Why are we where we are? Understanding replication origins and initiation sites in eukaryotes using ChIP-approaches

Aloys Schepers · Peer Papior

Abstract DNA replication initiates from origins of replication following a strict sequential activation programme and a conserved temporal order of activation. The number of replication initiation sites varies between species, according to the complexity of the genomes, with an average spacing of 100,000 bp. In contrast to yeast genomes, the location and definition of origins in mammalian genomes has been elusive. Historically, mammalian replication initiation sites have been mapped in situ by systematically searching specific genomic loci for sites that preferentially initiated DNA replication, potential origins by start-site mapping and autonomously replicating sequence experiments, and potential ORC and pre-replicative complex (pre-RC) sites by chromatin immunoprecipitation (ChIP) using antibodies for pre-RC proteins. In the past decade, ChIP has become an important method for analyzing protein/DNA interactions. Classically, ChIP is combined with Southern blotting or PCR. Recently, whole genome-ChIP methods have been very successful in unicellular eukaryotes to understand molecular mechanisms coordinating replication initiation and its flexibility in response to environmental changes. However, in mammalian systems, ChIP with pre-RC antibodies has often been challenging and genome-wide studies are scarce. In this review, we will appraise the progress that has been made in understanding replication origin organization using immunoprecipitation of the ORC and Mcm2-7 complexes. A special focus will be on the advantages and disadvantages of genome-wide ChIP-technologies and their potential impact on understanding metazoan replicators.

Abbreviations

ACS	ARS consensus motif
ARS	Autonomously replicating sequence
CDK	Cyclin-dependent kinase
C. elegans	*Caenorhabditis elegans*
ChIP	Chromatin immunoprecipitation
ChIP-Chip	Chromatin immunoprecipitation followed by micoarray analyis
ChIP-seq	Chromatin immunoprecipitation followed by direct high-thoughput sequencing
DDK	Dbf4-dependent kinase
D. melanogaster	*Drosophila melanogaster*
(mod)ENCODE	(model) Encyclopaedia Of DNA Elements

A. Schepers (✉) · P. Papior
Department of Gene Vectors, Helmholtz Zentrum München-German Research Center for Environmental Health, Marchioninistrasse 25,
81377 München, Germany
e-mail: schepers@helmholtz-muenchen.de

EBV	Epstein-Barr virus
FR	Family of repeats (of Epstein-Barr virus)
GINS	Heterotetrameric complex of Sld5; Psf1, Psf2, Psf3
HU	Hydroxyurea
HMG	High-mobility group
M	Mitosis
Mcm	Minichromosome maintenance
ORC	Origin recognition complex
post-RC	Post-replicative complex
pre-IC	Pre-initiation complex
pre-RC	Pre-replicative complex
S. cerevisiae; Sc	*Saccharomyces cerevisiae*
S. pombe; Sp	*Schizosaccharomyces pombe*

Transmitting genetic information from one cell generation to the next requires the accurate duplication of the genome, a process named DNA replication. DNA replication originates at defined starting sites, so-called replication origins. In bacteria, DNA replication initiates at one replication origin, which is recognized and bound by one specified initiator protein. This observation was summarized in the deterministic replicon model, which was proposed nearly 50 years ago and is still the dominating model of the control of DNA replication during the cell cycle. Jacob et al. suggested that a genetically defined *cis*-element, termed replicator, interacts with a specified *trans*-acting regulatory factor, denominated as initiator (Jacob et al. 1963). At least three self-evident activities are involved at replication origins.

(1) Specified proteins, termed initiators, bind replicators.
(2) After binding, the initiator must be transformed into a replication competent form. This can be achieved e.g. by modifying the initiator bound to origins, by associating additional factors, or by a combination of both.
(3) DNA synthesis initiates after an activation step at the replication competent site.

These three features are tightly linked with each other but have their own characteristics. How are replicators defined? This question is still open for higher eukaryotes. How and by which activity are replicators recognized and selected? Recent studies in different organisms have shown that many more potential replicators are present in eukaryotic genomes than are actually bound by DNA replication initiating proteins. It is also obvious that only a selection of replication competent origins initiate replication, whereas many others remain silent and are inactivated by passing replication forks during S phase. Different laboratories have developed their own techniques to address each of these topics. It is the aim of this review to discuss attempts to study replicator/initiator interactions by using antibody-based techniques. This cannot be achieved without touching related topics, such as the way origins might be defined and analyzed and the determination of replication initiation sites.

Eukaryotic replicators

In yeasts

The beauty of the replicon model is its simplicity. Results of initial studies aiming to identify replication origins in eukaryotes have led to the assumption that this model also holds true in higher systems. The first eukaryotic DNA replication origins were identified in *Saccharomyces cerevisiae* (Sc). In this unicellular system, replicators are sequence-defined genetic elements, too. *S. cerevisiae* origins support the extrachromosomal replication of plasmids and are therefore called autonomously replicating sequences (ARS) (Stinchcomb et al. 1979). They contain an ARS-consensus sequence (ACS or A element) and a second functionally conserved element (B1 element) that is essential for the binding of the origin recognition complex (ORC; Bell and Stillman 1992; Marahrens and Stillman 1992; Rao and Stillman 1995; Rowley et al. 1995; Theis and Newlon 1997; Xu et al. 2006). B1 often contains a WTW-motif (W=T or A) 28–20 bp distal to position 1 of the ACS. Mutations within the ACS or B1 modulate the affinity of ScORC origin binding and effect origin strength (Chang et al. 2008). However, it is clear that the consensus ACS is not sufficient for origin formation. Recent analyses suggest that positioned nucleosomes surrounding S. cerevisiae origins are characteristic for active origins and that most origins share a nucleosome-free region (G. Brown and D. MacAlpine, personal communications; Yin et al. 2009).

In another yeast, *Schizosaccharomyces pombe* (Sp), it was also possible to identify sequences that

function as autonomous replicators; however, no sequence-defined ACS-elements have been identified in this organism (Clyne and Kelly 1995; Dubey et al. 1994). In contrast, *S. pombe* elements are characterized by an extended 500–100 bp structure with multiple adenine-thymine (AT)-hook motifs that are required for the binding of SpORC. This feature has been used to bioinformatically identify 384 potential origins (Segurado et al. 2003). Further, the Orc4 subunit of the fission yeast complex exhibits a unique N-terminal extension of an AT-tract DNA-binding motif containing three copies of AT-hook triplets (Chuang and Kelly 1999). Reducing the AT-content of *S. pombe* origins results in delayed replication, indicating that sequence composition regulates origin features in this yeast, too (Wu and Nurse 2009).

In metazoans

In higher multi-cellular eukaryotes, replication origins are more difficult to define because of the size and increased complexity of their genomes. For example, it is estimated that in human cells replication initiates from 30.000 replication origins, a complexity that is difficult to harmonize with the simple replicon model that was proposed for a system with a single replication origin. The existence of genetically and sequence-defined replicators in metazoans seems questionable and other mechanisms are considered for determining ORC DNA-binding such as specified DNA elements (i.e. AT-rich sequences, CpG-islands, promoter regions, dinucleotide repeats, matrix attachment regions), the overall chromatin structure, DNA methylation and topology, the local epigenetic environment and ORC-chaperons that function as auxiliary ORC-targeting factors (Aladjem 2007; Gilbert 2004). We are only beginning to understand, how these features not only contribute to origin definition but also to origin usage. To elucidate how these processes, which are also involved in other nuclear processes, regulate replication initiation, it is necessary to identify and characterize elements in *cis* that serve as potential origins of replication in eukaryotes at a genome-wide level. Furthermore, it is becoming increasingly clear that only a small subset of potential origins is employed with high frequency. However, it is currently unclear whether this is cause or consequence of the observation that especially in higher eukaryotes the number of potential initiation sites is several-fold higher than the number of active replicators. It has been shown for human and hamster cells that ORC has to rebind to origins in each cell cycle and initiation sites are selected anew by an unknown mechanism during G1 at the origin decision point (Dimitrova and Gilbert 1999; Gerhardt et al. 2006; Wu and Gilbert 1996). This raises the question, which criteria, if any exist, determine and regulate initiator/replicator interactions and are crucial for the selection of origins to be activated. Towards this end, we need to know to which sites metazoan initiator complexes bind and how and when those qualify to be finally activated. This requires a genome-wide comparison of potential initiator binding and initiation sites.

The eukaryotic initiator complex

Although the nature of metazoan replicators is much more complex than that of unicellular eukaryotes, the principles of the eukaryotic initiator are conserved from yeast to man. The interaction between replicator and initiator is a tightly cell cycle controlled process that ensures in all systems a remarkably reliable and nearly error-free genome duplication. Origin selection and activation occurs in two steps. The first is characterized by the assembly of pre-replicative complexes (pre-RC; Bell and Dutta 2002; Mendez and Stillman 2003). Pre-RC formation is initiated by the binding of ORC to chromosomal DNA (Fig. 1). ORC provides a dynamic and interactive platform to allow the Cdc6- and Cdt1-dependent association of the Mcm2-7 complex. This completes the process of pre-RC assembly, also called licensing. Pre-RC formation renders chromatin competent for replication and is restricted to the G1 phase of the cell cycle, when cyclin-dependent kinase (CDK)-activity is low (for review: Sivaprasad et al. 2006). The so-called "window of opportunity" for pre-RC formation is closed by the activation of the S phase-promoting CDK and Dbf4-dependent kinase (DDK; Diffley 2004). These kinases initiate the second step of origin activation triggering the recruitment of additional initiator proteins to form the pre-initiation complex. The mechanism of origin activation is reasonably well understood in *S. cerevisiae* but not in higher eukaryotes and is therefore currently subject of investigation. The two protein kinases, CDK and DDK, promote the assembly of the pre-initiation complex
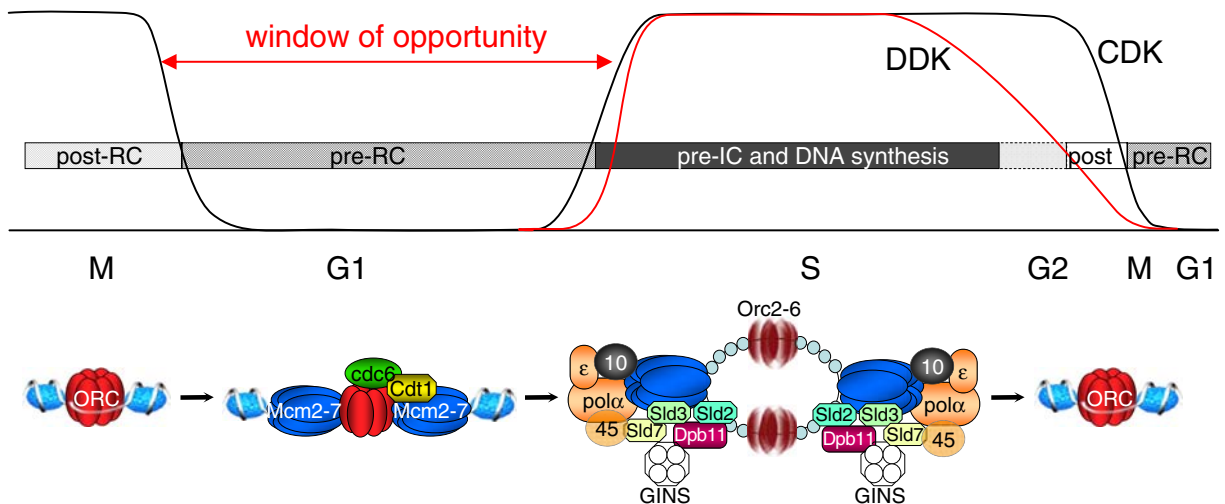
Fig. 1 Model of complex dynamics at eukaryotic replication origins. The origin recognition complex (*ORC*; *red ovals*) is the first step in licensing replication eukaryotic origins by the assembly of pre-replicative complexes (*pre-RCs*). Pre-RCs are assembled in a window of opportunity during the G1 phase of the cell cycle when the activities of cyclin-dependent kinases (*CDK*) and Dbf4-dependent kinases (*DDK*) are low. Cdt1 (*yellow ovals*) and Cdc6 proteins (*green ovals*) are required for the association of the Mcm2-7 protein family (*blue*). The window of opportunity is closed with the rise of CDK activity, but replication initiation can only occur, when DDK-activity is present. In a DDK- and CDK-dependent process, additional complexes bind to origins to form the pre-initiation (*pre-IC*) complex. This includes the Cdc45 (*45*), Mcm10 (*10*), the GINS complex (*Sld5, Psf1, Psf2, Psf3*), Sld2, Sld3, Sld7 and Dpb11 as well as the polymerase α-primase complex (*polα*) and pole (*empty square*). RPA (*light circles*) stabilizes locally unwound single-stranded DNA. In higher eukaryotes, Orc1 is degraded during S phase. ORC is associated with origins throughout the cell cycle (*post-RC*), but is not detected by ChIP at human and *S. pombe* origins after origin firing (Wu and Nurse 2009; Gerhardt et al. 2006). ORC binding rises again in mitosis. The figure is adapted from Kelly and Stillman (2006)

(pre-IC). This process includes the binding of additional proteins such as Mcm10, Cdc45, Sld2, 3 and 7, Dbp11 and the GINS complex and is essential for helicase activation as well as replication initiation at individual origins during S phase (Aparicio et al. 2006, 2009; Labib and Gambus 2007; Pacek et al. 2006; for recent reviews: Remus and Diffley 2009; Teer and Dutta 2006; Walter and Araki 2006).

Principles of identifying and describing origins

(1) A victim of its own success: autonomously replicating sequences?

In the past decades, many laboratories have developed different approaches to identify and characterize eukaryotic origins of DNA replication. Early attempts were stimulated by the successful identification of origins in the different types of yeast by the autonomous replicating sequence (ARS) assay. Although it turned out that the success of this approach is limited due to the poor autonomous

replication activity and stability of the plasmids, many of the up-to-date known mammalian origins have been identified by this technique (i.e. c-*Myc* (Berberich et al. 1995; McWhinney and Leffak 1990); HPRT (Sykes et al. 1988); HSP70 (Taira et al. 1994); mouse ADA (Virta-Pearlman et al. 1993)). The different ARS assays have led to disparate interpretations of the mechanism of initiation of chromosomal DNA synthesis (Burhans et al. 1990; Dijkwel et al. 1991; Handeli et al. 1989; Heintz and Hamlin 1982; Krysan et al. 1989; Ma et al. 1990; McWhinney and Leffak 1990; Vassilev and Johnson 1990). It became rapidly clear that origins of higher eukaryotes differ substantially from the bacterial or yeast paradigms, as they are not defined by sequence. The most successful variation of plasmid-based ARS-assays was introduced by the laboratory of Michelle Calos, who alleviated the problem of plasmid loss by cloning random mammalian DNA fragments into replication-defective Epstein-Barr virus oriP plasmids (Heinzel et al. 1991; Krysan et al. 1989). Replication of wild-type oriP plasmids occurs once per cell cycle in human cells under the control of cellular initiation factors and it depends on two

cis-acting factors. One of them, the family of repeats, is responsible for plasmid maintenance during cell division by tethering plasmids to segregating chromosomes. The results of the above-mentioned milestone papers were clear: Replication was supposed to be dependent on size rather than sequence and was initiated at many different sites, often declared as random or stochastic initiation events (Gilbert 2004). The finding of sequence independent initiation was later confirmed in a different plasmid system (Schaarschmidt et al. 2004). These results precluded the ARS assay as a general technique to identify mammalian replicators. They support a stochastic model in which most mammalian sequences might qualify as potential origins of replication, but these 'weak' origins are not activated in every cell cycle. Because of infrequent origin activation, an excess of replication competent origins has to be established to guarantee genome duplication. It was estimated that around 90% of licensed origins are inefficiently used and remain dormant in any given cell cycle (M. Mechali, personal communication; Blow and Ge 2008). This model is supported by a combined two-step origin-trapping assay that bypasses the size limitation: In a first step, short DNA fragments (<1,000 bp) that bind a human initiator protein (Orc2) had been enriched and were then tested in a functional ARS assay (Gerhardt et al. 2006). This origin-trapping assay turned out to be effective and the ARS-screen eliminated false-positive candidate sequences. However, none of the isolated origin sequences supported long-term maintenance of plasmids. This finding confirms that many human replication origins might not be very efficient and do not fire in every cell cycle. In contrast, longer eukaryotic sequences allow long-term maintenance of plasmids, probably because they contain multiple weak initiation sites (Heinzel et al. 1991).

(2) Identification of initiation start sites by early replication intermediates

The characterization of early replication intermediates is the straightest way to identify active origins of DNA replication. Many attempts are very labour-intensive and have been applied for single initiation sites. They are subject to recent reviews (Aladjem et al. 2006; Hamlin et al. 2008; MacAlpine and Bell 2005) and are the focus of other reviews of this issue (see Prioleau et al.; this issue). For the purpose of this review, we will only briefly summarize genome-wide approaches to describe replication initiation sites.

In yeasts

Genome-wide analyses have at first been published for *S. cerevisiae* and they confirmed that most of the known ARS-elements correlate with active initiation sites (Feng et al. 2006; Raghuraman et al. 2001). However, it was also obvious that not all potential ARS elements are active origins and that only some of them are used in >50% of cell cycles (Newlon and Theis 2002). Microarray analyses of BrdU-labelled newly synthesised DNA of *S. pombe* point in a similar direction as 153 of 460 of potential origins were considered as late and/or inefficient origins (Hayashi et al. 2007). In a parallel study, it was proposed that most *S. pombe* origins fire with low efficiency: approximately 400 potential origins have an efficiency of 10–60% and additional putative 500 origins are used in <10% of cell cycles (Heichinger et al. 2006). In both yeasts, screens for active initiation sites have been performed in the presence of hydroxyurea (HU) that depletes dNTPs resulting in a slow-down of replication forks and subsequently in the activation of normally passively replicated late or dormant origins. This observation implies that many more potential origins are set to 'replication competent' than are routinely required in an unperturbed cell cycle. It is suggested that only 160 origins of *S. pombe* initiate in every DNA replication cycle (Lygeros et al. 2008).

In metazoans

Different approaches to identifying replication intermediates have led to apparently contradictory results in higher eukaryotes. On the one hand, PCR-based techniques, such as nascent strand abundance, or polarity inversion of leading or lagging strands led to the identification of specific initiation sites suggesting that replication initiates at specific elements. This technique detects localized replication initiation events, but does not allow the detection of delocalized initiation events of initiation zones (Fig. 2). On the other hand, the more stringent 2D gel analysis tends to emphasize initiation derived from broad initiation zones, because many replication bubbles originating from adjacent locations are detected by the same

restriction fragment (for detailed discussion, see Aladjem et al. 2006). Two attempts have been made to generate libraries of replication intermediates by trapping replication bubbles and isolating nascent strand DNA (Mesner et al. 2006; Todorovic et al. 2005). However, these studies did not determine initiation events at a genome-wide level to provide the long-awaited answer to the question, whether replication occurs in zones or at specific sites that might be stochastically distributed in metazoan genomes.

Very recently, microarray-based technologies have been used to identify large numbers of replication
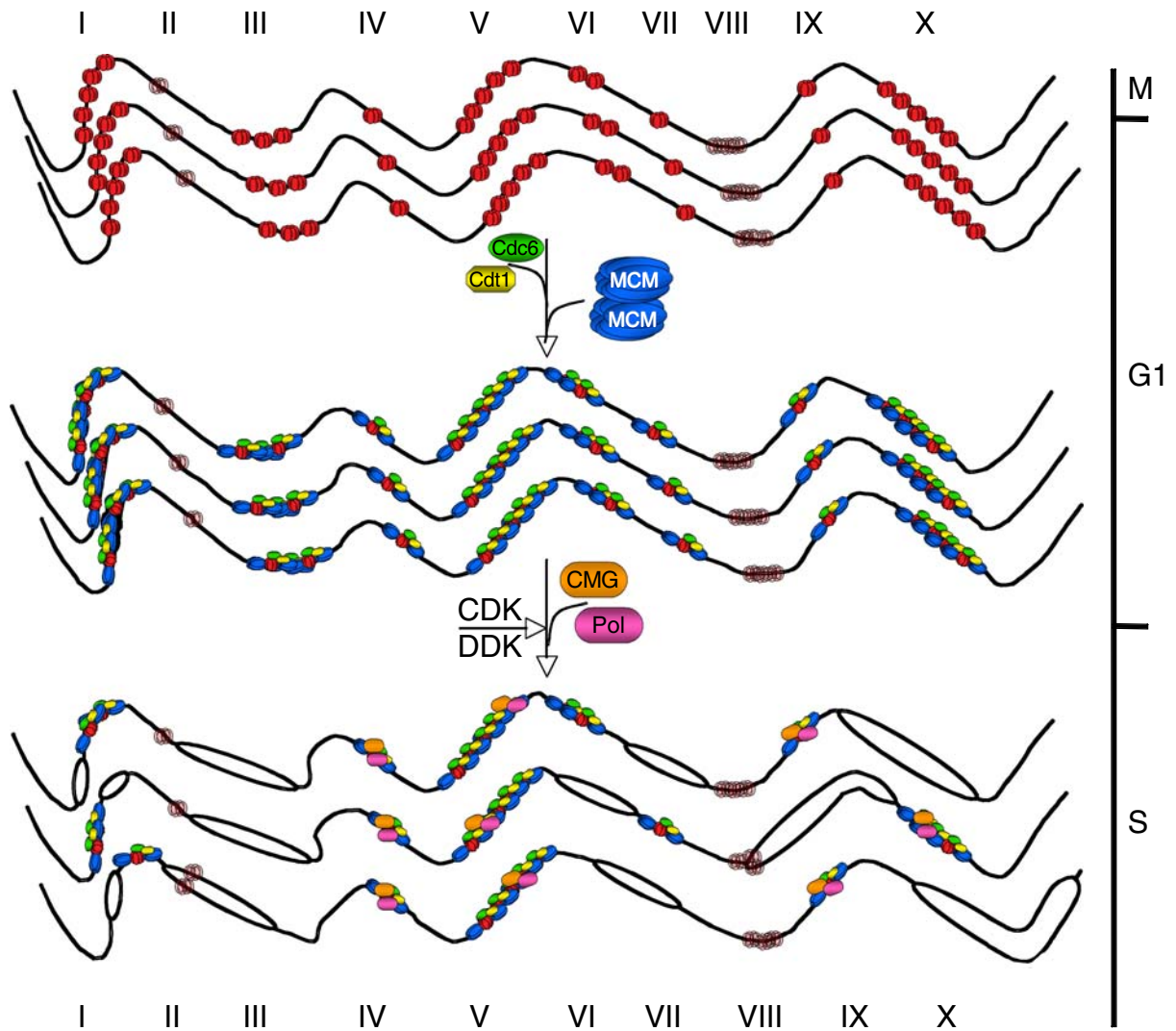
I	II	III	IV	V	VI	VII	VIII	IX	X	
mid	non-ori	early	late	late	early	early	non-ori	early	early	
no	-	yes?	yes	no	yes	yes	-	yes	no	NS
reference	yes	reference	yes	reference	reference	yes	yes	yes	reference	ChIP-PCR
no	yes	no	yes	no	no	yes	yes	yes	no	ChIP-Chip
no	yes	no	yes	no	no	yes	yes	yes	no	ChIP-seq

origins using nascent strand DNA: *Drosophila mela-nogaster* (MacAlpine et al. 2004; Schubeler et al. 2002); human-tiling arrays (Lucas et al. 2007; Watanabe et al. 2002; Woodfine et al. 2004), human ENCODE (ENCyclopedia Of DNA Elements, Cadoret et al. 2008), and a mouse-tiling array (Sequeira-Mendes et al. 2009). All these studies reported a clustering of replication initiation events at or near active promoter regions. In contrast to budding yeast (Raghuraman et al. 2001), microarray-based analyses of replication timing of *D. melanogaster*, human and mouse cells revealed a clear correlation between gene expression and early replication timing (reviewed in MacAlpine and Bell 2005). These data also suggest that origins are preferentially found in GC-rich regions and that very efficient origins are located within CpG-islands (Cadoret et al. 2008; Sequeira-Mendes et al. 2009). A causal correlation between replication initiation sites and specific epigenetic modifications has not been reported so far, and it is very likely that the heterochromatic H3K27me3

◀ **Fig. 2** ORC (*red ovals*) binds to DNA to promote pre-RC formation. ORC does not bind to origins only but might also associate with non-origin sequences (*light red*). Stepwise pre-RC formation is indicated by Cdc6 (*green*) and Cdt1 (*yellow*) that are required for the association of the Mcm2-7 protein family (*blue*). After activation of DDK- and CDK-complexes, the transition from the pre-RC to the pre-initiation complex (*pre-IC*) occurs at G1/S, characterized by the association of the CMG-complex (*orange*; consisting of Cdc45p, Mcm2-7p and GINS) and polymerases (*pink*) (Gerhardt et al. 2006; Moyer et al. 2006). The replication forks are depicted schematically. The scheme depicts different classes of ORC-binding sites. Some of them might not function as origins (*II, VII*). Different classes of origins are schematized: clusters with disperse ORC-sites (*I, III, V, VI, X*) or single sites (*IV, IX*). These potential origins assemble pre-RCs, but not all may be transformed into pre-ICs as is shown by the binding of CMG and polymerases (*pol*). Replication initiation sites can be identified by nascent strand techniques (*NS*); however, if replication initiates at different sites within an initiation zone nascent strand mapping might give unintelligible results. These initiation events can be detected by 2D gel methods. To identify ORC or pre-RC sites by ChIP, they can be paired with different readouts: Southern blot, quantitative and qualitative PCR or in combination with microarrays (*ChIP-Chip*) as well as high throughput sequencing (*ChIP-seq*). The interpretation of ChIP-PCR experiments is dependent on the choice of reference sites, which might be selected by subjective criteria (for a detailed discussion see text). Clusters of ORC and pre-RC sites might be difficult or impossible to detect by ChIP-Chip and ChIP-seq, because the enrichment is too low to allow the detection in relation to non-origin regions. Also, ORC might bind to non-origin regions making the detection of ORC-origin sites even more difficult

mark that shows a predominantly positive correlation with late-replicating domains, does not entail late-replication timing but transcriptional silencing (The_ENCODE_Consortium 2007). The average distance of initiation sites was experimentally determined to 63 kbp (Cadoret et al. 2008) and 103 kbp (Sequeira-Mendes et al. 2009), which is in the range of 100 kbp that originally has been estimated as replicon size (Huberman and Riggs 1968). However, approximately 50% of all origins are less than 60 kbp apart from each other suggesting a high degree of clustering (Cadoret et al. 2008; Sequeira-Mendes et al. 2009). These clusters are mainly observed in gene-rich regions, whereas gene poor regions have interorigin distances up to 500 kbp. A similar broad distribution of initiation sites is also observed in the recently published modENCODE (Model Organism ENCyclopedia of DNA Elements) study of *D. melanogaster* (Celniker et al. 2009).

The major advantage of mapping replication initiation sites on high-density microarrays is the high resolution of the data at a genome-wide or at least chromosome-wide level. One of the disadvantages is that repetitive elements, which i.e. make up more than 50% of the human genome, cannot be analyzed by this technique. This generates a bias towards unique regions of genomes representing the entire euchromatin but only part of the heterochromatin. For this reason, no or only very limited information about the initiation events in heterochromatic regions is available. A second technical disadvantage of microarray-based studies is that at least 2 µg of nascent strand (or other replication intermediates) are required for each hybridization. This requires either an amplification step prior to hybridization, or, for an unbiased analysis, a huge amount of starting material (Cadoret et al. 2008; Sequeira-Mendes et al. 2009). The latter allows relative abundant measurements to compare origin usage, a pre-requisite to determine origin strength. Using an unbiased experimental set up, Sequeira-Mendes et al. demonstrated that CpG islands containing promoter regions comprise the most efficient initiation sites (Sequeira-Mendes et al. 2009).

(3) Analysis of replication origins using antibodies

In order to study the association of specific proteins at defined genomic regions, the chromatin precipitation assay (ChIP) is a powerful tool. The experimental setup was developed in the laboratory of Alexander

Varshavsky aiming to study the association of HSP70 to promoters in *D. melanogaster* (Solomon et al. 1988). Covalent protein/DNA, protein/RNA and protein/protein bonds are introduced by limited cross-linking with formaldehyde at high resolution (2 Å). Chromosomal DNA is extracted and fragmentized using enzymatic activities or sonication. Immunoprecipitation with a specific antibody allows the affinity purification of associated DNA segments. After a DNA purification step, these fragments are analyzed by different techniques, originally by PCR or Southern blotting for a specific sequence. This limitation has been overcome with the development of high-throughput approaches using either micro-arrays (ChIP-Chip) or next generation sequencing (ChIP-seq) as readout, which will be discussed in detail later (Johnson et al. 2007; Wu et al. 2006).

In yeast genomes

After adapting the technique for more general chromatin studies (Orlando and Paro 1993), ChIP became a standard technique in studying the association between pre-RC components and chromatin. The laboratories of S. Bell and K. Nasmyth demonstrated that ORC of *S. cerevisiae* binds to known ARS sequences throughout the cell cycle and that Cdc6, the Mcm2-7 complex as well as Cdc45 bind to these elements in a sequential, cell cycle-regulated order (Aparicio et al. 1997; Tanaka et al. 1997). These studies were fundamental to prove directly the nature of pre-RCs originally observed by in vivo footprinting (Diffley et al. 1994). Microarrays of the *S. cerevisiae* DNA enabled a genome-wide analysis of ORC and Mcm2-7 protein-binding sites (Wyrick et al. 2001; Xu et al. 2006). The comparison of different microarray approaches (ChIP-Chip for location of pre-RC sites; density transfer as well as copy number experiments for replication initiation sites and replication timing) resulted in a comprehensive understanding of *S. cerevisiae* origins (reviewed in MacAlpine and Bell 2005). These data finally verified that most budding yeast origins fire with an average efficiency of once every second cell cycle. High-density tiling array used by Xu et al. allowed the fine mapping of 529 pre-RC sites and the development of criteria to predict *S. cerevisiae* origins (Wyrick et al. 2001; Xu et al. 2006). This in-depth understanding is a prerequisite to study the influence of the local chromatin context of origin

efficiency and replication timing. Manipulating the histone acetylation profile by deleting the histone deacetylase Rpd3L led to an advanced timing of 103 late-replicating origins, suggesting a link between replication timing, gene expression and histone acetylation as well as an impact of local constraints have on origin usage but not on origin definition (Knott et al. 2009).

The binding of *S. pombe* ORC to the SpARS-elements is not dependent on a distinct consensus motif, but is determined by AT-hook motifs present in the unique N-terminal extension of SpOrc4. Takahashi and Masukata used a combination of AT-box mutations of ARS2004 and Orc1- and Orc2-specific antibodies to demonstrate the dependency of SpORC/DNA binding on clustered A/T-stretches (Takahashi and Masukata 2001). ChIP-Chip experiments with pre-RC components showed a strong correlation between ORC and Mcm2-7 protein-binding sites, indicating that the majority of ORC sites serve as pre-RC sites (Feng et al. 2006). With the exception of centromeric and subtelomeric heterochromatin regions, these sites act as initiation sites. This study also demonstrated that ORC, but not Mcm2-7 proteins are bound to centromeres, indicating a replication independent function of ORC also within other nuclear processes. These observations clearly suggest that chromatin immunoprecipitation with pre-RC components and especially with ORC-specific antibodies are not sufficient to identify replication origins. A recent study by Wu and Nurse used ChIP and ChIP-Chip to study the sequential order of replication initiation protein binding in correlation with timing and origin efficiency (Wu and Nurse 2009). They demonstrate that SpORC DNA binding is cell cycle regulated and propose a model of ORC-rebinding and pre-RC formation during M and G1 establish timing and origin efficiency. In particular, Wu and Nurse draw a strong correlation between origin efficiency that is determined by AT-richness and early replication timing. The most efficient origins are activated every second cell cycle, whereas less efficient origins serve in less than every tenth cell cycle. These findings support the Jesuit model, which proposes that many more origins are established than are actually used (DePamphilis 1993).

In metazoan genomes

A specific interaction between ORC of a multi-cellular organism and specific DNA fragments was

first demonstrated by Austin et al., who used a DmOrc2-specific antibody to show that ORC localized to three different elements of the ACE1 and ACE3 chorion elements (Austin et al. 1999). However for several reasons, the use of ChIP in metazoans proved to be very difficult. In order to apply ChIP in combination with PCR or Southern blotting, one needs to know where to screen. This explains why the use of ChIP has mainly been limited to sites that had already been identified as origins or at least as likely candidate sequences. The first ChIP experiment using antibodies against human pre-RC components used oriP, the latent origin of Epstein-Barr virus, as readout sequence (Chaudhuri et al. 2001; Dhar et al. 2001; Ritzi et al. 2003; Schepers et al. 2001). The chromosomal MCM4- and TOP1-origins were the first examples demonstrating that ORC and Mcm2-7 proteins bind to human origins (Keller et al. 2002; Ladenburger et al. 2002; Schaarschmidt et al. 2002). The binding of ORC to human origins by ChIP has been examined at ten-initiation sites (see Table 1). The majority of them had originally been identified by different techniques and ChIP was used to confirm the binding of ORC and other pre-RC components to these sites (Table 1). With the exception of oriP, the

enrichment of pre-RC proteins at specific sites has always been very moderate in comparison to other chromatin-binding proteins. Only three human origins have been identified using a ChIP approach. Ori6 (RDH) was isolated in combination with a transient ARS assay (Gerhardt et al. 2006) and the Orc2 co-precipitated MCM4- and TOP1-origins, which have been confirmed as active origins after subcloning by nascent strand analysis (Keller et al. 2002; Ladenburger et al. 2002).

In principle ChIP experiments allow the identification of potential ORC-binding sites regardless of their use as active origins. The binding of ORC and the assembly of pre-RCs during G1 is a pre requisite for replication competence, but the low number of human origins analyzed and/or isolated by ChIP indicates that this technique is neither an efficient tool nor sufficient to identify and specify metazoan replication initiation sites. Several reasons account for this hurdle. One is that many more origins are established as replication competent sites than are actually activated in any S phase (see above) and another is that ORC does not only have a role in replication initiation, but also in other nuclear processes (Hemerly et al. 2009; Prasanth et al. 2002, 2004). Metazoan ORC binds with low sequence specificity. As a consequence, the chromatin associa-

Table 1 Human origins of replication

Origin	First identified	ChIP	References
DBF4	Yes	Yes	(Romero and Lee 2008)
DNMT1	Plasmid/NS	No	(Araujo et al. 1999)
FMR	NS	Yes	(Gray et al. 2007)
GM-CSF	NS	Yes	(Todorovic et al. 2005)
HPRT	Plasmid/NS	No	(Cohen et al. 2002; Sykes et al. 1988)
HSP70	Plasmid	No	(Taira et al. 1994; Tao et al. 1997)
IFN2A	Plasmid/NS	No	(Kholodii et al. 2008)
HBB	yes	No	(Kamath and Leffak 2001; Kitsberg et al. 1993)
LMNB2	NS	Yes (UV)	(Abdurashidova et al. 2000, 2003)
Mcm4/PRKDC	ChIP/NS[a]	Yes[a]	(Ladenburger et al. 2002)
c-*Myc*	Plasmid	Yes	(Ghosh et al. 2006; Kemp et al. 2005; McWhinney and Leffak 1990)
ori6 (RDH-prom)	Plasmid/ChIP/NS[a]	Yes[a]	(Gerhardt et al. 2006)
ors8	NS (homology to monkey ors8)	Yes	(Callejo et al. 2006)
RPE	NS	No	(Hu et al. 2004)
Top1	ChIP/NS[a]	Yes[a]	(Keller et al. 2002)
oriP	Plasmid/NS	Yes	(Chaudhuri et al. 2001; Dhar et al. 2001; Norio and Schildkraut 2001; Ritzi et al. 2003; Schepers et al. 2001; Yates et al. 1984)

[a] These origins have been identified by ChIP and verified by nascent strand and/or plasmid replication assays

tion of ORC is disperse and the enrichment measured at replication origins is low in comparison to other nuclear factors such as transcriptional activators that exhibit high sequence specificity or have sharp boundaries such as many epigenetic domains. Also, the expression level of ORC influences the specificity of ORC DNA binding, which varies between cell lines and ORC is over-expressed in many tumour cell lines. This might result in 'off-target effects', which interfere with the specificity of ORC DNA binding and blur the picture.

For these reasons, the choice of careful controls is very important to avoid false-positive results. Not only should the specificity of the antibody used and the input be monitored, but also reference sites might be misleading. Clearly, only sites with a localized ORC binding can be detected by ChIP, but regions that are bound by multiple ORC molecules cannot (Fig. 2). Because no 'ORC-free' regions are known so far, the search for negative sites is biased and uncertain. If ORC binding is stochastic, the measured enrichment of a potential ORC-binding site in relation to reference sites is getting arbitrary. Extreme examples of genomes that replicate without any specificity are early embryonic stages of *X. laevis* (Mechali and Kearsey 1984) and other amphibians as well as circular plasmids that contain DNA elements that tether plasmids to segregating plasmids (Krysan and Calos 1991; Schaarschmidt et al. 2004). Nevertheless, some origins have been identified in human cells that show some specificity in ORC binding (Table 1). Beside DNA sequence, a localized binding of ORC can be achieved by different means. Epigenetic features, permissive local chromatin structure, DNA topology, and, ORC-interacting proteins like Ku80, AlF-C, c-*Myc*, Trf2, EBNA1 and the architectural chromatin protein HMGA1a were shown to function as ORC chaperons in targeting ORC to chromatin regions and contributing to origin formation and a more specified binding of ORC (Atanasiu et al. 2006; Dhar et al. 2001; Dominguez-Sola et al. 2007; Minami et al. 2006; Schepers et al. 2001; Sibani et al. 2005; Thomae et al. 2008).

Chromatin-immunoprecipitation and high-throughput approaches of metazoan genomes

Up to date, only very limited data are available using metazoan replication-specific antibodies in ChIP-approaches. In contrast, genome-wide ChIP-approaches in different yeast organisms are state-of-the-art and widely used. Replication initiation sites mapped with microarrays as readouts have been published and are convincing. Why is there a lack of genome-wide chromatin immunoprecipitation experiments with antibodies directed against metazoan replication initiation proteins followed by microarray hybridization or sequencing?

General methodological considerations

Possible explanations are widespread. Of course, the above-discussed arguments regarding the specificity of ORC DNA binding and the abundance of pre-RC sites also apply for ChIP-Chip experiments. In addition, microarray-based ChIP experiments rely on high-quality and specific antibodies. As a general rule, increased complexity and, especially, increasing repetitiveness of the analyzed genomes, enforce the demand of high-quality antibodies that interact with high specificity towards the protein of interest and not with other factors or the chromatin. However, this is a general requirement and not specific for studying ORC and pre-RC sites. Careful controls are required, such as input and mock controls to normalize for biases introduced by cross-reactivity of the antibody and the input chromatin (Park 2008). The optimal length of chromatin for microarray experiments is 200–800 bp, the PCR step usually used to amplify co-precipitated chromatin before the hybridization favours even smaller DNA fragments. The input DNA always over-represents euchromatic DNA, because sonication and enzymatic activities such as micrococcal nuclease are easily accessible for enzymes and require less energy for physical fragmentation than for heterochromatin (Teytelman et al. 2009). Highly compacted heterochromatic DNA is usually either present in longer DNA fragments or remains insoluble and is therefore underrepresented in the ChIP samples.

Consequences for genome-wide analyses using pre-RC-specific antibodies

These technical considerations explain in addition to the DNA binding properties of ORC, why only human origins have been characterized by ChIP that are located within the euchromatin. ORC-binding

sites that have been confirmed by ChIP are situated in promoter regions or in regions that display open chromatin structures. Nucleosome-free regions allow the sequence-independent binding of metazoan ORC and imply that ORC is not only present at specific sites, but also in regions that might be of variable length. As a consequence, the enrichment of ORC ChIP experiments at a specific site is lower and much more disperse (Fig. 2). As mentioned above, the choice of reference sites is more difficult and nearly impossible to judge. This is probably one reason, why the published ORC data in human and other metazoans are relatively poor in comparison to gene expression profiling, methylated DNA immunoprecipitation and other epigenetic studies as well as profiling of microRNA. ChIP experiments of EBV-positive human B cells with different pre-RC-specific antibodies hybridized against the human ENCODE array and custom-made EBV-arrays confirm a broad and unspecific binding of replication initiation proteins to genomes (own unpublished observation). Peaks that indicate the enrichment of a ChIP'ed protein were detectable at low level for individual Orc2, Cdt1 and Mcm2-7 proteins, but they rarely correlated with each other or with replication initiation sites mapped by nascent strand analyses (Cadoret et al. 2008). These experiments also indicated that a sequence-dependent probe normalization procedure is required to avoid artefacts caused by the hybridization properties of different probes (Chung et al. 2007; Chung and Vingron 2009).

Taken these considerations into account, genome-wide studies with pre-RC specific antibodies are relatively easy feasible in different yeasts that have small genomes without large repetitive elements, in which pre-RCs are formed at sequence-defined elements. The genomes of D. melanogaster and C. elegans are 20 times smaller than the human genome, the genome of S. cerevisiae 100 times smaller, therefore less complex and easier to handle in genome-wide approaches. Microarray-based analyses of the yeast genomes have already been discussed above. ORC and Mcm2-7 data of D. melanogaster have recently been published on the basis of the modENCODE microarray and show the first ChIP-data of a metazoan species (Celniker et al. 2009). Using ORC and Mcm2-7-specific antibodies, MacAlpine and co-workers detected approximately 5000 pre-RC binding sites with a typical twofold to fivefold enrichment (D. MacAlpine, personal communication). The abundance is much lower than that of transcription factors that usually display more than tenfold enrichment in comparison to reference sites. The modENCODE data confirm that pre-RCs are often clustered at sites of actively transcribed genes, which are depleted of nucleosomes. These sites have two interesting features: they often co-localize with early replication origins as detected by BrdU incorporation during HU arrest and also display on average three times more ORC than later replicating regions (D. MacAlpine, personal communication). Heterochromatic regions have not been analyzed in this study, because the low probe-density of the array technology does not allow a reliable analysis of these regions. This highlights one of the major drawbacks of the microarray technology that does not allow the analysis of repetitive regions that are present in the heterochromatin. These repetitive elements make up significant parts of metazoans, i.e. more than 50% of the human genome.

ChIP-Chip or ChIP-seq?

Next generation sequencing (ChIP-seq) is the latest development to revolutionize genome-wide studies of ORC-binding, pre-RC assembly and replication initiation sites. Currently, experience with ChIP-seq is cumulating but up to date no objective comparison with ChIP-Chip is possible (for discussion, see Kharchenko et al. 2008; Park 2008). Although first studies in the replication field are pending, We will discuss advantages and disadvantages of this new technique with respect to eukaryotic replication origins. Similar to ChIP-Chip, using this technique the most important factor is high quality antibodies that should have minimal cross-reactivity towards chromatin. This is pre-requisite, because all co-precipitated chromatin is sequenced. One advantage of ChIP-seq is that this method is independent from hybridization artefacts such as dye effects, tiling resolution and the influence of GC-content of the oligoprobe. Another advantage is the fact that ChIP-seq is not dependent on a defined array-design, for example heterochromatic regions that are generally not represented on microarrays. Although ChIP-seq does not also allow the analysis of repetitive elements, unique regions within heterochromatin can be analyzed, e.g. to give some information of replication initiation of this part of the chromatin. The major

disadvantage is that 100–150 bp fragments are preferentially sequenced, which enhances the bias towards small DNA fragments and is the reason for large complexes and structures not being efficiently sequenced (Kharchenko et al. 2008; Teytelman et al. 2009). In analogy to the arguments discussed above, the preference of small fragments enhances the input effect. The fact that ORC binding and pre-RC formation are not sequence-specific in metazoans requires a high depth of sequencing, not only of the DNA co-precipitated with a specific antibody, but also of mock control experiments and input DNA. This increases the cost of each experiment dramatically. It is therefore very likely that ChIP-seq will in the near future only be available for eukaryotic genomes of lower complexity. We can also expect genome-wide analyses that are not dependent on an immunoprecipitation step such as the analysis of initiation sites.

As for the microarray technique a few years ago, one major challenge for ChIP-seq will be the interpretation of data obtained. The bioinformatical challenge for ChIP-seq are greater than for ChIP-Chip and it will be a major issue to find the right balance of sequencing-depth and the precise and reliable identification of ORC-binding regions, pre-RC assembly sites and replication initiation sites. As ChIP-seq will become cheaper, more genome-wide data of the different replication initiation aspects will be generated and in parallel with improved computational evaluation, it may shed some light onto the definition and the selection of replication origins in complex metazoan genomes. One can expect a more complex understanding of replication initiation sites in human cells and an answer to the question, whether the binding of ORC to metazoan genomes is sufficient to define a replication origin. Because ORC seems to display such an unspecific DNA binding and is involved in many other nuclear processes, it is conceivable that other criteria must be defined than the binding of a hard-core replication initiation protein.

References

Abdurashidova G, Deganuto M, Klima R et al (2000) Start sites of bidirectional DNA synthesis at the human lamin B2 origin. Science 287:2023–2026

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

Aladjem MI (2007) Replication in context: dynamic regulation of DNA replication patterns in metazoans. Nat Rev Genet 8:588–600

Aladjem MI, Falaschi A, Kowalski D (2006) Eukaryotic DNA replication origins. Cold Spring Harbor Press, Cold Spring Harbor

Aparicio OM, Weinstein DM, Bell SP (1997) Components and dynamics of DNA replication complexes in S. cerevisiae: redistribution of MCM complexes and Cdc45p during S phase. Cell 91:59–69

Aparicio T, Ibarra A, Mendez J (2006) Cdc45-MCM-GINS, a new power player for DNA replication. Cell Div 1:18

Aparicio T, Guillou E, Coloma J, Montoya G, Mendez J (2009) The human GINS complex associates with Cdc45 and MCM and is essential for DNA replication. Nucleic Acids Res 37:2087–2095

Araujo FD, Knox JD, Ramchandani S et al (1999) Identification of initiation sites for DNA replication in the human dnmt1 (DNA-methyltransferase) locus. J Biol Chem 274:9335–9341

Atanasiu C, Deng Z, Wiedmer A, Norseen J, Lieberman PM (2006) ORC binding to TRF2 stimulates OriP replication. EMBO Rep 7:716–721

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

Bell SP, Dutta A (2002) DNA replication in eukaryotic cells. Ann Rev Biochem 71:333–374

Bell SP, Stillman B (1992) Nucleotide dependent recognition of chromosomal origins of DNA replication by a multi-protein complex. Nature 357:128–134

Berberich S, Trivedi A, Daniel DC, Johnson EM, Leffak M (1995) In vitro replication of plasmids containing human c-*myc* DNA. J Mol Biol 245:92–109

Blow JJ, Ge XQ (2008) Replication forks, chromatin loops and dormant replication origins. Genome Biol 9:244

Burhans WC, Vassilev LT, Caddle MS, Heintz NH, DePamphilis ML (1990) Identification of an origin of bidirectional DNA replication in mammalian chromosomes. Cell 62:955–965

Cadoret JC, 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

Callejo M, Sibani S, Di Paola D, Price GG, Zannis-Hadjopoulos M (2006) Identification and functional analysis of a human homologue of the monkey replication origin ors8. J Cell Biochem 99:1606–1615

Celniker SE, Dillon LA, Gerstein MB et al (2009) Unlocking the secrets of the genome. Nature 459:927–930

Chang F, Theis JF, Miller J et al (2008) Analysis of chromosome III replicators reveals an unusual structure

for the ARS318 silencer origin and a conserved WTW sequence within the origin recognition complex binding site. Mol Cell Biol 28:5071–5081

Chaudhuri B, Xu H, Todorov I, Dutta A, Yates JL (2001) Human DNA replication initiation factors, ORC and MCM, associate with oriP of Epstein-Barr virus. Proc Natl Acad Sci USA 98:10085–10089

Chuang RY, Kelly TJ (1999) The fission yeast homologue of Orc4p binds to replication origin DNA via multiple AT-hooks. Proc Natl Acad Sci USA 96:2656–2661

Chung HR, Vingron M (2009) Comparison of sequence-dependent tiling array normalization approaches. BMC Bioinformatics 10:204

Chung HR, Kostka D, Vingron M (2007) A physical model for tiling array analysis. Bioinformatics 23:i80–i86

Clyne RK, Kelly TJ (1995) Genetic analysis of an ARS element from the fission yeast Schizosaccharomyces pombe. EMBO J 14:6348–6357

Cohen SM, Brylawski BP, Cordeiro-Stone M, Kaufman DG (2002) Mapping of an origin of DNA replication near the transcriptional promoter of the human HPRT gene. J Cell Biochem 85:346–356

DePamphilis ML (1993) Eukaryotic DNA replication: anatomy of an origin. An Rev Biochem 62:29–63

Dhar SK, Yoshida K, Machida Y et al (2001) Replication from oriP of Epstein-Barr virus requires human ORC and is inhibited by geminin. Cell 106:287–296

Diffley JF (2004) Regulation of early events in chromosome replication. Curr Biol 14:R778–R786

Diffley JFX, Cocker JH, Dowell SJ, Rowley A (1994) Two steps in the assembly of complexes at yeast replication origins in vivo. Cell 78:303–316

Dijkwel PA, Vaughn JP, Hamlin JL (1991) Mapping of replication initiation sites in mammalian genomes by two-dimensional gel analysis: stabilization and enrichment of replication intermediates by isolation on the nuclear matrix. Mol Cell Biol 11:3850–3859

Dimitrova DS, Gilbert DM (1999) The spatial position and replication timing of chromosomal domains are both established in early G1 phase. Mol Cell 4:983–993

Dominguez-Sola D, Ying CY, Grandori C et al (2007) Non-transcriptional control of DNA replication by c-myc. Nature 448:445–451

Dubey DD, Zhu J, Carlson DL, Sharma K, Huberman JA (1994) Three ARS elements contribute to the ura4 replication origin region in the fission yeast, Schizosaccharomyces pombe. EMBO J 13:3638–3647

Feng W, Collingwood D, Boeck ME et al (2006) Genomic mapping of single-stranded DNA in hydroxyurea-challenged yeasts identifies origins of replication. Nat Cell Biol 8:148–155

Gerhardt J, Jafar S, Spindler MP, Ott E, Schepers A (2006) Identification of new human origins of DNA replication by an origin-trapping assay. Mol Cell Biol 26:7731–7746

Ghosh M, Kemp M, Liu G et al (2006) Differential binding of replication proteins across the human c-myc replicator. Mol Cell Biol 26:5270–5283

Gilbert DM (2004) In search of the holy replicator. Nat Rev Mol Cell Biol 5:848–855

Gray SJ, Gerhardt J, Doerfler W, Small LE, Fanning E (2007) An origin of DNA replication in the promoter region of

the human fragile X mental retardation (FMR1) gene. Mol Cell Biol 27:426–437

Hamlin JL, Mesner LD, Lar O et al (2008) A revisionist replicon model for higher eukaryotic genomes. J Cell Biochem 105:321–329

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

Hayashi M, Katou Y, Itoh T et al (2007) Genome-wide localization of pre-RC sites and identification of replication origins in fission yeast. EMBO J 26:1327–1339

Heichinger C, Penkett CJ, Bahler J, Nurse P (2006) Genome-wide characterization of fission yeast DNA replication origins. EMBO J 25:5171–5179

Heintz NH, Hamlin JL (1982) An amplified chromosomal sequence that includes the gene for dihydrofolate reductase initiates replication within specific restriction fragments. Proc Natl Acad Sci U S A 79:4083–4087

Heinzel SS, Krysan PJ, Tran CT, Calos MP (1991) Autonomous DNA replication in human cells is affected by the size and the source of the DNA. Mol Biol Cell 11:2263–2272

Hemerly AS, Prasanth SG, Siddiqui K, Stillman B (2009) Orc1 controls centriole and centrosome copy number in human cells. Science 323:789–793

Hu L, Xu X, Valenzuela MS (2004) Initiation sites for human DNA replication at a putative ribulose-5-phosphate 3-epimerase gene. Biochem Biophys Res Commun 320:648–655

Huberman JA, Riggs AD (1968) On the mechanism of DNA replication in mammalian chromosomes. J Mol Biol 32:327–341

Jacob F, Brenner S, Cuzin F (1963) On the regulation of DNA replication in bacteria. Cold Spring Harb Symp Quant Biol 28:329–348

Johnson DS, Mortazavi A, Myers RM, Wold B (2007) Genome-wide mapping of in vivo protein-DNA interactions. Science 316:1497–1502

Kamath S, Leffak M (2001) Multiple sites of replication initiation in the human beta-globin gene locus. Nucleic Acids Res 29:809–817

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

Kelly TJ, Stillman B (2006) Duplication of DNA in eukaryotic cells. In: DePamphilis ML (ed) DNA replication and human disease. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, pp 1–29

Kemp MG, Ghosh M, Liu G, Leffak M (2005) The histone deacetylase inhibitor trichostatin A alters the pattern of DNA replication origin activity in human cells. Nucleic Acids Res 33:325–336

Kharchenko PV, Tolstorukov MY, Park PJ (2008) Design and analysis of ChIP-seq experiments for DNA-binding proteins. Nat Biotechnol 26:1351–1359

Kholodii G, Dantsevich O, Tarantul V (2008) New origins of DNA replication identified near the human IFNA2 gene. Cell Cycle 15:3949–3951

Kitsberg D, Selig S, Keshet I, Cedar H (1993) Replication structure of the human beta-globin gene domain. Nature 366:588–590

Knott SR, Viggiani CJ, Tavare S, Aparicio OM (2009) Genome-wide replication profiles indicate an expansive role for Rpd3L in regulating replication initiation timing or efficiency, and reveal genomic loci of Rpd3 function in *Saccharomyces cerevisiae*. Genes Dev 23:1077–1090

Krysan PJ, Calos MP (1991) Replication initiates at multiple locations on an autonomously replicating plasmid in human cells. Proc Natl Acad Sci USA 11:1464–1472

Krysan PJ, Haase SB, Calos MP (1989) Isolation of human sequences that replicate autonomously in human cells. Mol Biol Cell 9:1026–1033

Labib K, Gambus A (2007) A key role for the GINS complex at DNA replication forks. Trends Cell Biol 17:271–278

Ladenburger EM, Keller C, Knippers R (2002) Identification of a binding region for human origin recognition complex proteins 1 and 2 that coincides with an origin of DNA replication. Mol Cell Biol 22:1036–1048

Lucas I, Palakodeti A, Jiang Y et al (2007) High-throughput mapping of origins of replication in human cells. EMBO Rep 8:770–777

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 C, Leu TH, Hamlin JL (1990) Multiple origins of replication in the dihydrofolate reductase amplicons of a methotrexate-resistant Chinese hamster cell line. Mol Cell Biol 10:1338–1346

MacAlpine DM, Bell SP (2005) A genomic view of eukaryotic DNA replication. Chromosome Res 13:309–326

MacAlpine DM, Rodriguez HK, Bell SP (2004) Coordination of replication and transcription along a *Drosophila* chromosome. Genes Dev 18:3094–3105

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

McWhinney C, Leffak M (1990) Autonomous replication of a DNA fragment containing the chromosomal replication origin of the human c-*myc* gene. Nucleic Acids Res 18:1233–1242

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

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

Minami H, Takahashi J, Suto A, Saitoh Y, Tsutsumi K (2006) Binding of AlF-C, an Orc1-binding transcriptional regulator, enhances replicator activity of the rat aldolase B origin. Mol Cell Biol 26:8770–8780

Moyer SE, Lewis PW, Botchan MR (2006) Isolation of the Cdc45/Mcm2–7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase. Proc Natl Acad Sci USA 103:10236–10241

Newlon CS, Theis JF (2002) DNA replication joins the revolution: whole-genome views of DNA replication in budding yeast. Bioessays 24:300–304

Norio P, Schildkraut CL (2001) Visualization of DNA replication on individual Epstein-Barr virus episomes. Science 294:2361–2364

Orlando V, Paro R (1993) Mapping Polycomb-repressed domains in the bithorax complex using in vivo formaldehyde cross-linked chromatin. Cell 75:1187–1198

Pacek M, Tutter AV, Kubota Y, Takisawa H, Walter JC (2006) Localization of MCM2–7, Cdc45, and GINS to the site of DNA unwinding during eukaryotic DNA replication. Mol Cell 21:581–587

Park PJ (2008) Epigenetics meets next-generation sequencing. Epigenetics 3:318–321

Prasanth SG, Prasanth KV, Stillman B (2002) Orc6 involved in DNA replication, chromosome segregation, and cytokinesis. Science 297:1026–1031

Prasanth SG, Prasanth KV, Siddiqui K, Spector DL, Stillman B (2004) Human Orc2 localizes to centrosomes, centromeres and heterochromatin during chromosome inheritance. EMBO J 23:2651–2663

Raghuraman MK, Winzeler EA, Collingwood D et al (2001) Replication dynamics of the yeast genome. Science 294:115–121

Rao H, Stillman B (1995) The origin recognition complex interacts with a bipartite DNA binding site within yeast replicators. Proc Natl Acad Sci USA 92:2224–2228

Remus D, Diffley JF (2009) Eukaryotic DNA replication control: lock and load, then fire. Curr Opin Cell Biol (in press)

Ritzi M, Tillack K, Gerhardt J et al (2003) Complex protein-DNA dynamics at the latent origin of DNA replication of Epstein-Barr virus. J Cell Sci 116:3971–3984

Romero J, Lee H (2008) Asymmetric bidirectional replication at the human DBF4 origin. Nat Struct Mol Biol 15:722–729

Rowley A, Cocker JH, Harwood J, Diffley JFX (1995) Initiation complex assembly at budding yeast replication origins begins with the recognition of a bipartite sequence by limiting amounts of the initiator, ORC. EMBO J 14:2631–2641

Schaarschmidt D, Ladenburger EM, Keller C, Knippers R (2002) Human Mcm proteins at a replication origin during the G1 to S phase transition. Nucleic Acids Res 30:4176–4185

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

Schepers A, Ritzi M, Bousset K et al (2001) Human origin recognition complex binds to the region of the latent origin of DNA replication of Epstein-Barr virus. EMBO J 20:4588–4602

Schubeler D, Scalzo D, Kooperberg C et al (2002) Genome-wide DNA replication profile for *Drosophila melanogaster*: a link between transcription and replication timing. Nat Genet 32:438–442

Segurado M, de Luis A, Antequera F (2003) Genome-wide distribution of DNA replication origins at A+T-rich islands in *Schizosaccharomyces pombe*. EMBO Rep 4:1048–1053

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

Sibani S, Price GB, Zannis-Hadjopoulos M (2005) Ku80 binds to human replication origins prior to the assembly of the ORC complex. Biochemistry 44:7885–7896

Sivaprasad U, Dutta A, Bell SP (2006) Assembly of pre-replication complexes. In: DePamphilis ML (ed) DNA replication and human disease. Cold Spring Harbor Press, Cold Spring Harbor, pp 63–88

Solomon MJ, Larsen PL, Varshavsky A (1988) Mapping protein-DNA interactions in vivo with formaldehyde: evidence that histone H4 is retained on a highly transcribed gene. Cell 53:937–947

Stinchcomb DT, Struhl K, Davis RW (1979) Isolation and characterisation of a yeast chromosomal replicator. Nature 282:39–43

Sykes RC, Lin D, Hwang SJ, Framson PE, Chinault AC (1988) Yeast ARS function and nuclear matrix association coincide in a short sequence from the human HPRT locus. Mol Gen Genet 212:301–309

Taira T, Iguchi-Ariga SM, Ariga H (1994) A novel DNA replication origin identified in the human heat shock protein 70 gene promoter. Mol Cell Biol 14:6386–6397

Takahashi T, Masukata H (2001) Interaction of fission yeast ORC with essential adenine/thymine stretches in replication origins. Genes Cells 6:837–849

Tanaka T, Knapp D, Nasmyth K (1997) Loading of an Mcm protein onto DNA replication origins is regulated by Cdc6p and CDKs. Cell 90:649–660

Tao L, Nielsen T, Friedlander P, Zannis-Hadjopoulos M, Price G (1997) Differential DNA replication origin activities in human normal skin fibroblast and HeLa cell lines. Methods Mol Biol 273:509–518

Teer JK, Dutta A (2006) Regulation of S phase. Results Probl Cell Differ 42:31–63

Teytelman L, Ozaydin B, Zill O et al (2009) Impact of chromatin structures on DNA processing for genomic analyses. PLoS ONE 4:e6700

The_ENCODE_Consortium (2007) Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature 447:799–816

Theis JF, Newlon CS (1997) The ARS309 chromosomal replicator of Saccharomyces cerevisiae depends on an exceptional ARS consensus sequence. Proc Natl Acad Sci USA 94:10786–10791

Thomae A, Pich D, Spindler MP et al (2008) Interaction between HMGA1a and ORC creates site-specific origins. Proc Natl Acad Sci USA 105:1692–1697

Todorovic V, Giadrossi S, Pelizon C et al (2005) Human origins of DNA replication selected from a library of nascent DNA. Mol Cell 19:567–575

Vassilev L, Johnson EM (1990) An initiation zone of chromosomal DNA replication located upstream of the c-myc gene in proliferating HeLa cells. Mol Cell Biol 10:4899–4904

Virta-Pearlman VJ, Gunaratne PH, Chinault AC (1993) Analysis of a replication initiation sequence from the adenosine deaminase region of the mouse genome. Mol Cell Biol 13:5931–5942

Walter JC, Araki H (2006) Activation of pre-replication complexes. In: DePamphilis ML (ed) DNA replication and human disease. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, pp 89–104

Watanabe Y, Fujiyama A, Ichiba Y et al (2002) Chromosome-wide assessment of replication timing for human chromosomes 11q and 21q: disease-related genes in timing-switch regions. Hum Mol Genet 11:13–21

Woodfine K, Fiegler H, Beare DM et al (2004) Replication timing of the human genome. Hum Mol Genet 13:191–202

Wu JR, Gilbert DM (1996) A distinct G1 step required to specify the Chinese hamster DHFR replication origin. Science 271:1270–1272

Wu PY, Nurse P (2009) Establishing the program of origin firing during S phase in fission yeast. Cell 136:852–864

Wu J, Smith LT, Plass C, Huang TH (2006) ChIP-chip comes of age for genome-wide functional analysis. Cancer Res 66:6899–6902

Wyrick JJ, Aparicio JG, Chen T et al (2001) Genome-wide distribution of ORC and MCM proteins in S. cerevisiae: high-resolution mapping of replication origins. Science 294:2357–2360

Xu W, Aparicio JG, Aparicio OM, Tavare S (2006) Genome-wide mapping of ORC and Mcm2p binding sites on tiling arrays and identification of essential ARS consensus sequences in S. cerevisiae. BMC Genomics 7:276

Yates J, Warren N, Reisman D, Sugden B (1984) A cis-acting element from the Epstein-Barr viral genome that permits stable replication of recombinant plasmids in latently infected cells. Proc Natl Acad Sci USA 81:3806–3810

Yin S, Deng W, Hu L, Kong X (2009) The impact of nucleosome positioning on the organization of replication origins in eukaryotes. Biochem Biophys Res Commun 385:363–368