and they are not subjected to the same selective pressures as cellular chromosomes, so they are limited in their use as models for mammalian replicator function.

By contrast, the large Epstein–Barr virus (EBV) does not supply its own initiator, but instead requires ORC[86]. Latent (non-lytic) replication of EBV is restricted to one round per cell cycle[87], making this virus a good potential model for cellular replication. Interestingly though, a region known as oriP, which was shown to have ARS activity in the presence of a virally encoded protein, is dispensable for latent viral replication[87], and origin-mapping studies have shown that initiation-site selection in the full-sized EBV genome is just as degenerate as it is within cellular chromosomes, whether or not oriP is present[82].

Box 3 | **Regulation downstream of the initiator–replicator interaction**

In addition to regulating the initiator–replicator interaction, the activity of a potential replication origin in metazoans is also strongly influenced by events downstream of the ORC–DNA interaction. In budding yeast, there are many origins that do not fire in every cell cycle and some that do not fire at all in a normal cell cycle, but these silent origins are occupied by ORC and other pre-RC proteins[88]. In fact, ORC is even more tightly bound to silent origins[89]. Many of these silent origins are activatable under the appropriate conditions, which implies that functional pre-RCs are assembled at these sites. The conditions that silence some budding yeast origins or activate others are beginning to be understood, and they include regulators of chromatin structure[77], nucleosome-free transcriptional-control elements[69], proximity to nucleosomes[76], or simply proximity to other replication origins[90,91]. In *X. laevis* embryos, there are clearly many more pre-RCs assembled than are used[92], and the relative preference for initiating replication at specific sites can be influenced downstream of ORC binding by the presence of transcription factors[40]. In *D. melanogaster*, binding of a protein complex that contains the oncoprotein Myb is required for efficient initiation of replication at the chorion gene cluster during selective gene amplification, but is not required for ORC localization to this origin[93]. For the time being, it is important to appreciate that in many cases it has not been determined whether the factors that influence origin activity are upstream or downstream of ORC binding, and that this will be an important distinction for our future understanding of replication origins.

EBNA1 protein of Epstein–Barr virus (EBV) interacts with ORC and might recruit ORC to the EBV origin[63]. In *D. melanogaster* follicle cells, an interaction of ORC with E2F transcription factors[64,65] — possibly in conjunction with a developmentally regulated Myb-oncoprotein-containing complex[66] and the histone de-acetylase Rpd3 (REF. 67) — seems to orchestrate a developmentally regulated re-distribution of ORC to facilitate amplification of the chorion gene cluster. Chromosomal activities such as transcription can also desta-bilize nucleosomes and create superhelical tension, which might explain the transcrip-tion-dependent origin activity that has been described in several systems[68–70]. In short, any mechanism that removes a nucleosome and results in negative supercoiling could create a high-affinity binding site for ORC[54].

Another way to direct a nonspecific binding protein to specific sites is to prevent it from binding to other sites. Exclusion can occur at a global level, for example by tran-scriptional interference, which antagonizes the initiation of replication[68,71], or by the interaction with global repressors, as is sug-gested during amplification of the chorion gene cluster in *D. melanogaster*[64–67]. There can also be more localized exclusion effects. A recent study shows that methylated DNA will neither bind ORC nor initiate replica-tion in the otherwise omnipotent *X. laevis* egg extracts[72]. By cleverly constructing plas-mid templates that contain a small segment of unmethylated DNA, ORC binding was

could be due to a lack of sensitivity of the method, to a requirement for further unde-fined sequences[61], or to features of chromatin that prevent ORC from binding to some sites. Another confounding issue with these studies is that, counter-intuitively, stronger ORC binding sites have been shown to be less effi-cient replicators[89]. In any case, these studies require arrays that achieve complete unbiased coverage of the genome, which, at present, are available for very few species. An alternative approach is to clone and sequence DNA seg-ments that co-immunoprecipitate with ORC, which was recently done for a human cell line[62]. In both the human and yeast studies, many (but not all) ORC-binding sites coincide with replication-initiation activity (BOX 3).

If purified ORC does not preferentially bind to specific sequences, and replicators do not share a common sequence element, then what directs ORC to specific chromosomal sites *in vivo*? Several mechanisms can be imagined (FIG. 2). One possibility is that other proteins interact with ORC to direct it to spe-cific sequences. This could occur through the direct interaction of ORC with chromatin proteins, or such proteins could alter the local topological state of DNA. For example, the

Figure 2 | **Relaxed replicon model.** The hetero-hexameric origin-recognition complex (ORC) binds to naked DNA indiscriminately on its own, but the specificity of its binding to cellular chromatin *in vivo* is influenced by many factors. When presented with functionally inert (open or 'indiscriminate') chromatin (**a**), ORC binds nonspecifically[38]. However, ORC might be directed to specific sites by (**b**) interacting proteins that chaperone ORC to specific sites[63] or repress its binding to others[64–67]. Or (**c**), by superhelical tension[54], which can be created by nearby binding proteins (X) that might also position or remove nucleosomes, or otherwise create favourable ORC-binding sites. Superhelical tension can also be created by transcriptional activity, either upstream or downstream of gene promoters[69,70]. ORC can also be excluded from specific sites by (**d**) transcription[68,71] or by (**e**) general CpG DNA methylation[72], which can be quite prevalent[73]. Additional features of chromatin that lie downstream of the initiator–replicator interaction also influence origin specification (BOX 2). This model necessarily invokes complexity — origin specification results from a different combination of targeting mechanisms for each chromosomal site, making it no longer useful to think of the replicator as a single common genetic entity.

---
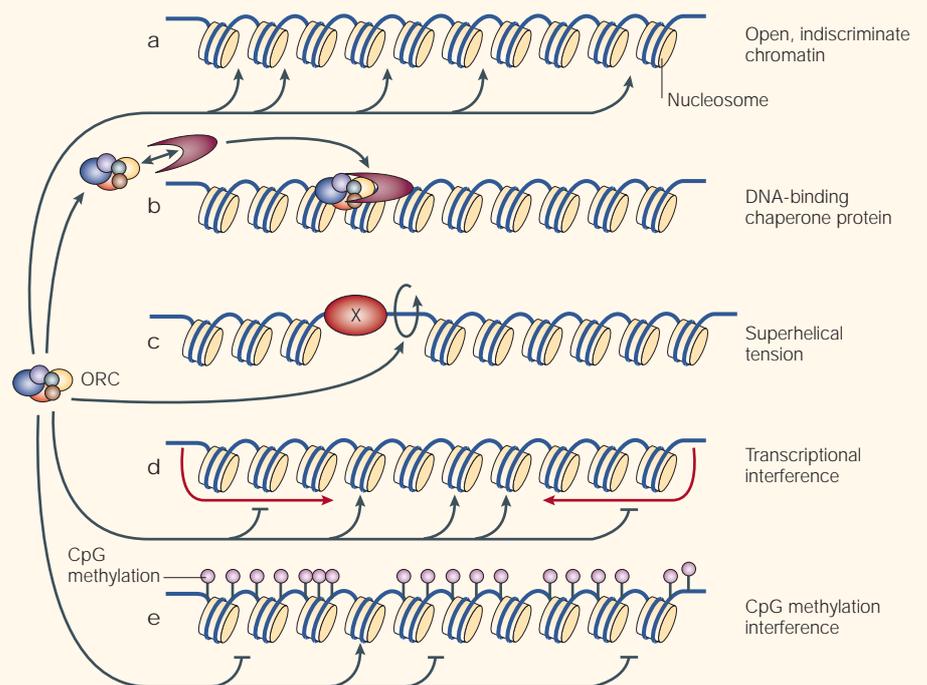
> Box 4 | **The status quo: a list of acceptable statements**
>
> **Replicators**
> - There are many more replicators (that is, potential replication origins that are occupied by the origin-recognition complex (ORC)) than are used during S phase.
> - More replicators are used than are needed to complete S phase.
> - Replicators fire asynchronously throughout S phase.
> - Replicator usage can be regulated during development and under different metabolic conditions.
> - Once-per-cell-cycle genome duplication does not require specific replicator sequences.
> - Different replicators have different DNA-sequence requirements; some have no DNA-sequence requirements.
> - Many different features of chromatin structure and function can influence the activity of replicators.
>
> **Initiators**
> - ORC homologues are found in all eukaryotes.
> - ORC occupies replicators *in vivo*.
> - The factors that influence the DNA-binding activity of ORC are highly divergent.

focused to a specific segment of the plasmid and replication was initiated within 2 Kb of this segment. This intriguing result could explain an observed loss of localized initiation in methylation-deficient cell lines[73], as well as the correlation between unmethylated CpG islands and replication origins[74].

Several less well-defined features of chromosomes also influence the usage of replication origins, but in most cases it is not known whether they are upstream or downstream of the initiator–replicator interaction (BOX 3). The proximity of nucleosomes can have both inhibitory[75] and stimulatory[76] effects on the activity of budding yeast replication origins, although, at least in some cases, these effects are known to be downstream of ORC binding[76]. There is evidence that histone modifications can influence the usage of replication origins in budding yeast[77], *D. melanogaster*[67] and *X. laevis*[40]. However, the effects of histone acetylation are not likely to be straightforward, as different early firing origins at the chicken β-globin locus have different patterns of epigenetic histone modifications[78]. Certainly, the fact that several origins are regulated during development[70,79,80] and under different growth conditions[81,82] implies that there are many complex factors that influence origin specification (FIG. 2).

### A preview of the 50th anniversary
What can we expect in the next decade? First and foremost, the initiator–replicator interaction will become appreciated as an interaction with far greater complexity than was previously expected. The concept of the replicator as a specific genetic element will cease to be useful, and will be replaced by that of a context-dependent element akin to our present concept of promoters. New methodologies to map replicators more rapidly and confidently will continue to drive discovery. The structure of metazoan ORC and the mechanisms by which it recognizes DNA will be revealed. Proteins that interact with ORC will be identified and their effects on ORC binding to specific DNA sequences will be evaluated. Comparative analysis of the many different factors that influence ORC–replicator interactions will reveal unique and common features of recognition. Further factors downstream of ORC binding and/or pre-RC assembly will also be found to influence the usage of replication origins. Perhaps most importantly, we will begin to address the biological significance of origin localization. Ultimately, the regulation of replication-origin usage during development and in response to different growth conditions will become integrated into our knowledge of the cell cycle and cell differentiation.

### Concluding remarks
The findings of the past decade (BOX 4) warrant a revision of the original concept of a replicator as it applies to eukaryotic cells. The idea that replication is initiated by the binding of an initiator to a specific DNA-sequence recognition element should be replaced by a 'relaxed replicator' model in which DNA sequence constitutes only one of several features that influence where replication will initiate. DNA topology, DNA methylation, transcriptional activity, ancillary protein chaperones, and long- and short-distance chromatin effects might each be absolutely critical or totally dispensable to origin specification under different conditions. This revision does not detract from the fundamental value and spirit of the replicon model. Jacob and Brenner intended the model to describe bacterial systems, and we should no more expect the replicon model to apply to metazoans than we do the operon model. Replication is clearly initiated by the positive interaction of an initiator with a replicator, but we must come to accept that the replicator can no longer be defined by a single specific DNA sequence. Our present, widely accepted, model for restricting replication to precisely once-per-cell-cycle does not invoke any requirement for specific DNA-sequence elements[30]. So, it is arguably the presence of specific initiation sites that is the greater mystery. In this regard, a highly significant question now is whether the 'relaxed replicator' arose by default or by design. Is the replicator simply any available ORC-binding site, or are origins actively positioned, possibly to coordinate the direction of replication forks with other chromosomal functions[30]? If so, selective pressures might have moulded a different blend of targeting mechanisms for each chromosomal site, and replicators might turn out to be no less complex than transcriptional promoters, where it is now a familiar concept that a common machinery can be targeted by many different mechanisms.

*David M. Gilbert is at the Department of Biochemistry and Molecular Biology, State University of New York Upstate Medical University, 750 East Adams Street, Syracuse, New York 13210, USA.*
*e-mail: gilbertd@upstate.edu*

1. Jacob, F., Brenner, S. & Cuzin, F. On the regulation of DNA replication in bacteria. *Cold Spring Harb. Symp. Quant. Biol.* **28**, 329–348 (1964).
2. Chakraborty, T., Yoshinaga, K., Lother, H. & Messer, W. Purification of the *E. coli* dnaA gene product. *EMBO J.* **1**, 1545–1549 (1982).
3. Bell, S. P. & Stillman, B. ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. *Nature* **357**, 128–134 (1992).
4. Timmis, K., Cabello, F. & Cohen, S. N. Cloning, isolation, and characterization of replication regions of complex plasmid genomes. *Proc. Natl Acad. Sci. USA* **72**, 2242–2246 (1975).
5. Yasuda, S. & Hirota, Y. Cloning and mapping of the replication origin of *Escherichia coli. Proc. Natl Acad. Sci. USA* **74**, 5458–5462 (1977).
6. Fujita, M. Q., Yoshikawa, H. & Ogasawara, N. Structure of the dnaA region of *Pseudomonas putida*: conservation among three bacteria, *Bacillus subtilis*, *Escherichia coli* and *P. putida. Mol. Gen. Genet.* **215**, 381–387 (1989).
7. Taylor, J. H. Rates of chain growth and units of replication in DNA of mammalian chromosomes. *J. Mol. Biol.* **31**, 579–594 (1968).
8. Huberman, J. A. & Riggs, A. D. On the mechanism of DNA replication in mammalian chromosomes. *J. Mol. Biol.* **32**, 327–341 (1968).
9. Stambrook, P. J. & Flickinger, R. A. Changes in chromosomal DNA replication patterns in developing frog embryos. *J. Exp. Zool.* **174**, 101–113 (1970).

10. Callan, H. G. Replication of DNA in the chromosomes of eukaryotes. *Proc. R. Soc. Lond. B* **181**, 19–41 (1972).

11. Blumenthal, A. B., Kriegstein, H. J. & Hogness, D. S. The units of DNA replication in *Drosophila melanogaster* chromosomes. *Cold Spring Harb. Symp. Quant. Biol.* **38**, 205–223 (1974).

12. Taylor, J. H. Increase in DNA replication sites in cells held at the beginning of S phase. *Chromosoma* **62**, 291–300 (1977).

13. Hand, R. Eucaryotic DNA: organization of the genome for replication. *Cell* **15**, 317–325 (1978).

14. Mueller, G. C. & Kajiwara, K. Early- and late-replicating deoxyribonucleic acid complexes in HeLa nuclei. *Biochim. Biophys. Acta* **114**, 108–115 (1966).

15. Balazs, I., Brown, E. H. & Schildkraut, C. L. The temporal order of replication of some DNA cistrons. *Cold Spring Harb. Symp. Quant. Biol.* **38**, 239–245 (1973).

16. Struhl, K., Stinchcomb, D. T., Scherer, S. & Davis, R. W. High-frequency transformation of yeast: autonomous replication of hybrid DNA molecules. *Proc. Natl Acad. Sci. USA* **76**, 1035–1039 (1979).

17. Huberman, J. A., Spotila, L. D., Nawotka, K. A., El-Assouli, S. M. & Davis, L. R. The *in vivo* replication origin of the yeast 2 microns plasmid. *Cell* **51**, 473–481 (1987).

18. Brewer, B. J. & Fangman, W. L. The localization of replication origins on ARS plasmids in *S. cerevisiae*. *Cell* **51**, 463–471 (1987).

19. Huberman, J. A., Zhu, J., Davis, L. R. & Newlon, C. S. Close association of a DNA replication origin and an ARS element on chromosome III of the yeast *Saccharomyces cerevisiae*. *Nucleic Acids Res.* **16**, 6373–6384 (1988).

20. Celniker, S. E., Sweder, K., Srienc, F., Bailey, J. E. & Campbell, J. L. Deletion mutations affecting autonomously replicating sequence ARS1 of *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **4**, 2455–2466 (1984).

21. Marahrens, Y. & Stillman, B. A yeast chromosomal origin of DNA replication defined by multiple functional elements. *Science* **255**, 817–823 (1992).

22. Harland, R. M. & Laskey, R. A. Regulated replication of DNA microinjected into eggs of *Xenopus laevis*. *Cell* **21**, 761–771 (1980).

23. Mechali, M. & Kearsey, S. Lack of specific sequence requirement for DNA replication in *Xenopus* eggs compared with high sequence specificity in yeast. *Cell* **38**, 55–64 (1984).

24. Heintz, N. H. & Hamlin, J. L. An amplified chromosomal sequence that includes the gene for dihydrofolate reductase initiates replication within specific restriction fragments. *Proc. Natl Acad. Sci. USA* **79**, 4083–4087 (1982).

25. Gilbert, D. & Cohen, S. N. Autonomous replication in mouse cells: a correction. *Cell* **56**, 143–144 (1989).

26. Masukata, H., Satoh, H., Obuse, C. & Okazaki, T. Autonomous replication of human chromosomal DNA fragments in human cells. *Mol. Biol. Cell* **4**, 1121–1132 (1993).

27. Krysan, P. J. & Calos, M. P. Replication initiates at multiple locations on an autonomously replicating plasmid in human cells. *Mol. Cell. Biol.* **11**, 1464–1472 (1991).

28. Burhans, W. C., Vassilev, L. T., Caddle, M. S., Heintz, N. H. & DePamphilis, M. L. Identification of an origin of bidirectional DNA replication in mammalian chromosomes. *Cell* **62**, 955–965 (1990).

29. Vaughn, J. P., Dijkwel, P. A. & Hamlin, J. L. Replication initiates in a broad zone in the amplified CHO dihydrofolate reductase domain. *Cell* **61**, 1075–1087 (1990).

30. Gilbert, D. M. Making sense of eukaryotic DNA replication origins. *Science* **294**, 96–100 (2001).

31. Kobayashi, T., Rein, T. & DePamphilis, M. Identification of primary initiation sites for DNA replication in the hamster *DHFR* gene initiation zone. *Mol. Cell. Biol.* **18**, 3266–3277 (1998).

32. Dijkwel, P. A., Wang, S. & Hamlin, J. L. Initiation sites are distributed at frequent intervals in the Chinese hamster dihydrofolate reductase origin of replication but are used with very different efficiencies. *Mol. Cell. Biol.* **22**, 3053–3065 (2002).

33. Hyrien, O. & Mechali, M. Plasmid replication in *Xenopus* eggs and egg extracts: a 2D gel electrophoretic analysis. *Nucleic Acids Res.* **20**, 1463–1469 (1992).

34. Mahbubani, H. M., Paull, T., Elder, J. K. & Blow, J. J. DNA replication initiates at multiple sites on plasmid DNA in *Xenopus* egg extracts. *Nucleic Acids Res.* **20**, 1457–1462 (1992).

35. Gilbert, D. M., Miyazawa, H. & DePamphilis, M. L. Site-specific initiation of DNA replication in *Xenopus* egg extract requires nuclear structure. *Mol. Cell. Biol.* **15**, 2942–2954 (1995).

36. Hyrien, O., Maric, C. & Mechali, M. Transition in specification of embryonic metazoan DNA replication origins. *Science* **270**, 994–997 (1995).

37. Sasaki, T., Sawado, T., Yamaguchi, M. & Shinomiya, T. Specification of regions of DNA replication initiation during embryogenesis in the 65-kilobase *DNApolα-dE2F* locus of *Drosophila melanogaster*. *Mol. Cell. Biol.* **19**, 547–555 (1999).

38. Vashee, S. *et al.* Sequence-independent DNA binding and replication initiation by the human origin recognition complex. *Genes Dev.* **17**, 1894–1908 (2003).

39. Wu, J.-R. & Gilbert, D. M. A distinct G1 step required to specify the Chinese hamster *DHFR* replication origin. *Science* **271**, 1270–1272 (1996).

40. Danis, E. *et al.* Specification of a DNA replication origin by a transcription complex. *Nature Cell Biol.* **6**, 721–730 (2004).

41. Kitsberg, D., Selig, S., Keshet, J. & Cedar, H. Replication structure of the human β-globin gene domain. *Nature* **368**, 588–590 (1993).

42. Aladjem, M. *et al.* Participation of the human β-globin locus control region in initiation of DNA replication. *Science* **270**, 815–819 (1995).

43. Kalejta, R. F. *et al.* Distal sequences, but not ori-β/OBR-1, are essential for initiation of DNA replication in the Chinese hamster *DHFR* origin. *Mol. Cell* **2**, 797–806 (1998).

44. Mesner, L. D., Li, X., Dijkwel, P. A. & Hamlin, J. L. The dihydrofolate reductase origin of replication does not contain any nonredundant genetic elements required for origin activity. *Mol. Cell. Biol.* **23**, 804–814 (2003).

45. Handeli, S., Klar, A., Meuth, M. & Cedar, H. Mapping replication units in animal cells. *Cell* **57**, 909–920 (1989).

46. Aladjem, M. I., Rodewald, L. W., Kolman, J. L. & Wahl, G. M. Genetic dissection of a mammalian replicator in the human β-globin locus. *Science* **281**, 1005–1009 (1998).

47. Paixao, S. *et al.* Modular structure of the human lamin B2 replicator. *Mol. Cell. Biol.* **24**, 2958–2967 (2004).

48. Wang, L. *et al.* The human β-globin replication initiation region consists of two modular independent replicators. *Mol. Cell. Biol.* **24**, 3373–3386 (2004).

49. Liu, G., Malott, M. & Leffak, M. Multiple functional elements comprise a mammalian chromosomal replicator. *Mol. Cell. Biol.* **23**, 1832–1842 (2003).

50. Altman, A. L. & Fanning, E. Defined sequence modules and an architectural element cooperate to promote initiation at an ectopic mammalian chromosomal replication origin. *Mol. Cell. Biol.* **24**, 4138–4150 (2004).

51. Gavin, K., Hidaka, M. & Stillman, B. Conserved initiator proteins in eukaryotes. *Science* **270**, 1667–1671 (1995).

52. Lee, D. G. & Bell, S. P. Architecture of the yeast origin recognition complex bound to origins of DNA replication. *Mol. Cell. Biol.* **17**, 7159–7168 (1997).

53. Chuang, R. Y. & Kelly, T. J. The fission yeast homologue of Orc4p binds to replication origin DNA via multiple AT-hooks. *Proc. Natl Acad. Sci. USA* **96**, 2656–2661 (1999).

54. Remus, D., Beall, E. L. & Botchan, M. R. DNA topology, not DNA sequence, is a critical determinant for *Drosophila* ORC–DNA binding. *EMBO J.* **23**, 897–907 (2004).

55. Aparicio, O. M., Weinstein, D. M. & Bell, S. P. Components and dynamics of DNA replication complexes in *S. cerevisiae*: redistribution of MCM proteins and Cdc45p during S phase. *Cell* **91**, 59–69 (1997).

56. Ogawa, Y., Takahashi, T. & Masukata, H. Association of fission yeast Orp1 and Mcm6 proteins with chromosomal replication origins. *Mol. Cell. Biol.* **19**, 7228–7236 (1999).

57. Bielinsky, A. K. *et al.* Origin recognition complex binding to a metazoan replication origin. *Curr. Biol.* **11**, 1427–1431 (2001).

58. Austin, R. J., Orr-Weaver, T. L. & Bell, S. P. *Drosophila* ORC specifically binds to ACE3, an origin of DNA replication control element. *Genes Dev.* **13**, 2639–2649 (1999).

59. Abdurashidova, G. *et al.* Localization of proteins bound to a replication origin of human DNA along the cell cycle. *EMBO J.* **22**, 4294–4303 (2003).

60. Wyrick, J. J. *et al.* Chromosomal landscape of nucleosome-dependent gene expression and silencing in yeast. *Nature* **402**, 418–421 (1999).

61. Breier, A. M., Chatterji, S. & Cozzarelli, N. R. Prediction of *Saccharomyces cerevisiae* replication origins. *Genome Biol.* **5**, R22 (2004).

62. Keller, C., Ladenburger, E. M., Kremer, M. & Knippers, R. The origin recognition complex marks a replication origin in the human *TOP1* gene promoter. *J. Biol. Chem.* **277**, 31430–31440 (2002).

63. Schepers, A. *et al.* Human origin recognition complex binds to the region of the latent origin of DNA replication of Epstein–Barr virus. *EMBO J.* **20**, 4588–4602 (2001).

64. Royzman, I., Austin, R. J., Bosco, G., Bell, S. P. & Orr-Weaver, T. L. ORC localization in *Drosophila* follicle cells and the effects of mutations in dE2F and dDP. *Genes Dev.* **13**, 827–840 (1999).

65. Bosco, G., Du, W. & Orr-Weaver, T. L. DNA replication control through interaction of E2F-RB and the origin recognition complex. *Nature Cell Biol.* **3**, 289–295 (2001).

66. Beall, E. L., Bell, M., Georlette, D. & Botchan, M. R. Dm-*myb* mutant lethality in *Drosophila* is dependent upon *mip130*: positive and negative regulation of DNA replication. *Genes Dev.* **18**, 1667–1680 (2004).

67. Aggarwal, B. D. & Calvi, B. R. Chromatin regulates origin activity in *Drosophila* follicle cells. *Nature* **430**, 372–376 (2004).

68. Saha, S., Shan, Y., Mesner, L. D. & Hamlin, J. L. The promoter of the Chinese hamster ovary dihydrofolate reductase gene regulates the activity of the local origin and helps define its boundaries. *Genes Dev.* **18**, 397–410 (2004).

69. Muller, M., Lucchini, R. & Sogo, J. M. Replication of yeast rDNA initiates downstream of transcriptionally active genes. *Mol. Cell* **5**, 767–777 (2000).

70. Maric, C., Benard, M. & Pierron, G. Developmentally regulated usage of *Physarum* DNA replication origins. *EMBO Rep.* **4**, 474–478 (2003).

71. Haase, S. B., Heinzel, S. S. & Calos, M. P. Transcription inhibits the replication of autonomously replicating plasmids in human cells. *Mol. Cell. Biol.* **14**, 2516–2524 (1994).

72. Harvey, K. J. & Newport, J. CpG methylation of DNA restricts prereplication complex assembly in *Xenopus* egg extracts. *Mol. Cell. Biol.* **23**, 6769–6779 (2003).

73. Rein, T., Kobayashi, T., Malott, M., Leffak, M. & DePamphilis, M. L. DNA methylation at mammalian replication origins. *J. Biol. Chem.* **274**, 25792–25800 (1999).

74. Delgado, S., Gomez, M., Bird, A. & Antequera, F. Initiation of DNA replication at CpG islands in mammalian chromosomes. *EMBO J.* **17**, 2426–2435 (1998).

75. Simpson, R. T. Nucleosome positioning can affect the function of a *cis*-acting DNA element *in vivo*. *Nature* **343**, 387–389 (1990).

76. Lipford, J. R. & Bell, S. P. Nucleosomes positioned by ORC facilitate the initiation of DNA replication. *Mol. Cell* **7**, 21–30 (2001).

77. Pasero, P., Bensimon, A. & Schwob, E. Single-molecule analysis reveals clustering and epigenetic regulation of replication origins at the yeast rDNA locus. *Genes Dev.* **16**, 2479–2484 (2002).

78. Prioleau, M. N., Gendron, M. C. & Hyrien, O. Replication of the chicken β-globin locus: early-firing origins at the 5′ HS4 insulator and the ρ- and β^A-globin genes show opposite epigenetic modifications. *Mol. Cell. Biol.* **23**, 3536–3549 (2003).

79. Lunyak, V. V., Ezrokhi, M., Smith, H. S. & Gerbi, S. A. Developmental changes in the *Sciara* II/9A initiation zone for DNA replication. *Mol. Cell. Biol.* **22**, 8426–8437 (2002).

80. Zhou, J. *et al.* The origin of a developmentally regulated Igh replicon is located near the border of regulatory domains for Igh replication and expression. *Proc. Natl Acad. Sci. USA* **99**, 13693–13698 (2002).

81. Anglana, M., Apiou, F., Bensimon, A. & Debatisse, M. Dynamics of DNA replication in mammalian somatic cells: nucleotide pool modulates origin choice and interorigin spacing. *Cell* **114**, 385–394 (2003).

82. Norio, P. & Schildkraut, C. L. Plasticity of DNA replication initiation in Epstein–Barr virus episomes. *PLoS Biol.* **2**, E152 (2004).

83. Tsurimoto, T., Melendy, T. & Stillman, B. Sequential initiation of lagging and leading strand synthesis by two different polymerase complexes at the SV40 DNA replication origin. *Nature* **346**, 534–539 (1990).

84. Weinberg, D. H. *et al.* Reconstitution of simian virus 40 DNA replication with purified proteins. *Proc. Natl Acad. Sci. USA* **87**, 8692–8696 (1990).

85. Bergsma, D. J., Olive, D. M., Hartzell, S. W. & Subramanian, K. N. Territorial limits and functional anatomy of the simian virus 40 replication origin. *Proc. Natl Acad. Sci. USA* **79**, 381–385 (1982).

86. Dhar, S. K. *et al.* Replication from oriP of Epstein–Barr virus requires human ORC and is inhibited by geminin. *Cell* **106**, 287–296 (2001).

87. Norio, P., Schildkraut, C. L. & Yates, J. L. Initiation of DNA replication within oriP is dispensable for stable replication of the latent Epstein–Barr virus chromosome after infection of established cell lines. *J. Virol.* **74**, 8563–8574 (2000).

88. Santocanale, C. & Diffley, J. ORC- and Cdc6-dependent complexes at active and inactive chromosomal replication origins in *Saccharomyces cerevisiae*. *EMBO J.* **15**, 6671–6679 (1996).

89. Palacios DeBeer, M. A., Müller, U. & Fox, C. A. Differential DNA affinity specifies roles for the origin recognition complex in budding yeast heterochromatin. *Genes Dev.* **17**, 1817–1822 (2003).

90. Santocanale, C., Sharma, K. & Diffley, J. F. Activation of dormant origins of DNA replication in budding yeast. *Genes Dev.* **13**, 2360–2364 (1999).

91. Vujcic, M., Miller, C. A. & Kowalski, D. Activation of silent replication origins at autonomously replicating sequence elements near the *HML* locus in budding yeast. *Mol. Cell. Biol.* **19**, 6098–6109 (1999).

92. Walter, J. & Newport, J. W. Regulation of replicon size in *Xenopus* egg extracts. *Science* **275**, 993–995 (1997).

93. Beall, E. L. *et al.* Role for a *Drosophila* Myb-containing protein complex in site-specific DNA replication. *Nature* **420**, 833–837 (2002).

94. Ehrenhofer-Murray, A., Gossen, M., Pak, D., Botchan, M. & Rine, J. Separation of origin recognition complex functions by cross-species complementation. *Science* **270**, 1671–1674 (1995).

95. Abdurashidova, G. *et al.* Start sites of bidirectional DNA synthesis at the human lamin B2 origin. *Science* **287**, 2023–2026 (2000).

96. Schaarschmidt, D., Baltin, J., Stehle, I. M., Lipps, H. J. & Knippers, R. An episomal mammalian replicon: sequence-independent binding of the origin recognition complex. *EMBO J.* **23**, 191–201 (2004).

## Online links

**DATABASES**
**The following terms in this article are linked online to:**
Entrez: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi
*DHFR* | dnaA | *HBB* | oriC
Flybase: http://flybase.bio.indiana.edu/
Myb | Rpd3
**Access to this links box is available online.**