Hereditary Genetics

Valenzuela, Hereditary Genetics 2012, S:1
DOI: 10.4172/2161-1041.S1-003

Review Article Open Access

Initiation of DNA Replication in the Human Genome

***Manuel S. Valenzuela**

Department of Biochemistry and Cancer Biology, School of Medicine, Meharry Medical College, 1005 D.B. Todd Jr. Boulevard, Nashville, TN 37208 , USA

Abstract

Replication of the human genome relies on the presence of thousands of origins distributed along each of the chromosomes. The activation of these origins occurs in a highly regulated manner to ensure that chromosomes are faithfully duplicated only once during each cell cycle. Failure in this regulation can lead to abnormal cell proliferation, or/ and genomic instability, the hallmarks of cancer cells. The mechanisms determining how, when, and where origins are activated remains still a mystery. However recent technological advances have facilitated the study of DNA replication in a genome-wide scale, and have provided a wealth of information on several features of this process. Here we present an overview of the current progress on our understanding of the initiation step of DNA replication in human cells, and its relationship to abnormal cell proliferation.

Introduction

DNA replication, a vital process in all biological systems, ensures the faithful transfer of genetic information from parent to progeny organisms, and constitutes the major determinant of cell proliferation. This process can be arbitrarily divided into three steps: Initiation, which activates the template DNA to facilitate the incorporation of the first nucleotide in the daughter strands; elongation, which allows further incorporation of nucleotides; and termination, which culminates with the production of two equal copies of the parental DNA. Of these three steps, initiation is the ultimate and critical step in the regulation of cell proliferation.

The replicon model proposed nearly 50 years ago by Jacob, Brenner, and Cuzin [1], has served as a good paradigm for understanding the initiation step of DNA replication. According to this model, initiation of DNA replication depends on the interaction of *trans*-acting factors (initiators) with *cis*- acting DNA sequences (replicators or origins).The origin, and adjacent DNA sequences whose replication depend on it, define an independent unit of replication, or replicon. Based on studies on the single replicon present in *E.coli*, the role of the initiator protein(s) has been expanded to not only to mark the position of the origin, but as a recruitment factor that facilitates the opening of the DNA helix, a step required for the initiation of DNA synthesis [2]. In contrast to bacterial genomes, eukaryotic genomes are composed of thousands of replicons, thus, their replication occurs in a segmental fashion [3-5]. Eukaryotic replication typically follows three cardinal rules: (a) replication is restricted to the S-phase of the cell cycle; (b) replicons are activated in a specific temporal order, some initiating early, and others late in the S phase [6], and (c) for any given replicon, replication occurs only once per S phase and thus re-initiation is prevented [7-8].

Initiation of Eukaryotic DNA Replication

The dissection of the initiation step of DNA replication in eukaryotes requires the answer to three specific questions; **How** are DNA sites containing origins activated only once during each cell cycle?; What determines **When** origins fire during the S phase?; and **Where** are origins localized along each of the chromosomes? At present our knowledge of the how, when, and where of initiation of DNA replication is not complete. A great deal is known about how origins are activated, but less information has been gathered about the underlying mechanisms behind the temporal order of origin firing, as well as, about their chromosomal distribution. Therefore, a coherent picture regarding the relationship between these properties, and how this provides the framework for its regulation, remains to be elucidated. In this paper I wish to present an overview of the current literature in the field with a focus on research in mammalian systems, and in particular, human cells. More comprehensive reviews dealing with several aspects of the initiation process in eukaryotes, methodological approaches used, and the potential misregulation of this process leading to cancer, are available. These include an excellent collection grouped in a special January 2010 issue of Chromosome Research [9], as well as others published elsewhere [10-14].

Activation of origins of DNA replication

Most of our current understanding about how origins are activated relies on the wealth of experimental information obtained studying this step in both the budding yeast *S. cerevisiae* and *Xenopus* oocytes. It turns out that in contrast to *E.coli*, where a single protein (DnaA) is required to load a helicase (DnaB) to make the double helix accessible to the DNA synthesizing machinery; in eukaryotes, this is achieved through a complex process that requires the formation of a multisubunit protein complex at replication origins. As shown in Figure 1, during the M to G1 transition of the cell cycle, a six subunit protein complex named, the origin recognition complex (ORC) binds to origin sequences, and serves as a landing pad for the assembly of other proteins known to be critical for the initiation step in DNA replication including Cdc6, Cdt1, and the MCM2-7 complex [15]. This multiprotein-DNA complex, named the pre-replicative complex (pre-RC), allows chromatin to be 'licensed' for replication in the subsequent S phase. Upon passage to the S phase, preRCs are activated by the action of specific S-phase cyclin-dependent kinases, and Cdc7/ASK kinase. These kinases, promote the recruitment of the GINS complex [82], and Cdc45 to MCM2-7, triggering the DNA helicase activity of the MCM2-7 complex [83-84]. The opening of the DNA helix and its stabilization by the binding of RPA facilitates the recruitment of DNA replication enzymes to begin DNA synthesis. The interplay of S phase-kinases with other components on the pre-RC also prevents the reformation of this complex, thus 'licensing' occurs only once per cell cycle at any given origin [15-16]. In animal cells the inactivation of the licensing system upon entry into the S phase is

***Corresponding author:** Manuel S. Valenzuela, Department of Biochemistry and Cancer Biology School of Medicine Meharry Medical College 1005 D.B. Todd Jr. Boulevard Nashville, TN 37208, USA, Tel: 615-327-6253; Fax: 615-327-6442; E-mail: mvalenzuela@mmc.edu

Received January 02, 2012; **Accepted** February 03, 2012; **Published** February 08, 2012

Citation: Valenzuela MS (2012) Initiation of DNA Replication in the Human Genome. Hereditary Genetics S1:003. doi:10.4172/2161-1041.S1-003

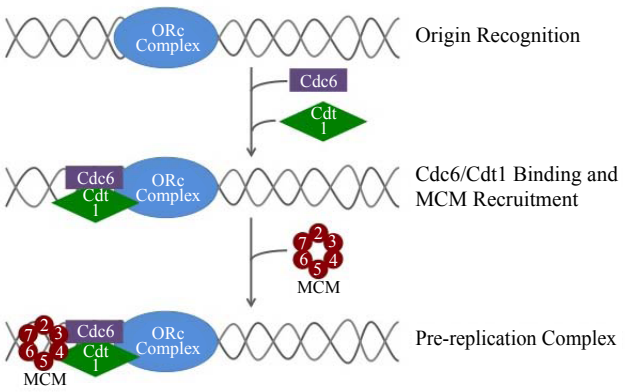

Figure 1: Assembly of the pre-replicative complex (pre-RC) renders chromatin competent for replication. In early G1 phase, the origin recognition complex (ORC) recruits Cdc6, which in turn promotes loading of MCM proteins onto chromatin. Upon passage to the S phase, activation of the Cdc7/Dbf4 kinase and S phase-promoting cyclin-dependent kinases (CDKs) induces a conformational change in the MCM complex that is required for unwinding of origin DNA by the recruitment of Cdc45 and GINS to the pre-RC. Initiation of DNA synthesis occurs when replication protein A (RPA) and DNA polymerase α- primase, are recruited to the unwound replication origin.

primarily through the regulation of Cdt1. This is achieved in two distinct ways: First, Cdt1 is subject to proteolysis as the cell cycle progresses through S and G2 [17-19], Second, Cdt1 is inhibited by geminin which specifically binds to Cdt1 during S, G2, and early mitosis [20-22]. The temporal separation of licensing and firing of origins creates a window of single opportunity for origin activation. DNA synthesis does not start until all pre-RCs are formed, and conversely pre-RC formation is inhibited during the DNA synthesis phase of the cell cycle. This provides the cell with another layer of control which prevents both, insufficient, and over-replication. The identification of structural and functional homologs of proteins involved in the initiation of DNA replication in multiple species prompted the speculation that these proteins, or at least ORC components, might bind to a common DNA sequence or motif present in all eukaryotic origins. However, such motifs have yet to be discovered. Nevertheless, it has revealed that components of the pre-RC are positioned downstream of complex networks of extra and intracellular signaling pathways that control cell growth, a finding that has elevated the status of pre-RC components as novel biomarkers of cell growth [23].

Temporal order of origin activation

It has been largely recognized that not all origins are activated at the same time, thus, during the duration of the S phase some origins fire early, others at mid S phase, and others still much later. At the nuclear level, a coordinated activation of origins has been observed, leading to the formation of replication foci, which depending on their timing of activation in the S phase, occupy distinct nuclear territories. It is believed that each foci contain hundreds of origins that are simultaneously activated in a replication factory. While the physical characterization of these putative factories is lacking, a nucleoprotein structure which resists high salt extraction (named the nuclear matrix) has been proposed to serve as the skeleton for replication factories. In agreement with this hypothesis, nuclear matrix preparations from mammalian cells have been found to be enriched for newly synthesized DNA, and for DNA structures containing forks and replicative bubbles [24-26]. At the chromosomal level, early activation has been associated with some basic chromosomal features such as Giemsa banding, G-C content and gene density [27]. In turn, late replication has been associated with chromosomal centromeric and telomeric regions. These findings point once more to the coordinated manner in which large chromosomal domains replicate simultaneously. Studies with individual genes have revealed a close correlation between transcription and early origin activation [28-30]. In general, housekeeping genes, which are constitutively transcribed in most cell types, replicate at early stages of the S phase, whereas the gene expression specificity of tissues determines the timing of replication of each gene. Thus, for a given tissue, expressed genes are replicated early whereas, repressed genes are replicated late, indicating that cellular development plays an important role in the temporal selection of origin activation.

Recent application of genome-wide analysis tools to query the relationship between transcription and replication timing in human cells [31-36], has confirmed that early replication correlates with transcriptional activity. However it has also revealed that this correlation is not air-tight. Thus some regions where active transcription is present appear to replicate late, conversely some silent regions occupy early replicating domains. One interesting example are CpG islands containing genes whose transcription persists even when located in late-replication zones [32-33,35,37]. These findings suggest that there maybe other factors other than transcription that regulate the timing of replication.

Origin specification in eukaryotes

The DNA sequences required for an origin specification vary considerably among eukaryotic organisms. At one end of the spectrum, in the baker's yeast S. cerevisiae, three or four 10-15 base pairs spread over a 100-150 bp DNA region are sufficient to function as an origin. These sequences include an essential but degenerate A-T rich11 bp sequence, named autonomously replicating consensus sequence (ACS), and less conserved but accessory DNA sequences called B elements which contain a DNA unwinding element (DUE) that presumably facilitates DNA melting [38]. In the fission yeast *S.pombe*, the minimum DNA region providing origin function extends to over 800-1000 bp. No ACS element has yet been determined; however several A-T rich sequences of 20-50 bp appear to be important for origin function in *S.pombe* [39]. Metazoan origins are even less defined and can extend over thousands of base pairs [40]. A requirement for specific DNA sequences for the origin of DNA replication in mammals has been suggested for a few chromosomal sites including a 5.7kb region of the amplified hamster dihydrofolate reductase domain [40-45], an 8kb DNA region in the human β globin gene [46-48], a 2.4kb fragment of the c-myc gene [49], and a 500 bp fragment downstream from the human lamin B2 gene [50-51]. Within each domain however, conflicting evidence has been accumulating suggesting that discrete DNA elements may not be required to initiate DNA replication. Finally, at the other extreme of origin sequence specification, the early embryos of D. *melanogaster* and *X.laevis* appear to require little or no sequence specificity to initiate DNA replication, which has been interpreted as a requirement for an extremely rapid S phase in these developmental stages [52]. From all the above, it can be said that although some DNA sites are consistently used as origins of DNA replication, the manner in which these sites are selected, and the sequences that determine their location in eukaryotic cells, remain still elusive. To gain a broader understanding about how DNA sequences are selected for origin specification, DNA microarray and more recently, second generation DNA sequencing technologies have been used to map out the location of origins of DNA replication in the human genome and correlate this location with other chromatin features. We will first describe some of the experimental approaches used to achieve this goal.

Methods to detect replication origins in a genome-wide scale

Two major experimental approaches have been used to identify and localize DNA sequences involved in the initiation step of DNA replication. 1) The first method for identifying sequences containing origins of DNA replication relies on the binding of pre-RC components to DNA to detect origins sequences. The methodology of choice in this case has been chromatin immunoprecipitation (ChIP). Briefly chromatin is chemically cross-linked in situ and later fragmented into smaller pieces. The resulting chromatin fragments are then immunoprecipitated with appropriate antibodies directed to the pre-RC component of choice. The DNA sequences contained in these fragments are then identified by PCR-based amplification of suspected targets, by hybridization to selected DNA microarrays, or by deep DNA sequencing. 2) The second approach identifies newly synthesized DNA at early stages of replicon activation. This approach detects initiation sites for DNA replication which may or not coincide with origin sequences. Unfortunately, both terms have been used interchangeable in the literature. Several methodologies have been used to identify newly synthesized DNA including leading-strand analysis, nascent-strand abundance, density transfer, or gel trapping of replicating DNA [36, 53-59; for further description see refs. 12 & 60]. Of all these methods, the most commonly used is the nascent DNA strand abundance assay which depends on the isolation of short DNA fragments contained in newly activated replicons. A variant of this approach as applied to DNA microarray technology is illustrated in Figure 2. Total DNA is collected from an asynchronous population of proliferating cells. Nascent DNA strands are then released by heat denaturation, size fractionated on a 5-30% sucrose gradient to select a pool of fractions containing DNA in the 0.7-1.5 kb size range. This pool is then subjected to digestion with λ-exonuclease in order to digest contaminant DNA fragments that do not harbor an RNA sequence at their 5' ends [61]. This fraction constitutes the test DNA. Total genomic DNA, obtained from the same cell line and sonicated to a similar size range constitutes the reference DNA. Usually both the test and reference DNAs are assayed by real time PCR to confirm the enrichment of *bonna fide* origins relative to adjacent non-origin sequences in the test DNA sample. Both test and reference DNAs are then labeled with Cy-5 and Cy-3 dUTP derivatives, respectively and hybridized to a custom made DNA tiling microarray containing 50-60 nt DNA probes staggered in 50-60 bp steps and spanning selected chromosomal regions. Signals of high Cy-5/Cy-3 ratios are then registered at potential sites of origin activation and their positions assigned to the annotated sequence of the human genome. This methodology provides comprehensive information about initiation sites which are active throughout the S phase, and thus does not require prior synchronization of the cells.

It is important to note however that these approaches are not perfect. Since arrays can only be built to cover regions where hybridization is unique (i.e. regions containing repetitive DNA sequences are excluded), and that this coverage is limited given the size of the genome, the information obtained with array platforms is incomplete and maybe somewhat biased. In contrast, the application of second generation DNA sequencing allows almost complete coverage of the genome and also is exempted from biases that may arise from querying selected regions of chromosomes or from differences in the efficiency of hybridization of the probes present in the array. Unfortunately, second generation DNA sequencing approaches are currently economically taxing. Therefore, most of the genome wide studies so far have used DNA microarrays. Another caveat that both of these approaches have in common is that since the DNA material arises from cell populations, the information obtained represent a cumulative profile of all the cells present in the population, thus differences in the response of individual cells can not be assessed. An alternative approach that circumvents all these problems is the direct visualization of DNA replication at the level of individual cells by the use DNA combing to stretch pulse labeled DNA to quantify both, the presence of initiation sites, as well as, the direction of replication in single DNA fibers [12]. This approach is quite robust to map origin distribution and elongation rates in anonymous DNA segments. By combining this technique with fluorescent in hybridization (FISH) to mark particular chromosomal regions, it is possible to determine the dynamics of the replication program along this region in a cell population. The major limitation however is the relatively short length of the DNA fibers (less than 1 Mb), making the information obtained to be somewhat biased toward fibers showing shorter inter-origin distances.

Genomic distribution of human replication origins

Classical studies using fiber autoradiography have revealed that the number of origins of replication in early embryos is much larger than in differentiated tissues, indicating that the genome contains an excess of origins, compared to the minimum required to replicate the genome [6]. In fact it has been estimated that the number of potential origins is about 4-5-fold times the number required to complete the replication of the genome. All these origins are supposed to be primed as pre-RCs by the end of the G1 phase of the cell cycle thus, paraphrasing the Jesuit dictum "*many are called, but few are chosen*" it has been proposed that some pre-RCs are activated upon passage to the S and mark the position of active replicons, whereas the remaining pre-RCs are located in regions that are passively replicated by adjacent replicons [62]. Recent studies using single fiber analysis have indicated a higher level of plasticity in the way origins are chosen. Origins can be classified as constitutive, flexible, or dormant. Constitutive origins fire upon entry into the S-phase in all cell types; flexible origins can be arbitrarily used in different cells of the same tissue; and dormant origins are those used only under special circumstances [10]. It is believed that the major determinant of origin usage is the completion of a single copy of the genome. Therefore the availability of extra origins offers the genome with enough potential resources to achieve this end, under a variety of physiological conditions.

What constitutes an origin, and what determines its localization? From the way origins are activated via the formation of the pre-RC, the simplest definition of an origin would be as that DNA sequence to which ORC binds. Therefore it was thought that the reconstitution of the purified human ORC complex on DNA could resolve this issue. However, when this task was achieved by Vashee et al [63], the reconstituted human ORC, while functionally active in an *X.laevis* DNA replication assay, and capable of DNA binding affinity, could not distinguish *bona fide* origin from non-origin DNA sequences. In fact the highest binding affinity of ORC was found with a synthetic polynucleotide containing A-T rich tracts (a 45-fold increase over origin sequences). These results suggested that origin recognition by ORC may either be DNA sequence independent or that other accessory elements provided DNA binding specificity.

Recent reports have addressed the question of localization of active origins in a large scale by applying microarray or second generation DNA sequencing technologies [53, 55-59]. Three major observations could be drawn from all these studies. First, several hundred origins of DNA replication could be localized in both genic and non-genic chromosomal regions. The distribution of origins was not uniform along the regions queried, since in some instances large DNA tracts appeared to be devoid of active origins. Second, while the cell lines as

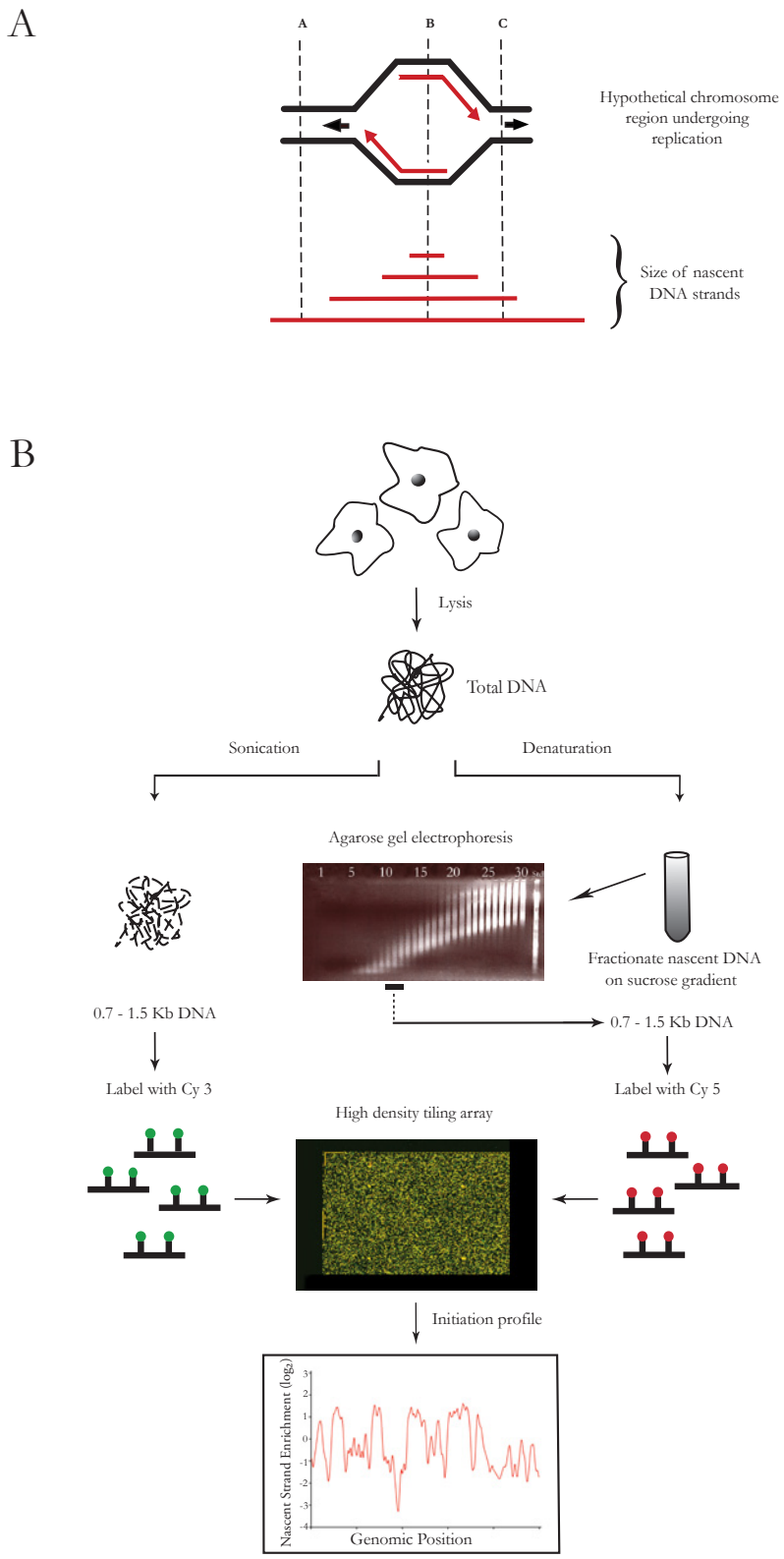

Figure 2: Experimental strategy to determine the enrichment of short nascent DNA using DNA microarrays. (A) A hypothetical chromosome region undergoing DNA replication through points A, B, and C will produce nascent strand DNA of increasing size. The shortest DNA, which includes point B, will contain an initiation site for DNA replication. (B) Total DNA isolated from synchronously growing cells is isolated and the nascent strand DNA released by heat denaturation is size fractionated through a 5-30% sucrose gradient. Fractions containing DNA in the 0.7-1.5kb size range are pooled and treated with λ exonuclease. A similar size range DNA is produced by sonication of total DNA. Both DNA's are differentially labeled and hybridized to a custom made high density tiling DNA array. By calculating the hybridization ratios at each one of the probes contained in the array a nascent DNA enrichment peak can be positioned within the genomic regions represented in the array.

well as the array platforms used in these studies varied, only a modest (less than expected) concordance of origin localization was found when overlapping chromosomal regions were compared. As recently pointed out by Gilbert [12], the best illustration of this lack of reproducibility is the less than 14 % overlapping shown by two independent studies using the same cell line and similar array platforms (ENCODE array) [36,55], More importantly however, these studies have revealed a common feature of origin localization. Origins are enriched at annotated transcriptional units, particularly at promoter regions harboring epigenetic markers that identify open chromatin structure, as well as at active PolII binding sites. These results suggest that origin selection may primarily depend on the accessibility of ORC to chromatin regions. Since nucleosome occupancy determines the availability of genomic DNA, it has been proposed that nucleosome depleted regions may serve as preferential sites for origin activity [12]. Support for this hypothesis has been gained by a recent mining of origin distribution data bases for both budding yeast and human genomes which indicated that active origins tended to be localized at nucleosome depleted locations [64]. Additional evidence has been provided through studies in other species such as a genome-wide study of ORC localization in *Drosophila melanogaster* cells [65], and a related study within the initiation region of the DHFR locus in Chinese Hamster cells [66], where it was found that ORCs are localized at nucleosome free regions. The nature of nucleosome free regions in the genome has been recently redefined by the discovery that nucleosomes containing the histone variant H3.3 [67], have a fast turnover rate and that at least in *D. melanogaster*, occupy the same chromosomal positions as ORC2, one of the subunits of the ORC complex [68]. These findings support the idea that rapid nucleosome turnover may be a determinant factor on ORC binding and therefore origin positioning. Based on these findings an "opportunistic" model for origin selection has been advanced, by which origins are defined by DNA accessibility rather than by specific sequence features [14]. It will be interesting to investigate if ORC occupancy in the human genome also follows these rules.

Initiation of DNA replication and cancer

As indicated earlier, the initiation step of DNA replication constitutes the most important determinant of cell proliferation. Therefore it is natural to assume that a lack of proper regulation of this step could lead to the abnormal cell growth observed in cancer cells. It is conceivable that changes in the licensing program of origins, their temporal activation, or their localization and number could have an impact in altering the replication program of normal cells. Thus, it might be instructive to analyze how each of these functions could affect cell growth. To begin with, the ultimate goal of the licensing program is to activate enough origins to ensure the complete replication of the genome. If fewer than needed origins are licensed, the genome carries the risk of reaching mitosis with large tracks of unreplicated DNA, which can lead to DNA rearrangements or breaks. On the other hand, if origins are licensed more than once in a single cell cycle, DNA amplifications will ensue which may lead to the genomic aneuploidy commonly observed in most cancers. Up-regulation of all the known components of the pre-RC has been observed in a variety of both cancer cell lines and tumor tissues compared to their normal counterparts. For instance, Cdt1 and Cdc6 were found to be overexpressed in cervical, lung, and brain cancers; up-regulation of MCM2 and MCM5 were observed in esophageal cancers; and increased expression of ORC in cervical cancer cell lines [13]. Interestingly, this up-regulation was not restricted to the proliferating cells in the tumor tissue, but was also observed on adjacent layers of low or non-proliferating cells [69]. Over-expression of pre-RC components also correlated with higher tumor grade and poor prognosis [70]. Based on these properties the use of pre-RC expression signatures are being considered as potential tumor diagnostic markers [23].

Reduced expression of some pre-RC components has also been shown to cause severe chromosomal instability ultimately leading to cancer but only under genotoxic conditions [71-72]. An increased incidence of chromosomal breaks, and chromosome instability was observed in mice harboring a hypomorphic mutation in Mcm4 (Chaos3 mutant) upon exposure to a replication inhibitor, and about 80% of Chaos3 females developed mammary adenocarcinomas [71]. A similar mutation in Mcm2 in mice also caused genomic instability, but in this case mice developed T- and B-lymphomas [72]. Overexpression of pre-RC components could alter the regulation upon passage of the cells into S phase leading to re-replication, a likely determinant of tumorigenesis. A direct demonstration as to how this deregulation could become oncogenic was obtained through the studies of Gonzales et al. [73], who showed that overexpression of Cdc6 in mammalian cells caused specific hypermethylation of the tumor suppressor INK4/ARF locus. This was achieved by the binding of Cdc6 at an origin located around a regulatory element of the INK/ARF locus and the subsequent recruitment of histone deacetylases which cause heterochromatization around this locus. As a result, repression of p14ARF, p16NH4A, and p14NK4B expression ensued. The product of these genes are important activators of both p53 and Rb tumor suppressor pathways [73].Thus enhanced cell proliferation is the result of derepression of cell cycle progression coupled to overexpression of Cdc6. To gain a better understanding about the oncogenic potential of pre-RC components, it would be interesting to find out how pre-RCs become deregulated in the first place.

The second area of misregulation leading to cancer could occur by the alteration of the temporal order of origin activation. It is conceivable that the change in the replication of an early or late origin could cause a disturbance in the temporal expression of crucial genes. Alternatively, origins may not function at their required time, leaving tracts of unreplicated DNA (a potential source of chromosomal abnormalities), upon passage through mitosis. Unexpectedly, the application of single DNA fiber analysis to study of the replication program around FRA3B, the most active common fragile site found in human lymphoblasts, has led to the discovery that the strength of this site depends on the proper functioning of late origins of DNA replication present at this region. In lymphoblastoid cells that had been subjected to replicative stress few origins fired at the location of the fragile site. In contrast in fibroblastoid cells, where the fragility at this site is not observed, the site contained several active origins [74]. Thus a tissue specificity on the stability of fragile sites was uncovered, which may depend on a differential activation of late origins of DNA replication occupying this region. It is not clear though how this specificity is determined, but this finding clearly illustrates a connection between the control of replication timing and chromosomal instability.

Finally, there is the possibility that in cancer cells alternative replication programs may be operative. According to this hypothesis, the changes in the distribution and activity of replication origins could constitute important determinants of the abnormal physiology observed in cancer cells. Since the pioneering studies by J.H Taylor [75,76], in the 1970s, it has been recognized that the number of active origins used in the genome could be increased under conditions of limited cell growth. More recently, a more detailed analysis using single fiber DNA technology at a specific chromosomal location has confirmed Taylor's results by demonstrating that under conditions of reduced nucleotide pool levels more origins are active [77]. This compensatory effect

ensures duplication of the genome while maintaining the duration of the S phase. Several reports have suggested that in both transformed and tumor derived cells the number and/or activity of origins per unit length of DNA is increased [78-80; 85-87]. Unfortunately, many of these studies have been limited in scope, thus support for this hypothesis is still lacking. More recently a real time PCR–based nascent DNA abundance assay was used to compare the location and number of active origins around a 78 kb region of the human chromosome 2q34 in both cancer and normal breast cell lines. The replication profile in three different breast cancer cell lines was found to be significantly different that a normal breast cell line. More origins appeared to be activated in the cancer cell lines, but their distribution was broader compared to normal cells [81]. These results suggest that at least in this chromosomal region more origins are recruited in breast cancer cell lines. The significance of this finding will require a large scale comparison of the replication profile between normal and their corresponding cancer cell homologs. This information, coupled with single DNA fiber studies could provide us with enough information to validate or discard this hypothesis.

Summary and Perspectives

DNA replication of human chromosomes occurs by the activity of thousands of units of replication, or replicons. Replicons function following a temporal program which is developmentally regulated and tissue specific. Each replicon harbors an origin of replication which has been primed by the formation a multi-protein complex (pre-RC) prior to S-phase. There is however more origins present in each chromosome, most of which remain dormant. This origin excess allows the cells to respond to situations of replicative stress by the generation of active replicons from dormant origins. While a great deal has been learned about how origins are primed, less information is available about what determines the temporal order, or the positioning of active origins. The recent application of genome-wide approaches to study these questions, have provided a wealth of information about the relationship of origin activation to transcription and chromatin structure but we still lack a coherent understanding about this process. Given the importance of the initiation step in the regulation of cell proliferation, it is not surprising that some components of pre-RC may have an oncogenic function. There is however much to be learned about the signaling pathways that connect mitogens with these components, and how the expression of pre-RC components is regulated. The application of more massive techniques such as second generation DNA sequencing should provide us more information about the cross-talk of DNA replication with other DNA transactions. The next frontier however will be to understand the relationship of chromosome structure and function, as well as how the spatial interactions between chromosomes and their positioning inside the nucleus influence their function.

Acknowledgements

This paper is dedicated to the memory of Susan H. Armor. The author wishes to acknowledge support from the National Institutes of Health (Grant CA138180), and to thank J. Stewart for help with Figure 2.

References

1. Jacob FS, Brenner S, Cuzin F (1963) On the regulation of DNA replication in bacteria. Cold Spring Harb Symp Quant Biol 28: 329-348.
2. Baker TA, Kornberg A (1992), Second Edition (New York: W.H. Freeman & Co).
3. Taylor JH (1968) Rates of chain growth and units of replication in DNA of mammalian chromosomes. J Mol Biol 31: 579-594.
4. Hsu TC (1964) Chromosome constitution in neoplasms. Proc Natl Cancer Conf 5: 49-56.
5. Aladjem MI, Fanning E (2004) The replicon revisited: an old model learns new tricks in metazoan chromosomes. EMBO Rep 5: 686-691.
6. Hand R (1978) Eukaryotic DNA: Organization of the genome for replication. Cell 15: 317-325.
7. Blumenthal AB, Kriegstein HJ, Hogness DS (1974) The units of DNA replication in Drosophila melanogaster chromosomes. Cold Spring Harb Symp Quant Biol 38: 205-223.
8. Blow JJ, Dutta A (2005) Preventing re-replication of chromosomal DNA. Nat Rev Mol Cell Biol 6: 476-486.
9. Jackson DA, Prioleau MN (2010) Eukaryotic DNA Replication: Is Time of the Essence? Chromosome Res 18: 1-5.
10. Mechali M (2010) Eukaryotic DNA replication origins: many choices for appropriate answers. Nat Rev Mol Cell Biol 11: 728-738.
11. Blow JJ, Gillespie PJ (2008) Replication licensing and cancer- a fatal entanglement? Nat Rev Cancer 8: 799-806.
12. Gilbert DM (2010) Evaluating genome-scale approaches to eukaryotic DNA replication. Nat Rev Genet 11: 673-684.
13. Lau E, Tsuji T, Guo L, Lu S-H, Jiang W (2007) The role of pre-replicative complex (pre-RC) components in oncogenesis. FASEB J 21: 3786-3794.
14. Sequeira-Mendes J, Gomez M (2012) On the opportunistic nature of transcription and Replication intitiation in the metazoan genome. Bioessays 34: 119-125.
15. Machida YJ, Hamlin JL, Dutta A (2005) Right place, right time, and only once: Replication initiation in metazoans. Cell 123: 13-24.
16. Bell SP, Dutta A (2002) DNA replication in eukayotic cells. Annu Rev Biochem 71: 333-374.
17. Nishitani H, Taraviras S, Lygerou Z, Nishimoto T (2001) The human licensing factor for DNA replication Cdt1 accumulates in G1 and is destabilized after initiation of S-phase. J Biol Chem 276: 44905-44911.
18. Li X, Zhao Q, Liao R, Sun P, Wu X (2003) The SCF(Skp2) ubiquitin ligase complex interacts with the human replication licensing factor Cdt1 and regulates Cdt1 degradation. J Biol Chem 278: 30854-30858.
19. Arias EE, Walter JC (2006) PCNA functions as a molecular platform to trigger Cdt1 destruction and prevent re-replication. Nat Cell Biol 8: 84-90.
20. McGarry TJ, Kirschner MW (1998) Geminin, an inhibitor of DNA replication, is degraded during mitosis. Cell 93: 1043-1053.
21. Tada S, Li A, Maiorano D, Mechali M, Blow JJ (2001) Repression of origin assembly in metaphase depends on inhibition of RLF-B/Cdt1 by geminin. Nature Cell Biol 3: 107-113.
22. Wohlschlegel JA, Dwyer BT, Dhar SK, Cvetic C, Walter JC, et al. (2000) Inhibition of eukaryotic DNA replication by geminin binding to Cdt1. Science 290: 2309-2312.
23. Giaginis C, Vgenopoulou S, Vielh P, Theocharis S (2010) MCM proteins as diagnostic and prognostic tumor markers in the clinical setting. Histol Histopathol 25: 351-370.
24. Berezney R, Coffey DS (1975) Nuclear protein matrix: association with newly synthesized DNA. Science 189: 291-293.
25. Dijkwel PA, Mullenders LH, Wanka F (1979) Analysis of the attachment of replicating DNA to a nuclear matrix in mammalian interphase nuclei. Nucleic Acids Res 6: 219-230.
26. Valenzuela MS, Mueller G, Dasgupta S (1983) Nuclear matrix-DNA complex resulting from EcoRI digestion of HeLa nucleoids is enriched for DNA replicating forks. Nucleic Acids Res 11: 2155-2164.
27. Cohen SM, Cobb ER, Cordeiro-stone M, Kaufman DG (1998) Identification of chromosomal bands replicating early in the S phase of normal human fibroblasts. Exp Cell Res 245: 321-329.
28. Gilbert DM (1986) Temporal order of replication in Xenopus laevis 5S ribosomal RNA genes in somatic cells. Proc Nat Acad Sci 83: 2924-2928.
29. Schmidt M, Migeon BR (1990) Asynchronous replication of homologous loci on human active and inactive X chromosomes. Proc Natl Acad Sci 87: 3685-3689.
30. Hansen RS, Canfield TK, Lamb MM, Gartler SM, Laird CD (1993) Association of fragile X syndrome with delayed replication of the FMR1 gene. Cell 73: 1403-1409.
31. Watanabe Y, Fujiyama A, Ichiba Y, Hattori M, Yada T, 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.
32. Jeon Y, Bekiranov S, Karnani N, Kapranov P, Ghosh S, et al. (2005) Temporal profile of replication of human chromosomes. Proc Natl Acad Sci 102: 6419-6424.
33. Woodfine K, Fiegler H, Beare DM, Collins JE, Mccann OT, et al. (2004) Replication timing of the human genome. Hum Mol Genet 13: 191-202.

34. Woodfine K, Beare DM, Ichimura K, Debernardi S, Mungall AJ, et al. (2005) Replication timing of human chromosome 6. Cell Cycle 4: 172-176.

35. White EJ, Emanuelsson O, Scalzo D, Royce T, Kosak S, et al.(2004) DNA replication-timing analysis of human chromosome 22 at high resolution and different developmental states. Proc Natl Acad Sci 101: 17771-17776.

36. Karnani N, Taylor CM, Malhotra A, Dutta A (2010) Genomic study of replication initiation in human chromosomes reveals the influence of transcription regulation and chromatin structure on origin selection. Mol Biol Cell 21: 393-404.

37. Hiratani I, Ryba T, Itoh M, Yokochi T, Schwaiger M, et al. (2008) Global reorganization of replication domains during embryonic stem cell differentiation. PLoS Biol 6: e245.

38. Newlon CS (1996) DNA Replication in Yeast In DNA Replication In Eukaryotic Cells. Cold Spring Harbor Laboratory Press pp 873-914.

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

40. DePamphilis ML (1996) Origins of DNA Replications. In DNA Replication In Eukaryotic Cells. Cold Spring Harbor Laboratory Press pp 45-86.

41. Heintz NH, Milbrandt JD, Greisen KS, Hamlin JL (1983) Cloning the initiation region of a mammalian chromosomal replicon. Nature 302: 439-441.

42. Burhans WC, Selegue JE, Heintz NH (1986) Replication intermediates formed during initiation of DNA synthesis in methotrexate-resistant CHOC 400 cells are enriched for sequences from a specific, amplified restriction fragment. Biochemistry 25: 441-449.

43. Pelizon C, Diviacco S, Falaschi A, Giacca M (1996) High-Resolution mapping of the origin of DNA replication in the hamster dihydrofolate reductase gene domain by competitive PCR. Mol Cell Biol 16: 5358-5364.

44. Kobayashi T, Rein T, DePamphilis ML (1998) Identification of primary initiation sites for DNA replication in the hamster dihydrofolate reductae gene initiation zone. Mol cell Biol 18: 3266-3277.

45. Altman A, Fanning E (1999) The Chinese hamster DHFR replication origin beta is active at multiple ectopic chromosomal locations and requires specific DNA sequence elements for activity. Mol Cell Biol 21: 1098-1110.

46. Kitsberg D, Seliq S, Keshet I, Cedar H (1993) Replication structure of the human β-globin domain. Nature 366: 588-590.

47. Aladjem, MI, Groudine M, Brody LL, Dieken ES, Fournier REK, et al. (1995) Participation of the human B-globin locus control region in initiation of DNA replication. Science 270: 815-819.

48. Aladjem MI, Rodewald LW, Kolman JL Wahl GM (1998) Genetic dissection of a mammalian replicator in the human β-globin locus. Science 281: 1005-1009.

49. Leffak M, James CD (1989) Opposite replication polarity of the germ line c-myc gene in HeLa cells compared with that of two Burkitt lynphoma cell lines. Mol Cell Biol 9: 586-593.

50. Giacca M, Zentilin L, Morio P, Diviacco S, Dimitrova D, et al. (1994) Fine mapping of a replication origin of human DNA. Proc Natl Acad Sci 91: 7119-7123.

51. Abdurashidova G, Deganuto M, Klima R, Riva S, Biamonti G, et al. (2000) Start sites of bi-directional DNA synthesis at the human lamin B2 origin. Science 287: 2023-2026.

52. Blow JJ (2001) Control of chromosomal DNA replication in the early Xenopus embryo. EMBO J 20: 3293-3297.

53. Karnani N, Taylor C, Malhotra A, Dutta A (2007) Pan-S replication patterns and chromosomal domains defined by genome-tiling arrays of ENCODE genomic areas. Genome Res 17: 865-876.

54. Sequeira-Mendes J, Diaz-Uriarte R, Apedaile A, Huntley D, Brockdorff N, et al. (2009) Transcription initiation activity sets replication origin efficiency in mammalian cells. PLoS Genet 5: e1000446.

55. Cadoret JC, Meisch F, Hassan-Zadeh V, Luyten I, Guillet C, et al. (2008) Genome-wide studies highlight indirect linksbetween human replication origins and gene regulation. Proc Natl Acad Sci 105: 15837-15842.

56. Valenzuela MS, Chen Y, Davis S, Yang F, Walker RL, et al. (2011) Preferential localization of human origins of DNA replication at the 5'-ends of expressed genes and at evolutionary conserved DNA sequences. PLoS One 6: e17308.

57. Lucas I, Palakodeti A, Jiang Y, Young DJ, Jiang N, et al. (2007) Highthroughput mapping of origins of replication in human cells. EMBO Rep 8: 770-777.

58. Martin MM, Ryan M, Kim R, Zakas AL, Fu H, et al. (2011) Genome-wide depletion of replication initiation events in highly transcribed regions. Genome Res 21: 1822-1832.

59. Mesner LD, Valsakumar V, Karnani N, Dutta A, Hamlin JL, et al. (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.

60. Cadoret JC, Prioleau MN (2010) Genome-wide approaches to determining origin distribution. Chromosome Res 18: 79-89.

61. Gerbi SA, Bielinsky AK (1997) Replication initiation point mapping. Methods 13: 271-280.

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

63. Vashee S, Cvetic C, Lu W, Simancek P, Kelly TJ, et al. (2003) Sequence-independent DNA binding and replication initiation by the human origin recognition complex.Genes Dev 17: 1894-1908.

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

65. MacAlpine HK, Gordân 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.

66. Lubelsky Y, Sasaki T, Kuipers MA, Lucas I, Michelle MLB, et al. (2011) 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.

67. Mito Y, Henikoff JG, Henikoff S (2005) Nat Genet 37: 1090-1097.

68. Deal RB, Henikoff JG, Henikoff S (2010) Science 328: 1161-1164.

69. Gonzalez MA, Pinder SE, Callagy G, Vowler SL, Morris LS, et al. (2003) Minichromosome maintenance protein 2 is a strong independent prognostic marker in breast cancer. J Clin Oncol 21: 4306-4313.

70. Gonzalez MA, Tachibana KE, Laskey RA, Coleman N (2005) Control of DNA replication and its potential clinical exploitation. Nat Rev Cancer 5: 135-141.

71. Shima N, Alcaraz A, Liachko I, Buske TR, Andrews CA, et al. (2007) A viable allele of Mcm4 causes chromosome instability and mammary adenocarcinomas in mice. Nat Genet 39: 93-98.

72. Pruitt SC, Bailey KJ, Freeland A (2007) Reduced Mcm2 expression results in severe stem/progenitor cell deficiency and cancer. Stem Cells 25: 3121-3132.

73. Gonzalez S, Klatt P, Delgado S, Conde E, Lopez-Rios F, et al. (2006) Oncogenic activity of Cdc6 through repression of the INK4/ARF locus. Nature 440: 702-706.

74. Letessier A, Millot GA, Koundrioukoff S, Lachagès AM, Vogt N, et al. (2011) Debatisse MCell-type- specific replication initiation programs set fragility of the FRA3B fragile site. Nature 470: 120-123.

75. Taylor JH, Adams AG, Kurek MP (1973) Replication of DNA in mammalian chromosomes II Kinetics of 3H-thymidine incorporation and the isolation and partial characterization of labeled subunits at the growing point. Chromosoma 41: 361-384.

76. Taylor JH (1977) Increase in DNA replication sites in cells held at the beginning of S phase. Chromosoma 62: 291-300.

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

78. Martin RG, Oppenheim A (1977) Initiation points for DNA replication in nontransformed and simian virus 40-transformed Chinese hamster lung cells. Cell 11: 859-869.

79. Oppenheim A, Martin RG (1978) Initiation points for DNA replication in nontransformed and simian virus 40-transformed BALB/c 3T3 cells. J Virol 25: 450-452.

80. Frumm RA, Khonder ZS, Kaufman DG (2009) Temporal differences in DNA replication during the S phase using single fiber analysis of normal human fibroblasts and glioblastoma T98G cells. Cell Cycle 8: 3133-3148.

81. Valenzuela MS, Hu L, Lueders J, Walker RL, Meltzer PS (2012) Broader utilization of origins of DNA replication in cancer cell lines along a 78 kb region of human chromosome 2q34. J Cell Biochem 113: 132-140.

82. Takayama Y, Kamimura Y, Okawa M, Muramatsu S, Sugino A, et al. (2003) GINS, a novel multiprotein complex required for chromosomal initiation in the budding yeast. Genes Dev 17: 1153-1165.

83. 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. Nucl Acid Res 37: 2087-2095.

84. Cohen S, Chastain PD, Cordeiro-Stone M, Kaufman DG (2009) DNA replication and the GINS complex:localization on extended chromatin fibers. Epigenetics Chromatin 2:6.

85. 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. J Mol Biol 273: 509-518.

86. Di Paola D, Price GB, Zannis-Hadjopoulos M (2006) Differentially active origins of DNA replication in tumor versus normal cell. Cancer Res 66: 5094-5103.

87. Di Paola D, Rampakakis E, Chan MK, Arvanitis DN, Zannis-Hadjopoulos M (2010) Increased origin activity in transformed versus normal cell: identification of novel protein players involved in DNA replication and cellular transformation. Nucleic Acids Res 38: 2314-2331.

This article was originally published in a special issue, **Hereditary Disorders** handled by Editor(s). Dr. X. Long Zheng , University of Pennsylvania, USA