High-resolution analysis of DNA synthesis start sites and nucleosome architecture at efficient mammalian replication origins

Rodrigo Lombraña^{1,2}, Ricardo Almeida^{1,2},
Isabel Revuelta^{1,2}, Sofia Madeira^{1,2},
Gonzalo Herranz¹, Néstor Saiz^{1,3},
Ugo Bastolla¹ and María Gómez^{1,*}

¹Functional Organization of the Genome Group, Centro de Biología Molecular ‘Severo Ochoa’, Consejo Superior de Investigaciones Científicas/Universidad Autónoma de Madrid (CSIC/UAM), Madrid, Spain

DNA replication origins are poorly characterized genomic regions that are essential to recruit and position the initiation complex to start DNA synthesis. Despite the lack of specific replicator sequences, initiation of replication does not occur at random sites in the mammalian genome. This has led to the view that DNA accessibility could be a major determinant of mammalian origins. Here, we performed a high-resolution analysis of nucleosome architecture and initiation sites along several origins of different genomic location and firing efficiencies. We found that mammalian origins are highly variable in nucleosome conformation and initiation patterns. Strikingly, initiation sites at efficient CpG island-associated origins always occur at positions of high-nucleosome occupancy. Origin recognition complex (ORC) binding sites, however, occur at adjacent but distinct positions marked by labile nucleosomes. We also found that initiation profiles mirror nucleosome architecture, both at endogenous origins and at a transgene in a heterologous system. Our studies provide a unique insight into the relationship between chromatin structure and initiation sites in the mammalian genome that has direct implications for how the replication programme can be accommodated to diverse epigenetic scenarios.

The EMBO Journal (2013) 32, 2631–2644. doi:10.1038/emboj.2013.195; Published online 30 August 2013

Subject Categories: chromatin & transcription; genome stability & dynamics; genomic & computational biology

Keywords: CpG islands; DNA replication initiation; high resolution; nucleosomes; ORC

Introduction

Genome-wide analysis of DNA replication origin (ORI) localisation in various human and mouse cell types have reported

a strong association between ORIs and transcriptional units (Cadoret *et al*, 2008; Sequeira-Mendes *et al*, 2009; Cayrou *et al*, 2011; Martin *et al*, 2011; Mesner *et al*, 2011; Besnard *et al*, 2012). This preference for DNA replication initiation to occur at specific chromosomal regions, the lack of identifiable replicator sequences, and the higher affinity of the origin recognition complex (ORC) for nucleosome-depleted chromatin (Vashee *et al*, 2003; Schaarschmidt *et al*, 2004; Lubelsky *et al*, 2010) suggest an open chromatin structure as a signature of mammalian replication origins. Several previous studies in different model systems have tackled this question but the answer is still open in the field, as they found opposite results in the different genomes analysed. In the yeast *Saccharomyces cerevisiae* or in *Drosophila* cells, for example, ORC binding at active chromosomal replication origins occurs at nucleosome-depleted regions (NDRs) (Eaton *et al*, 2010; MacAlpine *et al*, 2010). On the contrary, in the *Schizosaccharomyces pombe* genome ORC binding sites and NDRs do not overlap (de Castro *et al*, 2011). Moreover, recent work in *Arabidopsis* showed that ORI regions bound by ORC are enriched in nucleosomes (Stroud *et al*, 2012), and bioinformatics analysis of ORI sequences from mouse stem (ES) cells predict a high intrinsic nucleosome occupancy at the ORI centres (Cayrou *et al*, 2012). It is possible that the disparity of results found in the different model systems could indicate that nucleosome's contribution to ORI specification is unique in each genome. However, the resolution of these studies is low because their analysis is based on the averaged data of hundreds of regions. In addition, the precise location of the replication initiation sites within the ORI region has not been addressed in any of the cases.

To investigate the role of nucleosomes in regulating ORI activity in the mammalian genome we undertook a high-resolution analysis of nucleosomal architecture and replication initiation profiling at a selection of ORI regions. The chosen regions were representative of the reported ORI location relative to the genes and comprise different types of ORIs in terms of transcriptional status, associated histone modifications and relative firing efficiencies that we characterized in a previous work (Table I) (Sequeira-Mendes *et al*, 2009). The results revealed the nucleosomal diversity at mammalian ORIs and showed that replication start sites occur at positions of high-nucleosome occupancy. These findings from a limited number of ORIs were validated by genome-wide analysis of available data sets at efficient ORIs both in mouse ES and in HeLa cells. Analysis of replication initiation sites and ORC binding sites showed that they occur at distinct positions within the ORI region, the position of ORC being marked by unstable H3.3/H2A.Z double variant-containing nucleosomes. In addition, we found that ORI architecture depends on nucleosome architecture: replication initiation at higher efficient ORIs occurs at the same position in most cells in the

*Corresponding author. Centro de Biología Molecular ‘Severo Ochoa’, Consejo Superior de Investigaciones Científicas, CSIC/UAM, C/ Nicolás Cabrera 1, 28049 Madrid, Spain. Tel.: +34 911964702;

Fax: +34 91 1964420; E-mail: mgomez@cbm.uam.es

²These authors contributed equally to this work.

³Present address: Faculty of Life Sciences, University of Manchester, Manchester M13 9PT, UK.

Received: 4 March 2013; accepted: 7 August 2013; published online: 30 August 2013

Table 1 Localisation and characteristics of the mouse ORI regions analysed in this study

ORI Location ^a	Transcriptional activity ^b / histone modifications ^b	CGI ^c	Relative firing efficiency ^d
<i>Mecp2</i> promoter	Moderate/ H3K4me3 H3K9,14 Ac	Y	~140
<i>Haus7</i> promoter	Moderate/ H3K4me3 H3K9,14 Ac	N	~60
<i>Vps45</i> promoter	Moderate/ H3K4me3 H3K9,14 Ac	N	~90
<i>Atp2b3</i> promoter	Undetectable/ H3K4me3 H3K27 me3 H3K9,14Ac	Y	~140
<i>Hs6st2</i> promoter	Undetectable/ H3K4me3 H3K27 me3 H3K9,14 Ac	Y	~80
Exon <i>Plas3</i>	NA	N	~40
Exon <i>Syde1</i>	NA	N	~40
Exon <i>Scl7a14</i>	NA	N	~20
LaminB2 transgene	Undetectable/ NA	Y	~60

Abbreviation: NA, not analysed.

^aFrom Sequeira-Mendes *et al* (2009). ORI location was defined within 800 bp resolution by short nascent strands hybridisation on genomic microarrays.

^bFrom Ku *et al* (2008) and Sequeira-Mendes *et al* (2009).

^cCGIs were defined using the more stringent definition of Takai and Jones (2002).

^dRelative abundance of the ORI peaks in 300–800 nt SNS preparations (Supplementary Figures S5 and S6). Values were normalised to those obtained at a non-ORI region.

population whereas less-efficient ORIs showed a more dispersed initiation point within the cell population, coincident with a less strongly positioned nucleosome in those regions. Moreover, variations in nucleosome conformation due to increasing transcriptional activity or by employing a transgene in a heterologous system are accompanied by variations in ORI architecture. Our results indicate a relationship between the initiation of DNA replication and nucleosome positioning that might provide the mammalian genome with the necessary flexibility to accommodate replication initiation to diverse epigenetic programmes.

Results

Peaks of replication initiation profiles at efficient ORIs do not coincide with NDRs in the mouse genome

To determine the relationship between ORIs and nucleosome architecture in the mammalian genome, we first investigated the genome-wide distribution of replication initiation sites and nucleosome positioning around the transcription start sites (TSS) of CpG island (CGI)-associated ORIs in mouse ES cells. Replication initiation sites were identified by array hybridisation of short nascent strands (SNSs), whose abundance peaks close to the start sites of DNA synthesis (Sequeira-Mendes *et al*, 2009; Cayrou *et al*, 2011). We analysed CGI-ORIs as they are the most efficiently activated and more conserved ORIs across all cell types examined, both in mouse and in human systems (Sequeira-Mendes *et al*, 2009; Cayrou *et al*, 2011; Besnard *et al*, 2012). In addition,

CGIs harbour a stereotypical chromatin architecture consisting in an NDR immediately upstream of the TSS flanked by well-positioned nucleosomes (Valouev *et al*, 2011; Fenouil *et al*, 2012; Teif *et al*, 2012), allowing us to directly address the distribution of initiation sites at NDRs. Average profiles of SNS (Cayrou *et al*, 2011) and nucleosome densities (Teif *et al*, 2012) from mouse ES cells showed that the replication initiation sites do not coincide with the NDR (Figure 1A). Instead, the bimodal replication profile peaked at positions that are occupied by nucleosomes (~500 bp upstream and downstream of the TSS, as reported by Cayrou *et al*, 2011). Individual ORIs also showed similar distribution (Supplementary Figure 1), as well as a smaller data set of mouse CGI-ORIs from our previous work (Figure 1B) (Sequeira-Mendes *et al*, 2009). These results are consistent with the computational prediction of high-nucleosome occupancy at the centre of ORI regions in the mouse genome, regardless if they are CGI associated or not (Cayrou *et al*, 2012).

High-resolution analysis of nucleosome positioning and dynamics reveals the diversity of nucleosome patterns at mammalian origins

Genome-wide studies defined ORIs in the range of 1–2 kb, depending on the size of the SNS used for microarray hybridisation or sequencing, a resolution that does not allow us to address precisely the relationship between the initiation sites and individual nucleosomes. Moreover, different chromatin regions have different physical properties that require comprehensive studies including varying MNase

Figure 1 Genome-wide distribution of replication initiation sites and nucleosomes at CGI-ORIs in mouse ES cells. SNS signal strength or nucleosome density was determined by calculating the average number of hits per base in 200 bp windows over 6 kb for 323 highly efficient CGI-ORI regions (A) (data from Cayrou *et al*, 2011) or 4 kb for 36 CGI-ORI regions (B) (data from Sequeira-Mendes *et al*, 2009) centred on the TSS defined as the 5'-end of RefSeq genes. Files were converted to mm9 build using the liftOver tool of the UCSC. Genome-wide nucleosome positioning data were from Teif *et al* (2012). Composite profiles were generated by plotting the mean value in each window transformed into a Z score by subtracting the mean over the whole genome and dividing by the total standard deviation. The standard error of the mean was computed to test that the peaks were significant. Nucleosome profiles (red lines) show both that CGI promoters are depleted from bulk nucleosomes relative to the genome average and that the 200-bp NDR immediately upstream of the TSS do not coincide with the peaks of SNS enrichment (blue lines). Vertical filled lines mark the position of the TSS and dashed lines the maximum enrichment of SNS identifying the most frequent positions of replication initiation sites.

treatments and salt conditions to capture the epigenetic dynamics of chromatin, representing experimental and computational challenges for genome-wide analyses (Tolstorukov *et al*, 2010). Thus, to address the nucleosomal signature of mammalian replication origins representing different chromatin environments, we choose a candidate-based approach and employ a nucleosome-scanning assay (NuSA), as first developed in *S. cerevisiae* (Sekinger *et al*, 2005; Infante *et al*, 2012). This assay couples isolation of DNA after MNase digestion with quantitative real-time PCR (qPCR) using overlapping amplicons, producing a high-resolution, quantitative map of nucleosome location and occupancy (Figure 2A). The position, size and overlap of the amplicons are critical to resolve the location of nucleosomes and to achieve a high-resolution map. In all the regions studied, we used amplicons between 60 and 80 bp in size with a 15- to 25-bp overlap and empirically set up the appropriate qPCR conditions to ensure the specificity and quantitative performance of all primer pairs (Materials and methods; Supplementary Table SI). As different histone octamers show different sensitivities to MNase, we purified mononucleosome-protected DNA from both native and crosslinked chromatin treated with increasing amounts of the enzyme to unveil the different stabilities of individual nucleosomes (Figure 2B). We also used anti-H3 chromatin immunoprecipitation (ChIP) of mononucleosomal DNA derived from cross-linked chromatin followed by NuSA to confirm that peaks of presumed nucleosomal protection from MNase digestion contain histones. This highly detailed customized study allows us to capture some of the nucleosome dynamics at the analysed regions that we interpreted as follows: we consider a DNA region that remains protected on native chromatin under all MNase conditions employed as bound by a stable nucleosome particle, and a region that remains protected only on crosslinked chromatin or under low MNase treatment as bound by a labile histone octamer (Figure 2C).

The complete NuSA results of the eight ORI regions studied are shown in Supplementary Figures S2–S4 and representative nucleosome profiles of the various experimental conditions used are shown in Figure 3 and Supplementary Figure S7. As the ORI regions were previously identified with a resolution of 300–800 bp by SNSs hybridisation on tiled microarrays (Sequeira-Mendes *et al*, 2009), we analysed in detail ~600-bp length centred at each region. To validate our NuSA approach, we first examined ORI regions co-localising with the promoters of expressed genes (Supplementary Figure S2). At *Mecp2* CGI promoter region, we mapped two well-positioned nucleosomes downstream of the TSS that were stable throughout all MNase digestions assayed (Supplementary Figure S2A and B). We also identified two unstable nucleosomes at positions –1 and –2 that could only be detected at low MNase treatments (Supplementary Figure S2A) and by ChIP with anti-H3 antibodies performed in low-salt conditions (Supplementary Figure S2B). The results were similar at the *Haus7* promoter region: we identified a highly stable, but not a strongly positioned nucleosome, at location +1, and a labile nucleosome at location –1 that was only detectable when nuclei were treated with low MNase concentrations and H3-ChIP on crosslinked chromatin (Supplementary Figure S2A–D). At *Vps45* promoter region, we did not find any change in nucleosome occupancy within the experimental conditions used (Supplementary Figure S2A and B). Instead, we detected a constitutively exposed DNA region that could correspond to an NDR surrounding the TSS. To further confirm that this 200-bp region was indeed free of nucleosomes in most cells, we performed NuSA on mononucleosomal DNA derived from crosslinked chromatin and low MNase exposure (see Materials and methods). Under these conditions, we could detect the presence of a highly labile nucleosome exactly occluding the TSS (Figure 2C; Supplementary Figure S2C). Labile nucleosomes, or transiently exposed DNA, suggest

the existence of a continuous process that disrupts nucleosomes and maintains accessibility of cis-regulatory elements (Deal and Henikoff, 2010), consistent with the activity of the transcription machinery and chromatin remodelling complexes at these promoter regions. Our results are in agreement with the idea that promoters at transcriptionally active genes are not nucleosome-free but occupied by highly unstable particles that are preferentially lost under the

conditions normally used in nucleosome preparations for genome-wide studies (Jin *et al*, 2009). Indeed, alignment of our detailed nucleosome maps of individual promoter-ORIs to the corresponding ones reported by Teif *et al* (2012), which were used for the composite patterns shown in Figure 1, showed that several nucleosome particles were missing in the genome-wide data set (Supplementary Figure S5).

We next studied nucleosome configuration at two ORI regions co-localising with the CGI promoters of the *Atp2b3* and *Hs6st2* genes that show undetectable transcriptional activity in mouse ES cells. These two CGIs belong to the category of bivalent promoters because their associated nucleosomes carry both activating (H3K4me3) and repressive (H3K27me3) histone modifications (Table I) (Ku *et al*, 2008). At both ORI regions, the nucleosomal landscape remained unchanged along the experimental conditions used, indicating that these two promoter-ORIs are packaged by stable and well-positioned nucleosomes (Supplementary Figure S3).

Outside of promoter regions nucleosomes are not precisely positioned within the cell population (Valouev *et al*, 2011). This means that a particular position that is occupied by a nucleosome is occupied in some genomes in a population and unoccupied in others. At the ORIs located within the transcribed regions of the genes *Pias3*, *Syde1* and *Scl7a14*, we detected a relatively constant level of nucleosome density across the whole length of the regions, with much lower difference in enrichment between peaks and valleys than those found at promoters (Supplementary Figure S4). These data indicate that these ORI regions are occupied by stable but differently positioned nucleosomes in the cell population (Sekinger *et al*, 2005). Although qPCR profiles were very similar under all experimental conditions, nucleosome patterns were more easily inferred from higher MNase treatments and by anti-H3 mononucleosomal ChIP, supporting the notion that a comprehensive analysis is needed to reveal the variety of the *in vivo* nucleosome landscape at different genomic zones.

Replication initiation sites occur at positions of high-nucleosome occupancy

To generate a profile of replication initiation sites with comparable detail to the nucleosome landscapes at the eight ORI regions, we implemented a high-resolution analysis of SNS abundance by qPCR scanning. SNS purification was performed through size fractionation in sucrose gradients followed by specific enrichment in 5'RNA-primed replication intermediates by lambda-exonuclease digestion. This method has widely been used in mammalian cells to identify replication origins both at individual loci (Keller *et al*, 2002; Prioleau *et al*, 2003; Gmez and Brockdorff, 2004) and at genome-wide by microarray hybridisation or deep sequencing (Cadoret *et al*, 2008; Sequeira-Mendes *et al*, 2009; Cayrou *et al*, 2011; Martin *et al*, 2011; Besnard *et al*, 2012). To

determine the sites of replication initiation with maximum resolution, we improved the current SNS method by employing replication intermediates of small, increasing sizes (100–600, 300–800 and 400–1200 nt) and by scanning each ORI region with the same overlapping amplicons used to map nucleosome positions. This fine analysis allowed us to determine the sites of replication initiation with a resolution of 60–80 bp, below the size of a nucleosome. Since any tiny amount of contamination with small DNA fragments will hamper the detection of the initiation sites, we verified the quality of the SNS preparations by specifically blocking lagging-strand synthesis with emetine (Burhans *et al*, 1991; Abdurashidova *et al*, 2000). In agreement with previous studies, we found that nascent strands elongation was dramatically reduced when mouse ES cells were grown in the presence of emetine, indicating that our SNS preparations mainly contain replication intermediates (Supplementary Figure S6A). Consistently, qPCR measurements at an individual locus showed that replication elongation is impaired when Okazaki fragment synthesis is inhibited (Supplementary Figure S6B). Importantly, the replication initiation sites detected at non-treated and emetine-treated cells pointed to exactly the same genomic positions, indicating that they correspond to true sites of leading-strand synthesis.

The overlap of the replication initiation profiles with the nucleosome maps is shown in Figure 3 and Supplementary Figure S7. SNS abundance values at each ORI region were normalised to those obtained at a flanking primer pair. Due to the proximity of these flanking amplicons to the ORI peaks (between 700 and 1200 bp) and the small size of the SNS used as an input for the qPCRs, maximum SNS enrichments ranged between 3 and 4 times, in agreement with previous reports (Giacca *et al*, 1994; Keller *et al*, 2002; Ladenburger *et al*, 2002; Sequeira-Mendes *et al*, 2009). For the same reason, local SNS enrichments diminished slightly as the length of the SNS increased, as expected if bidirectional DNA synthesis elongates from those initiation sites and the baseline level is set at a proximal region (Figure 3D; Supplementary Figure S7D). However, the enrichments increased when SNS values were normalised to those obtained at a region identified in our previous work that locates far from an ORI and therefore detects background levels of SNS in all sucrose gradient fractions (Supplementary Figures S8 and S9) (Sequeira-Mendes *et al*, 2009). Altogether, these results indicate that the SNS peaks correspond to the sites from which DNA synthesis initiates and elongates to replicate the genome.

Figure 2 Nucleosome-scanning assay (NuSA) provides a quantitative map of nucleosome location and occupancy. (A) Example of the resolution achieved by NuSA. In this particular case, 14 overlapping primer pairs were used to scan a 514-bp region surrounding the major TSS located at the *Hs6st2* CGI promoter. Mononucleosomal DNA was purified from agarose gels after treating mouse ES nuclei with 1200 units of MNase for 6 min at 25C. Full interpretation of NuSA results is described in legend of Supplementary Figure S2. (B) 2.5% agarose gels illustrating examples of native chromatin digested during 6 min at 25C with the indicated units of MNase (left panel) or crosslinked chromatin digested with 2560 U MNase for 6 min at 37C used as an input for anti-H3 ChIP experiments (right panel). U, undigested genomic DNA; m, 1 kb plus DNA ladder. The position of mono-, di- and tri-nucleosomes is depicted at each side. (C) NuSA at the *Vps45* promoter-ORI region on mononucleosomal DNA derived from either crosslinked or native chromatin digested with increasing amounts of MNase at the indicated conditions. Vertical bars at *Vps45* map represent the distribution of CpG dinucleotides along the regions whose abundance is a hallmark of CGIs and that are used here also as a barcode to better resolve nucleosome positioning. Solid and open boxes represent exons and non-translated regions, respectively, and arrows indicate the location and orientation of the major TSS (Carninci *et al*, 2006). Red horizontal bars below each map indicate the length and location of the amplicons used in the qPCRs. Histograms represent nucleosomal DNA enrichments relative to MNase-untreated DNA. qPCRs were performed in duplicates in, at least, two independent mononucleosome preparations. Standard deviation bars are indicated. DNA regions that remain protected on native chromatin under all MNase conditions employed are considered bound by stable nucleosome particles (light grey ovals), and regions that are protected only on crosslinked chromatin or under mild MNase treatment are considered bound by a labile nucleosomes (light red ovals). Nucleosomes were numbered according to their location relative to the TSS.

Figure 3 High-resolution analysis of short replication intermediates abundance and nucleosome configuration at promoter-ORIs. (A) Maps of the promoter-ORI regions analysed and NuSA on mononucleosomal DNA derived from treating native chromatin with 300 U (upper row) or 2400 U (lower row) of MNase. Brackets indicate the position of the CGI, the rest of symbols are as in Figure 2. (B) NuSA on α -H3 immunoprecipitated DNA derived from crosslinked chromatin and digestion with 2560 U MNase 6 min at 37°C (Figure 2B). ChIP experiments were performed at low-salt conditions. (C) Same genomic maps as in (A) showing the nucleosome positioning and occupancies inferred from the data. The fully detailed nucleosome analyses are shown in Supplementary Figures S2 and S3. The horizontal blue line below each map represents the identified ORI region in Sequeira-Mendes *et al* (2009). (D) Histograms represent SNS enrichments relative to a flanking region in preparations of replication intermediates of the indicated sizes. Normalisations to a non-ORI region, what allows inferring the relative ORI usage within the cell population are shown in Supplementary Figure S8. qPCR experiments were performed in duplicates in pools of SNS preparations derived from mouse ES cells. Standard deviation bars are indicated. Note that only 281 bp could be analysed at the *Atp2b3* CGI due to the presence of repetitive DNA at both sides of the TSS (hatched lines).

Strikingly, at the two efficient ORIs associated with the CGIs of the genes *MeCP2* and *Atp2b3* (Figure 3, left two panels) we detected a point of maximum SNS enrichment located at the position where the highly stable +1 nucleosome binds (Figure 3D, upper graphs). These narrow initiation points were also observed when replication intermediates of longer sizes were used as an input for the qPCR measurements (300–800 nt and 400–1200 nt SNSs, medium and bottom graphs, respectively) indicating that, at these two efficient ORIs, DNA synthesis start sites occur at positions bound by well-positioned nucleosomes. In addition, the correlation between tight nucleosome locations, high firing efficiencies and sharp replication profiles—even when longer SNS was interrogated—suggest that, in the cells in which these ORIs fire, DNA replication starts at similar position in all genomes.

At promoter-ORI regions that fire with slightly lower efficiency (Table I), we detected two distinct replication initiation sites that, notably, also located at nucleosome-protected DNA in all three cases (Figure 3, three rightmost panels). At *Vps45* 5'-end, the initiation points localised within the -1 and +2 nucleosomes (SNS peaks separated by 290 bp), whereas at *Haus7* and *Hs6st2*, replication initiation sites mapped at the -1 and +1 nucleosomes (SNS peaks separated by 250 and 180 bp, respectively) (Figure 3D). Interestingly, at *Vps45* and *Haus7* ORIs one of the detected initiation points locate at a dynamically exposed DNA region, while all the other initiation points mapped at regions packed by stable nucleosomes. This suggests that is the positioning, and not the occupancy at a given time, what influences ORI architecture. Importantly, these high-resolution maps of individual efficient ORIs recapitulate and refine the patterns found on

aggregate building (Figure 1). These bimodal replication initiation profiles peaking upstream and downstream of the TSS of promoter-ORIs have been previously described by SNS-microarray hybridisation experiments and might correspond to different initiation points in different cells of the population, or to dual initiation events at two opposite start sites of leading-strand synthesis (Cayrou *et al*, 2011; Figure 1; Supplementary Figure S1). Bioinformatics analysis of a collection of mouse ORIs identified G-rich motifs prone to G-quadruplex formation upstream of each SNS peak, and the authors suggest that G-quadruplexes might play a role in coordinating replication and transcription of the associated genes (Cayrou *et al*, 2012).

We finally addressed replication initiation profiles at low-efficient ORIs not associated with promoter regions (Supplementary Figure S7). At the three regions studied, the peaks of SNS abundance were much less defined, suggesting higher variability at the initiation sites within the cell population than at promoter-ORIs. In addition, the replication profiles flattened rapidly as the length of the SNS increased, supporting the interpretation that the initiation sites at these ORI regions are dispersed across an ~600-bp interrogated. This cell-to-cell variability correlates well with the low nucleosome positioning packing these genomic regions (Supplementary Figure S4), although maximum SNS enrichments tend to co-localise with positions of higher nucleosome occupancy (Supplementary Figure S7). It should be noted that a region covered by loosely positioned nucleosomes means that its nucleosomal array is slightly different in every genome of the cell population (Sekinger *et al*, 2005). These results are reminiscent of the situation reported at clustered initiation sites, such as those constituting the 50-kb initiation zone of the DHFR locus in Chinese hamster cells (Dijkwel and Hamlin, 1995; Dijkwel *et al*, 2000). Recent analysis of the frequency of initiation sites along the locus in individual DNA fibres showed distinct replication initiation patterns in each cell broadly distributed within the entire zone. Notably, this initiation zone is packaged by loosely positioned nucleosomes coincident with scattered ORC binding along the region (Lubelsky *et al*, 2010).

We have previously reported that ORI activation during the S-phase of the cell cycle is accompanied by the overproduction of short abortive replication intermediates derived from nucleosome-free regions surrounding the TSS of CGIs in human cells (Gmez and Antequera, 2008). Due to their overrepresentation on genomic DNA preparations, these short DNA fragments are unveiled only by qPCR when primer set efficiencies are normalised using the cloned version of a locus instead of genomic DNA to construct the standard curves. To check whether local overreplication occurs at the mouse ORIs analysed here and, if so, whether its presence might affect the ORI maps, we re-analysed all qPCR measurements using cloned versions of each locus to construct the standard curves. We were unable to detect local overreplication at the mouse ORI regions analysed, indicating that the replication profiles reported in Figures 2 and 3 correspond to *bona fide* sites of DNA synthesis initiation and elongation rather than to the presence of short overabundant replication intermediates. In contrast, we found short overrepresented DNA fragments derived from the well-characterised human LaminB2 ORI at its previously reported location (Supplementary Figure S10) (Gmez and Antequera,

2008). The reason for these differences between human and mouse CGI-ORIs is currently unknown.

Replication initiation patterns reflect nucleosomal architecture at mammalian ORIs

Our data indicate that replication initiation sites occur at positions of high-nucleosome occupancy and suggest that there is a correlation between stronger nucleosome positioning and higher origin efficiency. To further study the relationship between nucleosomal architecture and ORI architecture, we analysed two scenarios where nucleosome conformation could be modulated locally.

First, we explored whether nucleosomal changes associated with variations in transcriptional activity were accompanied by variations in replication initiation profiles. We analysed the nucleosomal conformation and replication start sites at *Haus7* and *Vps45* ORI regions in ES cells and NIH3T3 fibroblasts, where their associated mRNAs are four times more abundant (Figure 4A). In the case of *Haus7* ORI region, NuSA analysis after increasing MNase exposure identified in both cell lines a labile nucleosome at location -1 and a highly stable, but loosely positioned nucleosome at location $+1$ (Figure 4B, left panels). Interestingly, the positioning of the $+1$ nucleosome was markedly different in both cell types: while the foremost 3'-end of the protected region was indistinguishable between ES and NIH3T3 cells, its 5'-end was clearly shifted in NIH3T3 (light-blue histograms), leaving a longer linker region between the -1 and $+1$ nucleosomes and exposing the TSS (Figure 4B and C). Most importantly, the replication initiation peak located at the $+1$ position was higher and sharper in NIH3T3 cells than in ES cells, suggesting a narrower distribution of this initiation site within the 3T3 cell population, a finding that is correlated with the narrower positioning of this nucleosome in this cell type (Figure 4D). Similar relationship between nucleosome positioning and replication initiation profiles was found at *Vps45* ORI region (Figure 4D, central panels). In this case, the sharp replication initiation site located at the -1 nucleosome position in ES cells (dark-blue lines) was much broader in NIH3T3 cells, correlated with a shifted and less well-positioned location of this nucleosome in these cells (light-blue histograms and lines).

Notably, the observed changes in the replication initiation patterns were not cell type specific; rather they were attributable to the different nucleosomal architecture of the studied regions in both cell types. We found that the replication profile at the ORI region associated with the promoter of the housekeeping *Mecp2* gene was indistinguishable between ES cells and NIH3T3 fibroblasts, paralleling in both cell lines the corresponding well-positioned nucleosomal array (Figure 4B–D, right panels). These results argue in favour of a link between nucleosome positioning and replication initiation patterns in the mammalian genome.

As a second approach, we developed a stable mouse ES cell line carrying a single copy of the best-characterised human ORI, and assayed the nucleosome packaging and replication initiation profile of the transgene in this heterologous system (see Materials and methods). This ORI, named LaminB2 ORI, maps close to the CGI associated with the promoter of the gene *TIMM13* and has been thoroughly studied along the years including: (i) the replication initiation site at single-nucleotide resolution (Abdurashidova *et al*, 2000), (ii) the

Figure 4 Nucleosomal landscape and replication initiation profiling at *Haus7*, *Vps45* and *Mecp2* ORI regions in ES and NIH3T3 cells. (A) RT-PCR analysis of *Haus7*, *Vps45* and *Mecp2* mRNA levels relative to *Hprt* mRNA levels in mouse ES cells (dark-blue histograms) and fibroblasts NIH3T3 (light-blue histograms). Total RNA was extracted with a guanidium isothiocyanate solution and cDNA was synthesised using oligo-dT and SuperScript II Reverse Transcriptase (Invitrogen). The amount of cDNA synthesised from 2 μ g of total RNA (– and + SuperScript II) was used as a template for each PCR, using the following primer pairs: *HprtA*: 5'-gcctaagatgagcgcaagt-3' and *HprtB*: 5'-gtgggaaaatacagccaacact-3'; *HausA*: 5'-acatccattccagtagtcc-3' and *HausB*: 5'-gcgtgaaccagttgtagtc-3'; *VpsA*: 5'-ggtcgaaggaaacaattgc-3' and *VpsB*: 5'-tcaattcgccaccatgaatg-3'; *MecpA*: 5'-atcattagggtccaaggagg-3' and *MecpB*: 5'-ctgaaggttgacacgaaag-3'. (B) NuSA on mononucleosomal DNA derived from native (*Haus7* and *Mecp2*) or crosslinked chromatin (*Vps45*) and digestion with the indicated amounts of MNase in ES and NIH3T3 cells. Data for ES cells are the same as in Figure 2 and Supplementary Figure S2. (C) Inferred nucleosomal patterns in both cell types. Nucleosomes showing a different positioning at each cell line are marked by colour-coded triangles. Note that the +1 nucleosome at *Mecp2* ORI region shows a higher occupancy in NIH3T3 than in ES cells, but its positioning is identical in both cell types. (D) SNS enrichments relative to a flanking region in preparations of replication intermediates of 100–600 nt. qPCR experiments were performed in duplicates in pools of SNS preparations derived from mouse ES cells (dark-blue lines) or NIH3T3 cells (light-blue lines). Standard deviation bars are indicated.

ORC binding site by footprinting protection and ChIP (Ladenburger *et al*, 2002; Abdurashidova *et al*, 2003), and (iii) the 1.2-kb zone capable of driving replication initiation when inserted at ectopic sites in human cells (Paixao *et al*, 2004). A detailed map of the LaminB2 ORI region with the above-mentioned features and the transgenic construct generated are shown in Supplementary Figure S11A.

First, we performed NuSA on mononucleosomal DNA purified after digestion with increasing MNase concentrations and quantify SNS abundance on human HEK293T cells. Primer pairs were carefully chosen and tested to verify their specificity for human DNA. For this reason, the amplicons interrogating the conserved coding regions are slightly longer and, therefore, the exact positioning and occupancy of the most 3' nucleosome is less precise (Supplementary Table S1). At the experimental conditions employed, two stable, well-

positioned nucleosomes (+1 and –1), and one labile nucleosome (–2), were identified along 826 bp at the endogenous LaminB2 ORI in human cells (Figure 5A and B). Of note, linker DNA at this human CGI (around 190 bp) was longer than those found at mouse CGI regions (35 bp on average), leaving open the possibility that some highly unstable nucleosome particles could transiently reside in these regions. In any case, NuSA on mouse nuclei clearly showed that the transgene recapitulates the nucleosome packaging of the endogenous locus, as the three mapped histone octamers occupy exactly the same locations in both cell types (Figure 5B and E). Nucleosome dynamics, however, differ in the mouse background: the position of the –2 labile nucleosome was occupied by a highly stable histone octamer in the transgene (Figure 5D and E). Most likely, this results from a lower propensity of this nucleosome to be either

Figure 5 Nucleosomal landscape and replication initiation profiling at the human LaminB2 ORI in its endogenous position and when knocked-in in mouse ES cells. **(A)** Map of the Lamin B2 ORI region and NuSA on mononucleosomal DNA derived from human HEK293T cells isolated after incubating nuclei with increasing concentrations of MNase. Blue double-headed arrow indicates the position where the 5'-end of the leading strands was mapped (Abdurashidova *et al*, 2000) and the blue horizontal line shows the ORC2 bound region identified by ChIP (Ladenburger *et al*, 2002). A more detailed map of the known features of this ORI region is shown in Supplementary Figure S11A. Rest of the symbols are as in Figure 2. **(B)** Inferred nucleosome pattern in human cells. **(C)** SNS enrichments normalised to those detected by the flanking 3'primer pair in preparations of replication intermediates of the indicated sizes derived from human cells. **(D)** NuSA analysis at the LaminB2 transgene in mouse ES cells. **(E)** Nucleosome landscape at the LaminB2 transgene in mouse ES cells **(F)** Same analysis as in **(C)** on SNS derived from mouse ES cells carrying the transgenic LaminB2 construct. qPCRs were performed in duplicates in two independent mononucleosome or SNS preparations. Standard deviation bars are indicated.

disrupted or moved by chromatin-remodelling enzymes in mouse cells, in agreement with the lack of *TIMM13* transcript detected at the transgene (Supplementary Figure S11B).

These results are consistent with the overall landscape of nucleosome locations and relative occupancy being dictated in part by intrinsic DNA sequence preferences of the

nucleosomes themselves, and by the action of nucleosome-remodelling complexes and competition between nucleosomes and sequence-specific DNA-binding proteins (Deal and Henikoff, 2010).

Interestingly, previously identified sites of ORC binding and replication initiation map exactly at the DNA sequence protected by the -2 unstable nucleosome in human cells (Figure 5B) (Abdurashidova *et al*, 2000, 2003; Ladenburger *et al*, 2002). Consistent with this, high-resolution qPCR scan analysis of SNS derived from human cells detected a replication initiation site precisely at the region where the 5'-end of the leading strand was previously mapped by a different method (Abdurashidova *et al*, 2000) (Figure 5C), reinforcing the accuracy of our SNS qPCR-scan technique. Along with this reported replication start site located at the position where the -2 nucleosome binds, we also identified another two initiation points that mapped at the $+1$ and -1 nucleosome-bound DNA positions. These results seem to suggest that the sites of DNA synthesis at this ORI are less confined than previously thought. It should be noted that these regions were not tested in the detailed initiation point studies of Abdurashidova *et al* (2000), but the initiation site located at the -1 protected position (Figure 5C) is immediately adjacent to a region that undergoes several rounds of synthesis during S-phase, generating overabundant replicated short DNA fragments of unknown function (Gómez and Antequera, 2008; Supplementary Figure S10). Therefore, replication initiation sites also coincide with regions of high-nucleosome occupancy at the efficient LaminB2 ORI in human cells. Strikingly, the replication profile was qualitatively but not quantitatively conserved at the LaminB2 transgene in mouse cells: three replication initiation sites mapping at the same nucleosome-bound DNA positions were detected (Figure 5F). However, relative SNS enrichments vary at the transgene. The peak located at the $+1$ nucleosome position increased and the one located at the -2 position diminished, coincident with a reduction in the accessibility of this DNA segment in mouse cells. These results indicate that this 2.2 kb human DNA fragment is capable to recruit mouse replication complexes and to direct DNA synthesis to start from sites of high-nucleosome occupancy. They also support our interpretation, together with those shown in Figure 4, that replication initiation patterns mirror nucleosome positioning and occupancies. Altogether, our findings fit with the idea that mammalian ORIs are highly flexible, a property that enhances the robustness of the replication process (Mechali, 2010; Cayrou *et al*, 2011; Sequeira-Mendes and Gómez, 2012).

Positioned nucleosomes at initiation sites are removed upon ORI activation

As mammalian ORIs are relatively inefficient—each given ORI in a cell is estimated to fire with $<30\%$ efficiency (Cayrou *et al*, 2011; Gilbert, 2012)—the most likely interpretation of our findings is that nucleosomes are preferentially positioned at ORIs and removed when they become activated. To test this possibility directly, we analysed nucleosome dynamics at the early-firing LaminB2 ORI in synchronised HeLa cells upon ORI activation (Figure 6). Cells were synchronised in G1/S by double thymidine block and mononucleosomes derived from cross-linked chromatin were purified from identical number of cells before and immediately after block release (0 h-sample and

45 min-sample, Figure 6A). At these time points, DNA synthesis was undetectable both by FACS analysis and by BrdU immunofluorescence (Figure 6A). Similarly, only background levels of SNS derived from the LaminB2 region were detected by qPCR in G1/S-blocked cells or early-S cells. Interestingly, NuSA analysis on synchronised cells revealed an increased DNA accessibility at the region bound by ORC and the -2 nucleosome in early-S cells (light-blue histograms) relative to G1/S-blocked cells (dark-blue histograms) (Figure 6B). It is worth mentioning that a prominent MNase protection exactly mapping at the ORC binding site was detected at synchronised cultures and not at exponentially growing HeLa cells (Figure 6C). This protection was mainly detected at low MNase exposure and decreases faster in early-S cells than in G1/S-blocked cells at increasing MNase digestions, suggesting that it might correspond to the previously reported footprinting signal detected at the same position in G1/S cells likely reflecting the recruitment of the pre-initiation complex (marked by red triangles in Figure 6B and Abdurashidova *et al*, 2003). This result is in line with recent findings that pre-RCs at the Epstein-Barr virus genome are characterised by a dynamic MNase pattern, which exhibits an increased sensitivity during S-phase (Papior *et al*, 2012). Altogether, these data indicate that nucleosomes located at initiation sites are specifically remodelled just before bidirectional DNA synthesis starts (Figure 6D).

H3.3/H2A.Z double-variant nucleosomes mark ORC binding sites at efficient ORIs

The finding that ORC is bound to a dynamically exposed DNA region at the LaminB2 ORI in 293T and HeLa cells (Figures 5 and 6) suggests that a similar scenario can also occur at other sites packed by labile nucleosomes. To address this, we investigated the genome-wide distribution of ORC1 binding sites recently generated for HeLa cells (Dellino *et al*, 2013), and of nucleosome particles containing the labile histone variants H3.3/H2A.Z in the same cell type (Jin *et al*, 2009). Labile nucleosome particles, specially the double-variant H3.3/H2A.Z, are easily disrupted when isolated at high-salt concentrations (Jin and Felsenfeld, 2007; Jin *et al*, 2009). Therefore, we analyse the distribution of both the labile variants only detectable at low-salt preparations and H2A.Z-containing nucleosomes isolated at higher salt, which reproduce previous reports of H2A.Z depletion at NDRs (Ozsolak *et al*, 2007; Schones *et al*, 2008). Composite profiles revealed that ORC1 is preferentially located immediately upstream of the TSS, exactly at the position occupied by the H3.3/H2A.Z double variant-containing nucleosomes that mark the NDRs defined by preparative methods used normally in studying nucleosome structure (Figure 7A).

We then analysed the genome-wide distribution of ORC1 and SNS derived from HeLa cells (Besnard *et al*, 2012). Despite the low overlap between the two data sets, the promoter regions that are positive for ORC1 and enriched in SNS (30% of the ORC1-positive sites) showed a clearly distinct localisation of the ORC and the initiation sites (Figure 7B). Similar to the pattern found in mouse ES cells (Figure 1), the biphasic replication profile of HeLa cells peaked at both sides of the TSS, at regions of high-nucleosome occupancy. Altogether, these results depict a complex architecture at efficient mammalian ORIs: labile histone

Figure 6 Nucleosome dynamics at the LaminB2 ORI upon S-phase entry. (A) HeLa cells arrested in early S-phase by a double-thymidine block (0 h sample) were released in thymidine-free medium, and their progress through S-phase was monitored by FACS analysis (upper panel) and BrdU immunofluorescence (lower panel). The profile of an exponential asynchronous culture is also shown (Exp). Black arrowheads point to the time points analysed in (B). (B) NuSA analysis on mononucleosomal DNA derived from crosslinked chromatin digested for 6 min at 37C with the indicated MNase concentrations on equivalent number of cells blocked at G1/S (0 h, dark-blue histograms) or early-S (45' after block release, light-blue histograms). The dark-red triangles above the map represent the precise localisation of the ORC2, ORC1 and CDC6 proteins in G1-phase (Abdurashidova *et al*, 2003). (C) Same analysis as in (B) for exponentially growing HeLa cells. qPCRs were performed in duplicates in two independent mononucleosome preparations. Standard deviation bars are indicated. (D) Diagrammatic interpretation of the results.

particles, whose instability likely facilitates ORC binding at those sites, and surrounding positioned nucleosomes where DNA synthesis starts.

Discussion

To identify some of the cis-acting factors that specify ORI usage in mammalian cells, we conducted a high-resolution analysis of nucleosome architecture at several examples representing diverse chromatin environments and firing efficiencies. The results revealed the nucleosomal diversity at mammalian ORIs and showed that replication start sites occur at positions of high-nucleosome occupancy. These findings from a limited number of ORIs were validated by genome-wide analysis of available data sets at efficient ORIs both in mouse ES and in HeLa cells (Figures 1 and 7). Strikingly, analysis of replication initiation sites and ORC

binding sites from HeLa cells showed that they occur at distinct positions within the ORI region: ORC occupies the classically referred NDR upstream of the TSS of active promoters (exactly at the position marked by unstable nucleosomes containing H3.3/H2A.Z histone variants) while replication initiation sites are maximal at both sides of the TSS, coinciding with sites of positioned nucleosomes. These findings suggest that, at least for the most efficient ORIs, mammalian cells are not so different from *S. cerevisiae* or *Drosophila* cells, where ORC localises to NDRs in the genome (Berbenetz *et al*, 2010; Eaton *et al*, 2010; MacAlpine *et al*, 2010). The relationship between ORC and dynamic histone variants described here in mammalian cells recapitulates previous findings in *Drosophila* cells (Deal *et al*, 2010; Eaton *et al*, 2010; MacAlpine *et al*, 2010) and likely represents a conserved property of metazoan origins. However, recent genome-wide data in *S. pombe* showed that

Figure 7 Distribution of H3.3/H2A.Z double-variant nucleosomes and ORC binding sites at efficient ORIs in HeLa cells. **(A)** Enrichment of labile histone variant-containing nucleosomes isolated either in low salt (H2A.Z, H3.3, and sequential immunoprecipitation of H3.3/H2A.Z) or in high salt (H2A.Z) (Jin *et al*, 2009), and ORC1 (Dellino *et al*, 2013). Reads were averaged in non-overlapping 50 bp windows centred at 3310 ORC-containing TSS. Labile H3.3/H2A.Z nucleosome particles (low salt) and ORC1 are enriched at the NDR, revealed by the profile of the H2A.Z-containing nucleosomes isolated in high salt (dark-blue lines). **(B)** Same analysis as in **(A)** centred at 1004 TSS positive for ORC1 (Dellino *et al*, 2013) and SNS enrichment (Besnard *et al*, 2012). Composite patterns were generated as described in Figure 1. Dashed vertical grey lines mark the TSS position.

ORC binding at ORI regions do not overlap with NDRs (de Castro *et al*, 2011), and enrichment of nucleosomes in ORI zones bound by ORC has also been described in the *Arabidopsis* genome (Stroud *et al*, 2012). These findings suggest that ORC might bind nucleosomes and this binding might be specially relevant to drive replication initiation at certain types of ORIs, such as those not located at promoter regions in mammalian cells (Supplementary Figure S7), or at those located in heterochromatic portions of the genome (Bartke *et al*, 2010), or at *orc1*Δ-sensitive replication origins in *S. cerevisiae* (Muller *et al*, 2010).

Quite likely, the typical nucleosome configuration of active promoters that invariably contain an NDR transiently occupied by labile nucleosomes could account for the reported higher efficiency of CGI-associated ORIs (Sequeira-Mendes *et al*, 2009; Cayrou *et al*, 2011; Besnard *et al*, 2012). Unstable H3.3/H2A.Z-containing nucleosomes could prevent regulatory regions from being occupied by adjacent stable nucleosomes or non-specific factors, as it might occur if the region was completely free of nucleosomes (Jin *et al*, 2009). At the same time, being unstable particles they could be more easily displaced by the replication machinery, as shown here upon LaminB2 ORI firing (Figure 6). This scenario fits well with the finding that nucleosome turnover rates are higher at regions of ORC binding compared to its surrounding sequences in *Drosophila* cells (Deal *et al*, 2010). Interestingly, ORC1 distribution genome-wide is strongly associated with TSS of coding or non-coding RNAs, and higher transcription levels at ORC sites correlate with earlier replication timing (Dellino *et al*, 2013). This suggests that enhanced transcription (i.e., higher enrichment of H3.3/H2A.Z at the NDR, Jin *et al*, 2009) might favour ORC1 recruitment to DNA sites that are transiently available at G1-phase and thus increase ORI firing efficiency. In agreement with this, in our previous work we found that promoter-ORIs are significantly enriched in CAGE tags derived from early embryos relative to the rest of promoters (Sequeira-Mendes *et al*, 2009). This scenario is consistent with the view that mammalian ORIs cannot be defined as unique genomic features. Instead, we propose that DNA synthesis initiates opportunistically from sites of favourable

chromatin configuration, enhancing the robustness of the replication programme (Sequeira-Mendes and Gómez, 2012).

Recent high-resolution analysis of Okazaki fragment distribution in the *S. cerevisiae* genome demonstrated that fragment ligation junctions occur at the nucleosome dyad (Smith and Whitehouse, 2012). Moreover, the authors showed that disrupting chromatin assembly, or lagging-strand processivity, affects both the size and the distribution of Okazaki fragments, suggesting a role of nascent chromatin in the termination of Okazaki fragment synthesis. It is tempting to speculate that our finding that DNA synthesis start sites occur at nucleosome-protected regions might reflect a link between newly deposited nucleosomes and leading-strand synthesis in mammalian cells. Our results showing that nucleosome-positioning variations at regions that are differentially expressed in two cell types result in parallel changes in their corresponding replication profiles support such hypothesis (Figures 4 and 5). Coupling histone repositioning to replication initiation at promoter-ORIs could provide a way to change the chromatin structure that could promote a switch during cell differentiation and development, as well as to accommodate the replication programme to change epigenetic scenarios.

Materials and methods

Cell culture

The mouse embryonic stem cell line PGK12.1 was grown as described previously (Sequeira-Mendes *et al*, 2009). For lagging-strand synthesis inhibition, exponentially growing cells were incubated for 1 h on media containing 2 μ M emetine (Sigma) before DNA extraction or nuclei purification (Burhans *et al*, 1991). All human cell lines were cultivated as recommended in the ATCC. The 129/1 XY ES cell line carrying a floxed PGKneo cassette inserted 5.5 kb upstream of the TSS of the *Xist* gene on the X chromosome was a gift from Dr Tatyana Nesterova and cells were grown as described. The floxed PGKneo cassette inserted at this site did not affect X-inactivation choice in heterozygous female mice (Nesterova *et al*, 2003), indicating that insertions at this genomic region would not affect the regulation of neighbouring genes.

Nucleosome-scanning assay

Nuclei purification and MNase treatment were performed as previously described (Gong *et al*, 1996). MNase-treated chromatin and

purified DNA samples were electrophoretically separated on 2.5% agarose gels, and mononucleosome-sized fragments (140 bp) were excised from the gel and purified (Promega). The resulting material was analysed by sets of overlapping primer pairs, each generating amplicons of 60–80 bp with 15–25 bp overlap between neighbouring amplicons.

Chromatin immunoprecipitation

Formaldehyde-crosslinked chromatin was fragmented by extensive MNase digestion using the following modifications. 1×10^7 purified nuclei were incubated with 2560 U MNase (Fermentas) for 6 min at 37°C and reactions were stopped by shifting samples to 4°C and adding 18 µl of 500 mM EDTA and 7 µl of 200 mM EGTA. The slurry was then purified four times through a 20G needle and four times through a 25G needle. In all, 50 µg of chromatin was immunoprecipitated with 4 µg of polyclonal anti H3 antibody (Abcam) as described previously (Sequeira-Mendes *et al*, 2009), or using a low-salt buffer during both the antibody incubation and the washes (10 mM Tris-HCl pH 7.4; 15 mM NaCl; 0.2 mM EDTA) (Jin and Felsenfeld, 2007).

SNS purification

Genomic DNA isolation and nascent strands fractionation were performed as described previously (Gómez and Antequera, 2008). Sucrose gradient fractions containing replication intermediates ranging between 100–600 nt, 300–800 nt and 400–1200 nt in size were subjected twice to digestion by lambda-exonuclease, which degrades contaminating nicked DNA leaving untouched replication intermediates that are protected by a 5'RNA primer.

qPCR

Quantitative real-time PCR was performed with an ABI Prism 7900HT Detection System (Applied Biosystems), with HotStar Taq polymerase (Qiagen) and SYBR Green (Molecular Probes). Four serial five-fold dilutions of sonicated genomic DNA or the corresponding cloned DNA fragments were amplified using the same reaction mixture as the samples to construct the standard curves. qPCR conditions were empirically tested for each primer pair and only those showing an slope of $-3.3 \pm 10\%$ and R^2 value > 0.99 were accepted. Primer sequences, amplicon sizes and amplification reactions are indicated in Supplementary Table S1. All real-time PCRs were performed in duplicate and in at least two independent preparations of mononucleosomal DNA, SNS or immunoprecipitated material. Quantitative analyses were carried out using the ABI Prism 7900HT SDS Software (version 2.4).

Generation of knock-in LaminB2 ES cells

A 2.2-kb human DNA fragment containing the LaminB2 ORI was amplified from human genomic DNA with the primers LB1: 5'-CTGCACGCGGCAGTTCTTTG-3' and LB2: 5'-TACCCTACCCAAGGACGGG-3' (positions 2366165–2368445, NCBI build 37.3) and cloned in a pBS plasmid flanked by two loxP sites in the same orientation. 129/1 XY ES cells line carrying a floxed PGKneo cassette

was first transfected with a plasmid encoding a tamoxifen-inducible Cre recombinase (pCAGMerCreMer, gift from Dr Mitsuhiro Endo) and stable integrants were obtained by standard puromycin selection. Clones showing appropriate excision of the neo cassette after Cre-mediated recombination were subject of a second round of transfection with the LoxPLaminB2LoxP construct. Neo-resistant selected recombinant clones were extensively characterised by Southern blot analysis and PCR amplification to verify correct and single copy integration of the human LaminB2 ORI.

Cell synchronisation and immunofluorescence with anti-BrdU antibodies

A double-thymidine block was performed by growing HeLa cells for 16 h in DMEM containing 2.5 mM thymidine followed by 9 h of incubation in fresh medium and further incubation for 16 h in the presence of 2.5 mM thymidine. Progression through S-phase was checked at 45 min and then every 1.5 h after block removal by FACS analysis. Synchronous entry in S-phase was checked by immunofluorescence with anti-BrdU antibodies (BD) on parallel cultures of cells grown on slides labelled during 15 min with 10 µM BrdU prior to cell harvest (Leonhardt *et al*, 1992).

Supplementary data

Supplementary data are available at *The EMBO Journal* Online (<http://www.embojournal.org>).

Acknowledgements

We thank Tatyana Nesterova for cell lines and continuous advice on their manipulation, and T Nesterova and Dionisio Martín-Zanca for plasmids. We are also grateful to Francisco Antequera, Pieter Dijkwel and Crisanto Gutierrez for critical review of the manuscript; Lorena Benito for technical assistance; Pepe Belio for artwork; Jean-Marc Lemaître, Joana Sequeira-Mendes and members of MG laboratory for discussions; and the editor and three anonymous referees for excellent constructive criticism. This work was supported by the Spanish Ministry of Economy and Competitiveness (BFU2010-18992) to MG and (BFU2012-40020) to UB. RA is supported by a grant from the Portuguese Foundation for Science and Technology (SFRH/BD/81027/11).

Author contributions: RL, RA, IR and SM designed the experiments and carried out the experimental work with the help of GH and NS. UB performed the computational analyses assisted by RL. MG designed the experiments, supervised the general strategy of the work and wrote the article. All authors analysed the data, discussed the results and approved the final version of the manuscript.

Conflict of interest

The authors declare that they have no conflict of interest.

References

- Abdurashidova G, Deganuto M, Klima R, Riva S, Biamonti G, Giacca M, Falaschi A (2000) Start sites of bidirectional DNA synthesis at the human lamin B2 origin. *Science* **287**: 2023–2026
- Abdurashidova G, Danailov MB, Ochem A, Triolo G, Djeliovic V, Radulescu S, Vindigni A, Riva S, Falaschi A (2003) Localization of proteins bound to a replication origin of human DNA along the cell cycle. *EMBO J* **22**: 4294–4303
- Bartke T, Vermeulen M, Xhemalce B, Robson SC, Mann M, Kouzarides T (2010) Nucleosome-interacting proteins regulated by DNA and histone methylation. *Cell* **143**: 470–484
- Berbenetz NM, Nislow C, Brown GW (2010) Diversity of eukaryotic DNA replication origins revealed by genome-wide analysis of chromatin structure. *PLoS Genet* **6**: e1001092
- Besnard E, Babled A, Lapasset L, Milhavet O, Parrinello H, Dantec C, Marin JM, Lemaître JM (2012) Unraveling cell type-specific and reprogrammable human replication origin signatures associated with G-quadruplex consensus motifs. *Nat Struct Mol Biol* **19**: 837–844
- Burhans WC, Vassilev LT, Wu J, Sogo JM, Nallaseth FS, DePamphilis ML (1991) Emetine allows identification of origins of mammalian DNA replication by imbalanced DNA synthesis, not through conservative nucleosome segregation. *EMBO J* **10**: 4351–4360
- Cadoret JC, Meisch F, Hassan-Zadeh V, Luyten I, Guillet C, Duret L, Quesneville H, Prioleau MN (2008) Genome-wide studies highlight indirect links between human replication origins and gene regulation. *Proc Natl Acad Sci* **105**: 15837–15842
- Carninci P, Sandelin A, Lenhard B, Katayama S, Shimokawa K, Ponjavic J, Semple CA, Taylor MS, Engstrom PG, Frith MC, Forrest AR, Alkema WB, Tan SL, Plessy C, Kodzius R, Ravasi T, Kasukawa T, Fukuda S, Kanamori-Katayama M, Kitazume Y *et al*. (2006) Genome-wide analysis of mammalian promoter architecture and evolution. *Nat Genet* **38**: 626–635
- Cayrou C, Coulombe P, Vigneron A, Stanojic S, Ganier O, Peiffer I, Rivals E, Puy A, Laurent-Chabalier S, Desprat R, Mechali M (2011) Genome-scale analysis of metazoan replication origins reveals their organization in specific but flexible sites defined by conserved features. *Genome Res* **21**: 1438–1449

- Cayrou C, Coulombe P, Puy A, Rialle S, Kaplan N, Segal E, Mechali M (2012) New insights into replication origin characteristics in metazoans. *Cell Cycle* **11**: 658–667
- de Castro E, Soriano I, Marin L, Serrano R, Quintales L, Antequera F (2011) Nucleosomal organization of replication origins and meiotic recombination hotspots in fission yeast. *EMBO J* **31**: 124–137
- Deal RB, Henikoff JG, Henikoff S (2010) Genome-wide kinetics of nucleosome turnover determined by metabolic labeling of histones. *Science* **328**: 1161–1164
- Deal RB, Henikoff S (2010) Capturing the dynamic epigenome. *Genome Biol* **11**: 218–225
- Dellino GI, Cittaro D, Piccioni R, Luzi L, Banfi S, Segalla S, Cesaroni M, Mendoza-Maldonado R, Giacca M, Pelici PG (2013) Genome-wide mapping of human DNA-replication origins: levels of transcription at ORC1 sites regulate origin selection and replication timing. *Genome Res* **23**: 1–11
- Dijkwel P, Hamlin JL (1995) The Chinese hamster dihydrofolate reductase origin consists of multiple potential nascent-strand start sites. *Mol Cell Biol* **15**: 3023–3031
- Dijkwel PA, Mesner LD, Levenson VV, d'Anna J, Hamlin JL (2000) Dispersive initiation of replication in the Chinese hamster rhodopsin locus. *Exp Cell Res* **256**: 150–157
- Eaton ML, Galani K, Kang S, Bell SP, MacAlpine DM (2010) Conserved nucleosome positioning defines replication origins. *Genes Dev* **24**: 748–753
- Fenouil R, Cauchy P, Koch F, Descostes N, Cabeza JZ, Innocenti C, Ferrier P, Spicuglia S, Gut M, Gut I, Andrau JC (2012) CpG islands and CG content dictate nucleosome depletion in a transcription-independent manner at mammalian promoters. *Genome Res* **22**: 2399–2408
- Giacca M, Zentilin L, Norio P, Diviacco S, Dimitrova D, Contreas G, Biamonti G, Perini G, Weighardt F, Riva S, Falaschi A (1994) Fine mapping of a replication origin of human DNA. *Proc Natl Acad Sci USA* **91**: 7119–7123
- Gilbert DM (2010) Evaluating genome-scale approaches to eukaryotic DNA replication. *Nat Rev Genet* **11**: 673–684
- Gilbert DM (2012) Replication origins run (ultra) deep. *Nat Struct Mol Biol* **19**: 740–742
- Gómez M, Antequera F (2008) Overreplication of short DNA regions during S phase in human cells. *Genes Dev* **22**: 375–385
- Gómez M, Brockdorff N (2004) Heterochromatin on the inactive X chromosome delays replication timing without affecting origin usage. *Proc Natl Acad Sci USA* **101**: 6923–6928
- Gong QH, McDowell JC, Dean A (1996) Essential role of NF-E2 in remodeling of chromatin structure and transcriptional activation of the epsilon-globin gene *in vivo* by 5' hypersensitive site 2 of the beta-globin locus control region. *Mol Cell Biol* **16**: 6055–6064
- Infante JJ, Law GL, Young ET (2012) Analysis of nucleosome positioning using a nucleosome-scanning assay. *Methods Mol Biol* **833**: 63–87
- Jin C, Felsenfeld G (2007) Nucleosome stability mediated by histone variants H3.3 and H2A.Z. *Genes Dev* **21**: 1519–1529
- Jin C, Zang C, Wei G, Cui K, Peng W, Zhao K, Felsenfeld G (2009) H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. *Nat Genet* **41**: 941–945
- 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
- Ku M, Koche RP, Rheinbay E, Mendenhall EM, Endoh M, Mikkelsen TS, Presser A, Nusbaum C, Xie X, Chi AS, Adli M, Kasif S, Ptaszek LM, Cowan CA, Lander ES, Koseki H, Bernstein BE (2008) Genome-wide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. *PLoS Genet* **4**: e1000242
- 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
- Leonhardt H, Page AW, Weier HU, Bestor TH (1992) A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell* **71**: 865–873
- Lubelsky Y, Sasaki T, Kuipers MA, Lucas I, Le Beau MM, Carignon S, Debatisse M, Prinz JA, Dennis JH, Gilbert DM (2010) Pre-replication complex proteins assemble at regions of low nucleosome occupancy within the Chinese hamster dihydrofolate reductase initiation zone. *Nucleic Acids Res* **39**: 3141–3155
- MacAlpine HK, Gordan R, Powell SK, Hartemink AJ, MacAlpine DM (2010) Drosophila ORC localizes to open chromatin and marks sites of cohesin complex loading. *Genome Res* **20**: 201–211
- Martin MM, Ryan M, Kim R, Zakas AL, Fu H, Lin CM, Reinhold WC, Davis SR, Bilke S, Liu H, Doroshow JH, Reimers MA, Valenzuela MS, Pommier Y, Meltzer PS, Aladjem MI (2011) Genome-wide depletion of replication initiation events in highly transcribed regions. *Genome Res* **21**: 1822–1832
- Mechali M (2010) Eukaryotic DNA replication origins: many choices for appropriate answers. *Nat Rev Mol Cell Biol* **11**: 728–738
- Mesner LD, Valsakumar V, Karnani N, Dutta A, Hamlin JL, Bekiranov S (2011) Bubble-chip analysis of human origin distributions demonstrates on a genomic scale significant clustering into zones and significant association with transcription. *Genome Res* **21**: 377–389
- Muller P, Park S, Shor E, Huebert DJ, Warren CL, Ansari AZ, Weinreich M, Eaton ML, MacAlpine DM, Fox CA (2010) The conserved bromo-adjacent homology domain of yeast Orc1 functions in the selection of DNA replication origins within chromatin. *Genes Dev* **24**: 1418–1433
- Nesterova TB, Johnston CM, Appanah R, Newall AE, Godwin J, Alexiou M, Brockdorff N (2003) Skewing X chromosome choice by modulating sense transcription across the Xist locus. *Genes Dev* **17**: 2177–2190
- Ozsolak F, Song JS, Liu XS, Fisher DE (2007) High-throughput mapping of the chromatin structure of human promoters. *Nat Biotechnol* **25**: 244–248
- Paixao S, Colaluca IN, Cubells M, Peverali FA, Destro A, Giadrossi S, Giacca M, Falaschi A, Riva S, Biamonti G (2004) Modular structure of the human lamin B2 replicator. *Mol Cell Biol* **24**: 2958–2967
- Papior P, Arteaga-Salas JM, Günther T, Grundhoff A, Schepers A (2012) Open chromatin structures regulate the efficiencies of Pre-RC formation and replication initiation in Epstein-Barr virus. *J Cell Biol* **198**: 509–528
- Prioleau MN, Gendron MC, Hyrien O (2003) Replication of the chicken beta-globin locus: early-firing origins at the 5' HS4 insulator and the rho- and betaA-globin genes show opposite epigenetic modifications. *Mol Cell Biol* **23**: 3536–3549
- 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
- Schones DE, Cui K, Cuddapah S, Roh TY, Barski A, Wang Z, Wei G, Zhao K (2008) Dynamic regulation of nucleosome positioning in the human genome. *Cell* **132**: 887–898
- Sekinger EA, Moqtaderi Z, Struhl K (2005) Intrinsic histone-DNA interactions and low nucleosome density are important for preferential accessibility of promoter regions in yeast. *Mol Cell* **18**: 735–748
- Sequeira-Mendes J, Diaz-Uriarte R, Apedaile A, Huntley D, Brockdorff N, Gómez M (2009) Transcription initiation activity sets replication origin efficiency in mammalian cells. *PLoS Genet* **5**: e1000446
- Sequeira-Mendes J, Gómez M (2012) On the opportunistic nature of transcription and replication initiation in the metazoan genome. *Bioessays* **34**: 119–125
- Smith DJ, Whitehouse I (2012) Intrinsic coupling of lagging-strand synthesis to chromatin assembly. *Nature* **483**: 434–438
- Stroud H, Otero S, Desvoves B, Ramirez-Parra E, Jacobsen SE, Gutierrez C (2012) Genome-wide analysis of histone H3.1 and H3.3 variants in Arabidopsis thaliana. *Proc Natl Acad Sci* **109**: 5370–5375
- Takai D, Jones PA (2002) Comprehensive analysis of CpG islands in human chromosomes 21 and 22. *Proc Natl Acad Sci USA* **99**: 3740–3745
- Teif VB, Vainshtein Y, Caudron-Herger M, Mallm JP, Marth C, Höfer T, Rippe K (2012) Genome-wide nucleosome positioning during embryonic stem cell development. *Nat Struct Mol Biol* **19**: 1185–1192
- Tolstorukov MY, Kharchenko PV, Park PJ (2010) Analysis of primary structure of chromatin with next-generation sequencing. *Epigenomics* **2**: 187–197
- Valouev A, Johnson SM, Boyd SD, Smith CL, Fire AZ, Sido A (2011) Determinants of nucleosome organization in primary human cells. *Nature* **474**: 516–520
- Vashee S, Cvetic C, Lu W, Simanek P, Kelly TJ, Walter JC (2003) Sequence-independent DNA binding and replication initiation by the human origin recognition complex. *Genes Dev* **17**: 1894–1908