# Genome organization of DNA replication timing and its link to chromatin and transcription

**Inauguraldissertation**

zur
Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

von

**Michaela Schwaiger**
aus Österreich

Basel, 2008

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät der Universität Basel auf Antrag von Dr. Dirk Schübeler, Professor Dr. Frederick Meins Jr., und Professor Dr. Anne Donaldson.

Basel, den 14. Oktober 2008

Professor Dr. E. Parlow
Dekan

# Table of Content

# 1. Summary

The replication of the genome is a highly organized process. Not every sequence replicates at the same time, instead some genes replicate early, while others replicate later during S phase. The timing of DNA replication is conserved within consecutive cell divisions of a given cell type. The aim of this PhD thesis was a better understanding of the regulation of DNA replication. In particular, I determined the genomic landscape of the timing of DNA replication in the *Drosophila* genome, and defined the dynamics of replication timing and its connection with chromatin and transcription.

Recent genome-wide studies of replication timing and transcription suggested a strong relation between both processes since early replicating genes are more likely to be expressed than genes replicating later during S phase. This correlation is not absolute, therefore raising the question if replication timing is dynamic between different epigenetic states, or if it is static and this correlation is driven mostly by a distinct set of constitutively expressed genes. To create a defined replication timing program, initiation of DNA replication needs to be controlled in space and time. The location and time of firing of the closest origin of replication (ori) defines the replication timing of a certain sequence. However, only few metazoan origins of replication have been identified, and they lack a consensus sequence. Therefore it has been suggested that replication initiation is defined epigenetically.

To address this problem I generated datasets for replication timing in two *Drosophila* cell types representing different developmental states and gender, using high-resolution tiling arrays. This detailed analysis permitted the identification of zones of replication initiation throughout the whole genome. Surprisingly, I could identify a higher number of initiation zones in early and late S phase than in mid S phase. This work also shows that about 20% of the *Drosophila* genome replicates at different times in the two cell types. These differences in replication timing correlate with differences in gene expression, chromatin modifications and position in the nucleus relative to the nuclear periphery. Interestingly, the dosage compensated male X chromosome replicates predominantly in early S phase. This correlates with chromosome-wide hyperacetylation, often independent of transcription differences. High levels of acetylation on Lysine 16 of Histone H4 were also detected at initiation zones, supporting the model of epigenetically defined replication initiation.

In addition, I addressed the potential role of chromatin-bound proteins in modulating replication timing. Using RNA interference, I could show that the absence of

Heterochromatin Protein 1 (HP1) has distinct effects on replication timing many of which appear transcription independent.

Together, my results reveal organizational principles of DNA replication of the *Drosophila* genome and indicate that replication timing is dynamic and chromatin-dependent.

# 2. Introduction

For a cell to divide, its genome has to be replicated completely, yet not more than once, before cell division occurs. Furthermore, transcription of genes has to occur on the same DNA molecule to provide essential proteins. Thus, DNA replication and transcription have to be coordinated in space and time. This is further complicated by the fact that eukaryotic genomes are confined into the small volume of a cell nucleus, and therefore need to be packaged into higher order structures. This organization has a major effect on DNA readout, since the chromatin packaged DNA has to be made accessible for transcription, replication and DNA repair to occur, and some regions of the genome are differently accessible than others. Such differences in compaction state, at the same time, provide an opportunity for differential regulation of gene expression in different cell types. These gene expression programs have to be remembered by a cell through multiple divisions. Epigenetic differences between cell types could provide such a cellular memory over the lifetime of an organism. The following section will give an introduction to the organization of DNA in a eukaryotic cell nucleus. It will also introduce the regulation of two chromatin-templated events, transcription and DNA replication. Finally, the link between replication to chromatin and transcription will be discussed.

## 2.1. Chromatin structure and transcription

In contrast to prokaryotes, the cells of eukaryotes, from yeast to humans, contain a specialized compartment, called the nucleus, into which almost all the DNA is confined. Furthermore, the DNA in the nucleus is wrapped around an equal mass of proteins, forming a nucleoprotein complex called chromatin, which controls not only genetic inheritance, but also the activity of genes.

### 2.1.1. The nucleosome

The most abundant proteins within chromatin are called histones (reviewed in (Felsenfeld *et al.* 2003)). There are equimolar ratios of four canonical core histones inside the cell nucleus: H2A, H2B, H3 and H4. They are highly basic proteins, and highly conserved throughout all eukaryotes (Sullivan *et al.* 2003), suggesting that all eukaryotes might harbor a common chromatin architecture. Histones H3 and H4 form hetero-tetramers and H2A and H2B form hetero-dimers. The H3/H4 tetramer together with two H2A/H2B dimers then form the histone octamer, around which 146bp of DNA are wrapped in 1 ¾ superhelical turns to form the nucleosome (Kornberg 1974). The

histones' N-terminal tails extend outwards from the nucleosome core particle, allowing for their interaction with other proteins (Luger *et al.* 1997) (Figure 1A) and are subject to a variety of post-translational modifications including methylation, acetylation, phosphorylation, and ubiquitination which are known to exert various degrees of regulation on gene-specific transcription (reviewed in (Jenuwein *et al.* 2001)). Metazoan cells also contain a linker histone (such as histone H1) associated with each nucleosome, which might stabilize regions of chromatin into a condensed state (reviewed in (Felsenfeld et al. 2003)).

The nucleosomes are connected to each other via a short (10-60bp) stretch of "linker" DNA in between them (Olins *et al.* 1974). Such an array of nucleosomes is about 10nm in diameter, but can condense further to form a 30nm fiber, where the DNA is compacted about 50-fold (Figure 1B). The exact structure of this 30nm fiber remains subject of debate (reviewed in (Tremethick 2007)). Interestingly, it has been shown that histone tails play an important role in higher order folding of nucleosomes (Dorigo *et al.* 2003; Shogren-Knaak *et al.* 2006). It is conceivable that modifications on those tails modulate interactions between nucleosomes, which might have a regulatory function in array folding, and therefore modulate the accessibility of chromatin.

Chromatin structure *in vivo* is difficult to visualize, but in recent years, evidences have piled up to indicate that it is packed at a level higher than the 30nm fiber and such organization is crucial for long-range control of gene-specific transcription (reviewed in (Tremethick 2007)). The higher-order chromatin structures can range between 60-80 nm in interphase chromatin, and finally form the 500-750nm metaphase chromatids during mitosis, which are stabilized by the condensin complex (reviewed in (Tremethick 2007)).

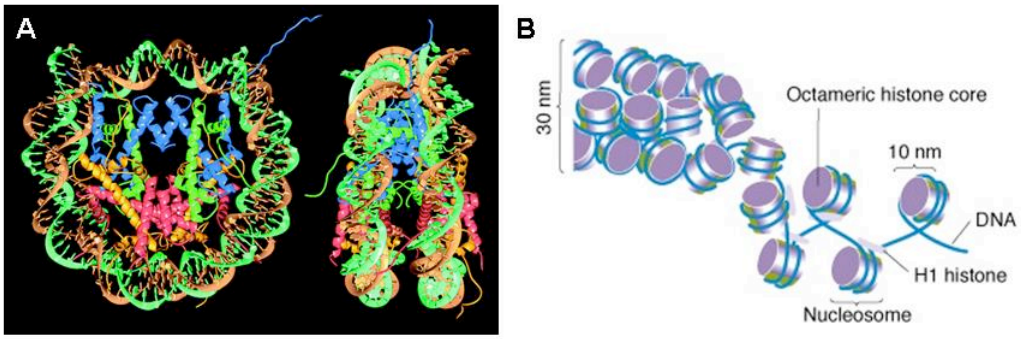



**Figure 1:** Packaging of DNA. A) Structure of the nucleosome core particle as determined by X-ray crystallography at 2.8Å (Luger et al. 1997). B) The organization of DNA into nucleosomes and higher order chromatin structure.

The compaction of interphase chromatin is not uniform within a single nucleus. Some regions are packed only very lightly and some parts of the genome are highly condensed. Based on electron microscopic observations of light or dark staining regions of interphase chromatin, these parts of the genome were termed 'euchromatin' and 'heterochromatin', respectively (Heitz 1928). Euchromatin is generally more sensitive to nuclease digestion and easily transcribed, while heterochromatin is less accessible for biochemical processes and contains fewer genes (Dillon *et al.* 2002). Constitutive heterochromatin is present in all cell types usually in the repeat-rich sequences around the centromeres of chromosomes, and at telomeres. Facultative heterochromatin consists of regions in the genome that are variably silenced between cell types, such as the inactive X chromosome in female mammals.

2.1.2. Mechanism of transcription initiation

All messenger RNA (mRNA) in eukaryotic cells is synthesized by RNA Polymerase II (Pol II). Before transcription can initiate, sequence specific activators have to bind at enhancers and upstream elements of the core promoter of a gene. Then coactivators, (such as chromatin-remodeling enzymes and the large Mediator complex), are recruited to increase the accessibility of DNA for general transcription factors (GTFs) (Thomas *et al.* 2006). Next, the pre-initiation complex is formed by the regulated binding of Pol II, TFIID, TFIIA and TFIIB to the promoter (Figure 2A). Finally, the DNA is melted and Pol II is released to start transcribing, which is mediated by phosphorylation of the carboxy-terminal domain (CTD) of Pol2 by TFIIH. Together the proteins required for the initiation of transcription make up a large complex of over 3 million Daltons (reviewed in (Kornberg 2007)). It is easily conceivable that the promoter sequence of a gene to be transcribed must be made free of nucleosomes for activator binding to occur. Indeed, it has been shown that nucleosomes can have a negative effect on such binding and transcription elongation *in vitro* (Lorch *et al.* 1987; Izban *et al.* 1991). Thus, chromatin, besides playing other regulatory roles, has a 'physical' role in regulating transcription. Such regulation of gene expression mostly occurs at the step of Pol II binding. However, transcription can also be regulated downstream of Pol II binding, since some genes have Pol II bound at their promoter, yet are not transcribed (Muse *et al.* 2007; Zeitlinger *et al.* 2007). This form of regulation can facilitate rapid activation of genes, as is the case for heat-shock genes in *Drosophila* (Lis *et al.* 1993) . It is also possible that polymerase

stalling allows a better coordination of the activation of key regulatory genes during development.

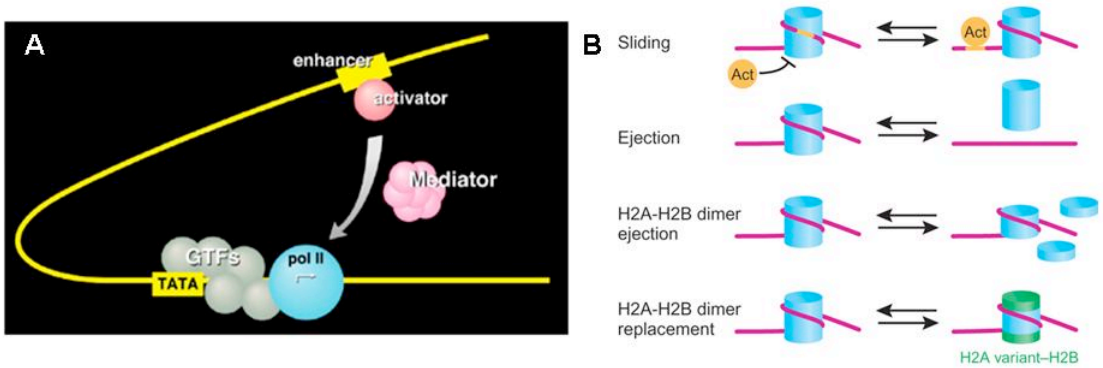



Figure 2: A) Illustration shows the transduction of regulatory information from a gene activator protein bound to an enhancer DNA element to the Pol II transcription machinery at a promoter. GTFs=general transcription factors, pol II=RNA polymerase II. Adapted from (Kornberg 2007). B) Chromatin remodels can allow access of activators (Act) to DNA in different ways, such as sliding, ejection or H2A-H2B dimer removal. Adapted from (Cairns 2007).

## 2.2. Chromatin dynamics

Nucleosomes are not stably bound all the time at a certain DNA sequence, instead they can be removed from the DNA or move along it in order to make DNA accessible. In addition, nucleosomes can incorporate different variants of histones and histones are posttranslationally modified in different ways, most of which are related to chromatin-templated events like transcription, replication and DNA repair.

### 2.2.1. ATP-dependent nucleosome remodeling

Nucleosomes are distributed across the whole genome with a certain preference for some DNA sequences versus others (Segal *et al.* 2006). Regions of the genome which are in the linker sequence between nucleosomes or on their surface are more accessible than regions buried inside the nucleosome. Therefore, nucleosomes often have to be removed for the binding of sequence specific factors, which is required for most chromatin templated events. This can be achieved by chromatin remodeling complexes, using the energy of ATP hydrolysis to alter histone-DNA interactions. Nucleosomes can be removed completely or slid to either side along the DNA (Figure 2B). Nucleosome remodeling complexes can not only increase, but also decrease the accessibility of DNA (reviewed in (Cairns 2007)). All remodeling enzymes belong to the Swi2/Snf2 super-family of helicases (due to their shared ATPase domain), and are divided into several subfamilies, which are conserved in all eukaryotes (Flaus *et al.* 2004).

*Drosophila* ISWI (imitation Swi/Snf) is the catalytic subunit of three remodeling complexes (ACF, CHRAC and NURF) (Ito *et al.* 1997; Mizuguchi *et al.* 1997; Varga-Weisz *et al.* 1997). Null mutation of ISWI results in larval lethality and decondensation of the male X chromosome (Deuring *et al.* 2000). The hyperactivated male X is also hyperacetylated at Lysine 16 of Histone H4 (H4K16). Blocking H4K16 acetylation in ISWI mutants rescues the chromatin structure of the male X chromosome (Corona *et al.* 2002). This suggests that acetylation of H4K16 residues interferes with ISWI-mediated chromatin compaction on the male X chromosome.

CHD-type remodeling complexes contain a pair of chromodomains and consist of several subgroups. The Chd1 subgroup associates with sites of active transcription, where it seems to promote permissive chromatin structure (Krogan *et al.* 2003; Srinivasan *et al.* 2005). The Chd3/4 (Mi-2) subgroup resides in a nucleosome remodeling histone deacetylase (NuRD) complex, which might render N-terminal histone tails accessible for modification. It localizes to sites of active transcription through interaction with methylated Histone 3 Lysine 36 (H3K36) at the 3' end of genes. Thereby, it promotes a compact chromatin structure to prevent transcription initiation from intragenic sequences (Carrozza *et al.* 2005; Joshi *et al.* 2005).

SWI/SNF (mating type switching/sucrose non-fermenting) -type chromatin remodeling complexes are believed to facilitate the binding of activators to nucleosomal DNA by eviction of nucleosomes from the promoter. Indeed, SWI/SNF facilitates Gal4 binding to nucleosomal DNA in an ATP-dependent reaction (Cote *et al.* 1994). Chromatin decompaction by SWI/SNF can be stimulated by histone acetylation to facilitate transcriptional elongation (Carey *et al.* 2006). SWI/SNF cooperates with the histone chaperone Asf1 in *Drosophila* (Moshkin *et al.* 2002), and has been shown to be involved in nucleosome eviction at the PHO5 promoter in yeast (Reinke *et al.* 2003). Depletion of nucleosomes at active promoters occurs throughout the genome, as revealed by recent genome-wide analysis of nucleosome occupancy (Lee *et al.* 2004; Yuan *et al.* 2005; Mavrich *et al.* 2008).

### 2.2.2. Replacement histones

Every histone, except H4, has at least one variant form which is incorporated into chromatin in place of the canonical histone variant in a highly localized way. For example, the nucleosomes at centromeres contain CenH3 instead of H3, whereby the kinetochores can be distinguished from the chromosome arms. Some histone variants

are very similar to the canonical histone in their amino acid sequence. For example, H3.3 differs from H3 in only 4 amino acids. The major difference is that H3.3 is expressed outside of S phase and incorporated into DNA in a transcription-dependent manner, while H3 can only be incorporated during S phase in a replication dependent way (reviewed in (Henikoff 2008)). The replication coupled incorporation of H3 requires the N-terminal tail of the histone. H3 was isolated from the CAF1 histone chaperone complex which interacts with proteins at the replication fork, such as proliferating cell nuclear antigen (PCNA) (Loyola *et al.* 2004). H3.3 on the other hand, was co-purified with the replication independent histone chaperone HIRA, and can be incorporated without its N-terminal tail (Tagami *et al.* 2004). This suggests that the two variants are incorporated by different pathways at different times during the cell cycle. Recent evidence shows that upon transcription induction, displacement of nucleosomes containing H3 is counteracted by incorporation of nucleosomes containing H3.3 throughout the coding regions of all active genes (Mito *et al.* 2005; Wirbelauer *et al.* 2005). This mechanism might mark active genes to provide a transcriptional memory.

H2A has several variants, the most common to all species is H2A.Z. Its levels on chromatin correlate with transcription. H2A.Z is located predominantly at the ends of genes. While it is located just downstream and upstream of the promoter in yeast (Albert *et al.* 2007), it is only found upstream of the transcription start site at *Drosophila* genes (Mavrich et al. 2008). This might be functionally related to paused RNA polymerase which is found engaged at many promoters in *Drosophila* irrespective of the transcription state of the gene (Zeitlinger et al. 2007).

### 2.2.3. Posttranslational histone modifications

Histones are subject to at least 8 different kinds of covalent posttranslational modifications, primarily on their N-terminal tails, including acetylation, methylation, phosphorylation, and ubiquitination (Figure 3A). There are at least 60 residues on histones which can be modified in many different ways. Those modifications are not randomly distributed across the genome. Recent genome-wide studies of several histone modifications have shown that some modifications promote active chromatin while others occur in silent regions of the genome. In addition, certain modifications, such as the phosphorylation of Serine 10 on H3 during mitosis, occur in a cell-cycle dependent manner (reviewed in (Kouzarides 2007)). Histone modifications have first been mapped at the resolution of genes, showing that many modifications often coincide

on the same gene (Schubeler *et al.* 2004). The first genome-wide maps of histone modifications at a resolution of single nucleosomes have been generated in the yeast genome (Liu *et al.* 2005a). Recently, histone modifications have been mapped at high resolution in mammalian genomes (Barski *et al.* 2007; Mikkelsen *et al.* 2007). These maps were generated by combining chromatin immunoprecipitation with an antibody raised against a certain modification and microarray or high throughput sequencing technology. Such high resolution maps revealed characteristic distributions of different histone modifications (Figure 3B). For example, some modifications are localized specifically at the promoter of genes, while others are enriched at the 3' end (reviewed in (Kouzarides 2007)).

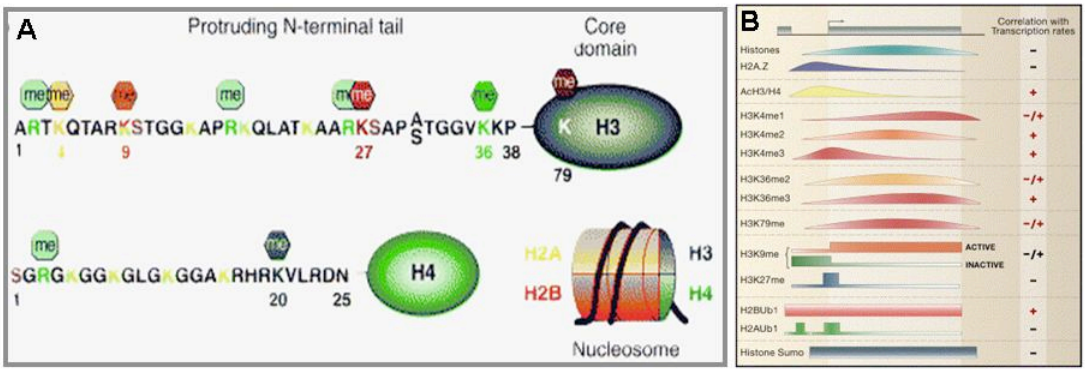



**Figure 3**: A) Selected posttranslational modifications along N-terminal tails of histones H3 and H4. Indicated are sites of lysine (K – color-coded) and arginine (R – turquoise) methylation, lysine acetylation (K – green letters) and serine (S – violet letters) phosphorylation. Only the mono-methylated states are presented. H3K9 can either be methylated or acetylated. Adapted from (Peters *et al.* 2005). B) Genome-Wide Distribution Pattern of Histone Modifications from a Transcription Perspective. The distribution of histones and their modifications are mapped on an arbitrary gene relative to its promoter (5′ IGR), ORF, and 3′ IGR. The curves represent the patterns that are determined via genome-wide approaches. The squares indicate that the data are based on only a few case studies. With the exception of the data on K9 and K27 methylation, most of the data are based on yeast genes. Adapted from (Kouzarides 2007).

Several models have been put forward to explain the function of histone modifications in gene regulation. It is known that histone acetylation or phosphorylation can change the overall charge of the chromatin. The acetylation of histones neutralizes positive charges of histones and phosphorylation adds a negative charge to chromatin. The charge neutralization model suggests that histone acetylation leads to a decondensation of the chromatin fiber by destabilizing the interaction among nucleosomes and between nucleosomes and DNA. Indeed, there is evidence that histone acteylation can relax chromatin structure *in vivo* and *in vitro* (Wolffe *et al.* 1999; Shogren-Knaak et al. 2006).

Since in addition to acetylation, so many lysine residues can be modified in three different methylation states, it was proposed that there is a combinatorial complexity of histone modification patterns, resulting in a "histone code". The histone code hypothesis states that multiple histone modifications act in combination to form a "code" which regulates downstream functions on chromatin (Jenuwein et al. 2001).However, several studies have shown that histone modifications promoting a certain activity state often occur on the same genes and even nucleosomes (Schubeler et al. 2004; Liu et al. 2005a; Pokholok *et al.* 2005), suggesting that the level of complexity is lower than predicted from the number of possible modifications. Instead, it is possible that these modifications are partially redundant, ensuring a robust chromatin regulation (Schreiber *et al.* 2002). The signaling-pathway model postulates that histone modifications can serve as signaling platforms onto which enzymes bind for their function on chromatin (Li *et al.* 2007). This also suggests that multiple histone modifications can provide bi-stability, robustness and specificity through feedback loops, redundancy and combination. In the following section I will focus on the most widely studied histone modifications, acetylation and methylation.

2.2.3.1. Histone acetylation

Acetylation can occur at multiple lysine residues of histones H3, H4 and H2A and is almost always associated with an active chromatin state, promoting transcription and possibly also the initiation of DNA replication (reviewed in (Chakalova *et al.* 2005)). Acetylation changes the overall charge of histones and thereby is thought to reduce the interaction between the highly basic histone proteins, especially histone tails, and the negatively charged DNA. This could then lead to an overall loosening of the DNA packaging into nucleosomes and thereby promote the binding of other proteins, such as transcription factors. Interestingly, the structure of the nucleosome suggests an interaction between the tail of histone H4 and H2A/H2B at an adjacent nucleosome, which might promote a tighter chromatin structure (Luger et al. 1997). Indeed, deletion of different histone tails showed that the histone H4 tail, and specifically the region from amino acid 14-19, mediates chromatin compaction *in vitro* (Dorigo et al. 2003). Furthermore it has been shown by chemically ligating modified tail peptides to core histones, that acetylation of lysine 16 on histone H4 interferes with this higher order chromatin folding leading to decompaction of chromatin (Shogren-Knaak et al. 2006). Different combinations of mutations of the four acetylable lysine residues on the histone

H4 tail resulted in cumulative effects on gene expression in yeast, with the exception of H4K16. Mutation of this lysine residue had additional specific consequences for transcription (Dion *et al.* 2005). This implies a cumulative mechanism of acetylation for transcriptional activation, and an additional specific role for H4K16 acetylation. As suggested above, this might be due to direct effects of this modification on chromatin structure. Alternatively, acetlylated H4K16 could recruit specific binding molecules, which then regulate gene expression.

Histone acetylation is generally very dynamic. Acetylation levels are increased by enzymes called histone acetyl transferases (HATs) and removed by histone deacetylases (HDACs). There are many different HATs and HDACs, which target different lysine residues on histones and sometimes also other proteins. Most of these enzymes modify more than one lysine residue while some are specific for individual lysines. HATs and HDACs often show broad activity, but usually they are part of larger complexes, which confer specificity to defined regions on chromatin. HATs are divided in three main families, GNAT, MYST and CBP/p300 (reviewed in (Kouzarides 2007)).

Histone H4 acetylation at lysine 16 (H4K16ac) in *Drosophila* can be catalyzed by MOF (male absent on first), which acts in a complex with MSL (male specific lethal) proteins (Gu *et al.* 1998), or the recently identified ATAC2, which is a part of the ATAC complex (Suganuma *et al.* 2008). Note that in *Drosophila* males, compensation of sex-specific differences in X-linked gene dosage is achieved by doubling the expression from the single X chromosome and involves the activity of MOF in the MSL complex (Lucchesi *et al.* 2005; Mendjan *et al.* 2007; Straub *et al.* 2007). The dosage compensated X chromosome is associated with highly elevated levels of H4K16ac as measured by immunofluoresence, polytene staining, and chromatin-IP (Turner *et al.* 1992; Smith *et al.* 2001a; Lucchesi et al. 2005; Gilfillan *et al.* 2006; Kind *et al.* 2008). The male X chromosome also seems to be decondensed (Lucchesi et al. 2005), which further underscores the role of H4K16ac in chromatin decompaction as described above. H4K16ac in generally enriched at active genes, especially at their promoter. H4K16ac at dosage compensated genes, however, increases towards the 3' end (Bell *et al.* 2007; Bell *et al.* 2008; Kind et al. 2008). Depletion fo MOF affects H4K16ac and transcription of autosomal and X-linked genes (Kind et al. 2008). This suggests an important role for H4K16ac not only in gene activation, but also in the two-fold upregulation of genes on the male X chromosome.

2.2.3.2. Histone methylation

Histones can be methylated at lysine or arginine residues, mostly on the N-terminal tails of H3 and H4. Lysine can be mono-, di-, or trimethylated, while arginine can be mono- or dimethylated (symmetric or asymmetric). Methylation is more stable than acetylation, and for a long time was thought to be irreversible unless diluted through DNA replication. Furthermore, unlike acetylation, it does not change the overall charge of histones, but instead it functions in recruiting effector proteins to chromatin, which then conduct enzymatic activities such as chromatin remodeling. These proteins can bind to methylated residues via different conserved domains, such as chromodomains (Lachner *et al.* 2001) or PHD domains (Wysocka *et al.* 2006). Lysine methylation is carried out by lysine-methyltransferases, most of which carry a SET domain (reviewed in (Kouzarides 2007)), and can be removed by the recently identified lysine-demethylases, such as LSD1 (Shi *et al.* 2004). Depending on the modified residue, histone methylation can have an activating or repressing effect on transcription. At least 24 sites of methylation have been identified. The best characterized activating sites are H3K4, H3K36 and H3K79, while H3K9, H3K27 and H4K20 are thought to be repressive.

H3K4 trimethylation is located specifically at active promoters where it activates transcription, while H3K4 mono- and dimethylation gradually decrease towards the 3' end of active genes in yeast (Pokholok et al. 2005). This distribution is very similar to metazoa, except that H3K4 methylation is not exclusively at promoters of active genes in mammalian cells (Roh *et al.* 2006). In metazoa, the H3K4 methyltransferases Trithorax (TRX) and ASH1 function as anti-repressors (Klymenko *et al.* 2004), suggesting that H3K4 methylation propagates an active transcription state throughout cell division by blocking repressive histone modifications (reviewed in (Schwartz *et al.* 2007)).

H3K36 methylation was shown to localize to active genes, where the dimethylated state (H3K36me2) is distributed throughout the gene except for the promoter (Bell et al. 2007), and the trimethylated state (H3K36me3) is enriched at the 3' end of genes (Barski et al. 2007; Bell et al. 2007). In *Drosophila* cells, different H3K36 methylation states are catalyzed by specific enzymes, dMes4 for H3K36me2, and dHypB for H3K36me3 (Bell et al. 2007). In yeast, H3K36me3 was shown to suppress initiation from cryptic transcription start sites in coding regions by recruiting HDACs to deacetylate histones after the passage of the RNA polymerase (Carrozza et al. 2005; Joshi et al. 2005).

Only very little is known about H3K79 methylation. It is laid out by Dot1, which is the only known methyltransferase without a SET domain. It locates to transcribed regions, is

involved in the activation of HOXA9 and probably has an indirect role in maintaining heterochromatin by limiting the spreading of heterochromatic proteins to euchromatin (reviewed in (Kouzarides 2007)).

Methylation of H3K27 depends on Polycomb (PcG) complexes, which were initially identified in *Drosophila* as regulators of homeotic (Hox) genes, together with Trithorax (Trx) complexes, which are required for H3K4 methylation (see above). Further studies have shown that PcG complexes are negative regulators of transcription, while Trx proteins are associated with gene activation. There are three polycomb complexes: PRC1, PRC2, which contains E(Z) (the H3K27 methylase), and the PhoRC complex. H3K27 trimethylation (H3K27me3) is mostly a repressive chromatin mark, while H3K27 di- and monomethylation are not yet well understood (reviewed in (Schwartz et al. 2007)). Recent genome-wide profiles of H3K27me3 in *Drosophila* (Schwartz *et al.* 2006; Tolhuis *et al.* 2006) and mammalian cells (Barski et al. 2007; Mikkelsen et al. 2007) contributed to our knowledge of this histone modification. Interestingly, in *Drosophila* this modification occurs mostly in rather large regions, covering not only the Hox gene clusters, but many developmentally regulated genes (Schwartz et al. 2006; Tolhuis et al. 2006). PcG proteins localize to the same regions as H3K27me3, but with higher specificity to Polycomb Response Elements (PREs) within those regions (Schwartz et al. 2006; Tolhuis et al. 2006). This suggests that PREs are the initial target sequences of PcG proteins, but E(Z) can methylate a large region around those sites (reviewed in (Schwartz et al. 2007). It is conceivable that this mechanism ensures the propagation of a repressed epigenetic state through cell division, even after the initial cue to silence a gene has gone.

2.2.3.3. Interaction of histone modifications with effector proteins

As mentioned above, one function of histone methylation is the recruitment of effector molecules to chromatin, which in turn can propagate the spread of the methylated state, lead to other histone modifications, or perform certain chromatin remodeling tasks. This can be the case for most histone modifications, such as H3K27 methylation, which interacts with different Polycomb complexes (reviewed in (Schwartz et al. 2007)), or H3K9 methylation and heterochromatin protein 1 (HP1) (Lachner et al. 2001). Below I will discuss the role of HP1 and H3K9 methylation in heterochromatin formation and gene regulation.

The SuVAR3-9 class of HMTs establishes H3K9 di- and trimethylation at constitutive heterochromatin, which is mainly present in the repeat-rich regions around centromeres, and remains condensed throughout the cell cycle (reviewed in (Peters et al. 2005)). HP1 binds this modification via its chromodomain, and then recruits more SuVAR3-9 via its chromo-shadow domain, which leads to the propagation and spreading of heterochromatin (Lachner et al. 2001). Interestingly, cell cycle dependent transcription of pericentric repeats, which is processed into small RNAs by the RNAi machinery, is required for the establishment and propagation of heterochromatin by H3K9 methylation (Volpe *et al.* 2002; Pal-Bhadra *et al.* 2004; Verdel *et al.* 2004; Kloc *et al.* 2008).

While mammalian pericentric heterochromatin contains mostly H3K9me3, H3K9me2 is the predominant mark of heterochromatin in *Drosophila* (Schotta *et al.* 2002). The SuVAR3-9 HMT in *Drosophila* is a suppressor of position effect variegation (PEV), which describes the observation that euchromatic genes, when translocated near heterochromatin, show a variegated expression pattern (reviewed in (Schotta *et al.* 2003)). This suggests that the role of H3K9 methylation lies in gene silencing. Interestingly however, there are hundreds of genes which are embedded in heterochromatin and rely on this location for their expression (reviewed in (Yasuhara *et al.* 2006)). This implies that certain genes directly depend on H3K9 methylation and or HP1 binding for their expression, for example by possessing a different kind of promoter structure. Alternatively, H3K9 and HP1 might be required to silence nearby repetitive sequences. Consequently, the loss of heterochromatin structure would result in activation of repeats and thereby silence genes. However, transferring those genes into euchromatin abolishes their transcription, and a recent study does not support a specific promoter structure of heterochromatic genes (Yasuhara *et al.* 2005). Therefore, it seems possible that heterochromatin specific proteins facilitate a certain type of long range enhancer-promoter interaction, on which those genes depend. A recent profile of H3K9me2 across a large portion of the *Drosophila* genome showed that, besides transposable elements, this modification is also enriched at heterochromatic genes, except at their promoter, where H3K9 is acetylated (Yasuhara *et al.* 2008). H3K9 methylation is not restricted to heterochromatin. It has also been found in coding regions of active genes in mammals (Vakoc *et al.* 2005). In *Drosophila* cells, SuVAR3-9 localizes predominantly to pericentric repetitive sequences, where it recruits HP1, but can also be found at a few heterochromatic and euchromatic genes (Greil *et al.* 2003).

Not all H3K9 methylation depends on SuVAR3-9. Mutation of this enzyme does not erase, but rather increases H3K9 methylation at heterochromatic genes (Yasuhara et al. 2008), even though global H3K9 methylation levels in the nucleus decrease. dG9a is another methyltransferase known to methylate H3K9 (Mis *et al.* 2006), and recent work has shown that H3K9 methylation on chromosome 4 in *Drosophila* is established by dSETDB1 (Seum *et al.* 2007; Tzeng *et al.* 2007). HP1 seems to bind to H3K9 methylation independent of the enzyme which established the modification, since HP1 can bind many, mostly non-pericentric regions of the genome independent of SuVAR3-9 (Greil et al. 2003; de Wit *et al.* 2005; de Wit *et al.* 2007). Furthermore, on chromosome 4, HP1 binding depends on dSETDB1 (Seum et al. 2007; Tzeng et al. 2007). However, it is also possible that in some places HP1 binds chromatin independently of H3K9 methylation.

HP1 binding is inhibited by Serine 10 phosphorylation, which is established by the Aurora B kinase at the beginning of mitosis (Fischle *et al.* 2005; Hirota *et al.* 2005). This suggests that HP1 binding has to be re-established after mitosis, potentially resulting in easy reversibility of HP1 binding. In *Drosophila,* next to HP1 (also called HP1a), there are several other heterochromatin proteins: HP1b and HP1c, which is thought to localize mostly to euchromatin. HP2, HP3, HP4, HP5 and HP6 are targeted to heterochromatin by HP1 and seem to be suppressors of position effect variegation, suggesting that they contribute to heterochromatin function (Shaffer *et al.* 2002; Greil *et al.* 2007).

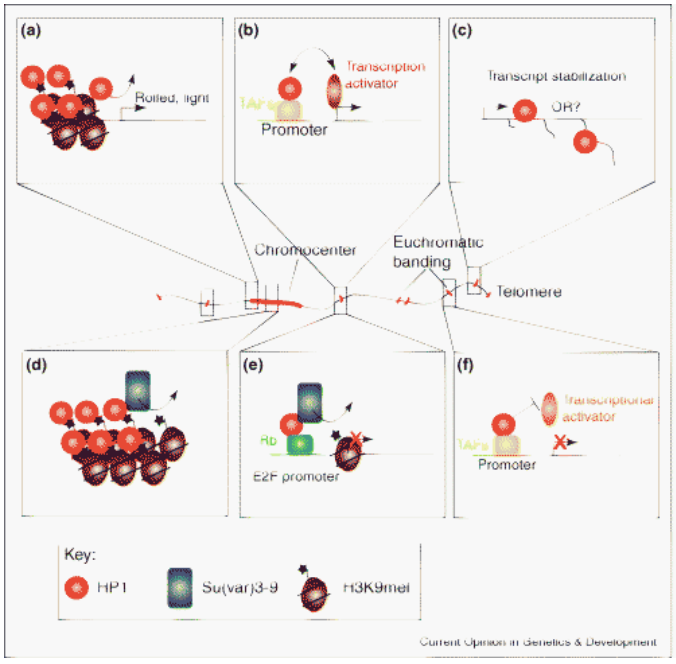



**Figure 4: Multiple roles of HP1. (a)** At heterochromatin, HP1 co-localizes with Su(var)3-9 and H3K9me, promoting a compact chromatin structure. **(b)** At genes regulated by the cyclin E promoter, HP1 can be recruited through association with Retinoblastoma, promoting histone methylation and gene repression. **(c)** HP1 interacts with factors associated with the general transcriptional machinery, possibly blocking the later association of activators. **(d)** Heterochromatic genes require a compact chromatin structure for their normal expression level. **(e)** The association of HP1 at promoters with transcription factors could promote activation through recruitment of co-activators. **(f)** HP1 can bind throughout coding regions, where its positive effect on their transcription suggests a role for HP1 in stabilizing mRNA transcripts. Adapted from (Hediger *et al.* 2006).

Overall, the classical view that H3K9 methylation and HP1 binding are strictly repressive has recently changed. In fact, it seems as if the choice between repression and activation for this modification is context and location dependent. It is conceivable that the interactions of HP1 with different proteins contribute to its diverse functions.

2.3. The timing and initiation of DNA replication

Each time a cell to divides it must first duplicate its genome in an organized and error-free manner. In eukaryotic cells, DNA replication starts at many sites throughout the genome termed origins of replication, and then proceeds during the S phase of the cell cycle. DNA replication is a highly regulated process, ensuring that each sequence replicates once, and only once, during every cell cycle. This regulation occurs mostly at the level of the initiation of DNA replication: The more origins are used, the faster the replication of the whole genome will be finished (Figure 5A). Some cells, as in the early embryos of *Xenopus laevis* and *Drosophila melanogaster*, can replicate their genome in only 20 minutes. However, the typical length of S phase for cells of higher eukaryotes is about 10 hours. Not all regions of the genome replicate at the same time. Instead, some regions replicate at the beginning of S phase, others more towards the end. This temporal order of DNA replication is highly conserved between consecutive cell cycles in a given cell type. The replication timing of a certain sequence depends on its distance to the closest origin and the time during S phase at which this origin is activated (reviewed in (Gilbert 2004; Aladjem 2007)).

More than 40 years ago, Francois Jacob and Sydney Brenner postulated the "replicon model" to explain the initiation of DNA synthesis in bacteria (Jacob *et al.* 1963). They proposed that replication was regulated by an initiator protein, which is encoded by a structural gene, and a specific genetic element termed "replicator", which is recognized by the initiator to start DNA synthesis at this site. The model, which proved to be true for prokaryotes, was later adapted for eukaryotic cells, where multiple replicators would exist on each chromosome. Indeed, origins of replication (replicators) in budding yeast share a consensus sequence and their location has been mapped throughout the genome (MacAlpine *et al.* 2005; Nieduszynski *et al.* 2006), and initiator proteins have been identified, which are highly conserved from yeast to man. However, higher eukaryotes often lack genetically defined replicators (reviewed in (Gilbert 2004)). Therefore it has been proposed that epigenetic features define the initiation of DNA replication in metazoan cells (Gilbert 2004; Aladjem 2007).

Replication initiation in eukaryotes proceeds in two temporally distinct steps during the cell cycle. First, origins are licensed for replication during G1 phase. This involves the binding if the pre-replicative complex (pre-RC) to multiple sites across the genome. The regulation of pre-RC formation is important for the coordination of DNA replication with the cell cycle. During S phase, many, but not all, pre-RCs are activated (Figure 5B). This occurs via the assembly of additional replication factors, which facilitate DNA unwinding and lead to the binding of DNA polymerases (reviewed in (Bell *et al.* 2002)).

2.3.1. Pre-replicative complex (pre-RC) formation (G1 phase)

DNA replication begins with the assembly of a pre-RC consisting of at least 14 different proteins. This process is called origin licensing.The origin recognition complex (ORC) is a six-subunit ATPase complex (consisting of ORC1-6 proteins) that acts as the initiator in eukaryotic cells. It was originally identified in *S. cerevisiae* as binding to the autonomously replicating sequence (ACS) representing yeast origins of replication (Bell *et al.* 1992), and is highly conserved in all eukaryotes. As the first pre-RC component to bind DNA, ORC associates with replication origins in early G1 phase, and its binding is required for the initiation of DNA replication (Figure 5B). In mammalian cells ORC2-5 form a core complex, with ORC 1 and 6 only loosely associated. In addition, ORC6 is not required for the DNA binding of the ORC complex to DNA. In *Drosophila* however, ORC6 is tightly associated with the other ORC proteins, and is essential for DNA binding (reviewed in (Sasaki *et al.* 2007)). ChIP-chip studies have shown that the *Drosophila* ORC locates preferentially to AT-rich sequences (MacAlpine *et al.* 2004), which seems to be mediated by the ORC6 subunit (Balasov *et al.* 2007). Nevertheless, no consensus sequence of ORC binding to DNA has been found in *Drosophila* (MacAlpine et al. 2004). Rather, ORC seems to bind preferentially to negatively supercoiled DNA, suggesting that DNA topology is more important for ORC binding than DNA sequence (Remus *et al.* 2004).

Besides its role in the initiation of DNA replication, ORC has several additional functions. It has been implicated in the establishment of transcriptionally silent chromatin at the budding yeast silent mating type loci, where it recruits Sir1 (Bell *et al.* 1993; Hou *et al.* 2005; Hsu *et al.* 2005). Similarly, in *Drosophila* and in human cells, ORC2 interacts with HP1, and its mutation or knock-down results in reduced heterochromatin spreading due to HP1 de-localization (Pak *et al.* 1997; Prasanth *et al.* 2004). Besides a direct function of ORC in silencing, it is also possible that the highly condensed state of

heterochromatin makes this part of the genome more difficult to replicate, and therefore requires more ORC (Leatherwood *et al.* 2003). Additional functions of ORC have been implicated in many more processes. For example, it has been implicated in mitotic chromosome condensation and centromere function (reviewed in (Sasaki et al. 2007)). In addition, it establishes sister chromatid cohesion in *S. cerevisiae* independently of the cohesin complex or its role in DNA replication (Shimada *et al.* 2007).

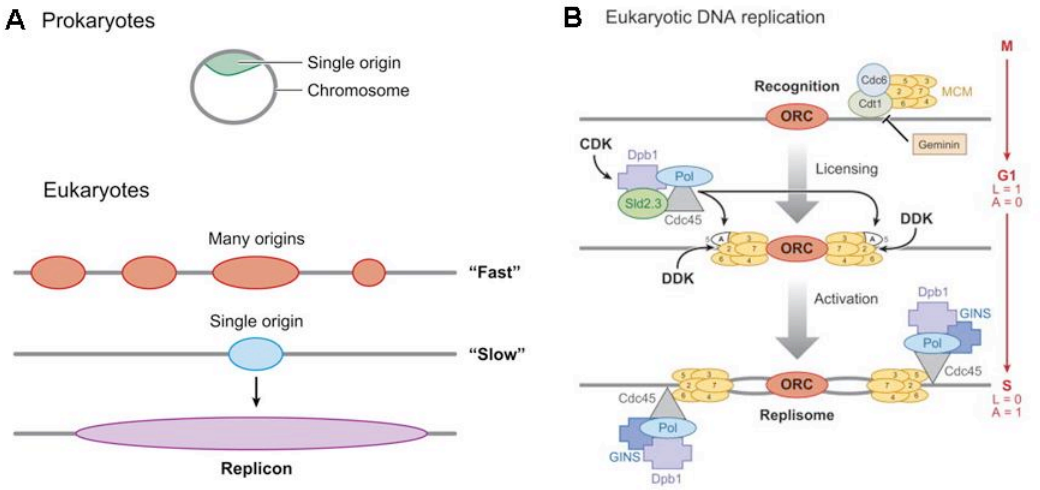



**Figure 5: The initiation of DNA replication. A)** Prokaryotes have a single origin on a circular chromosome (*above*). **B)** In eukaryotes, multiple origins are found on a single chromosome. When replication is "fast," many origins are used, whereas only one origin is used in this region when replication is "slow". Replication proceeds bidirectionally from an origin to form a replicon (*below*). **C)** An origin is recognized by ORC, then Cdc6 and Cdt1 protein load the hexameric MCM helicase to form the "licensed" (L) pre-RC in G1 phase (L = 1, A = 0). Geminin inhibits Cdt1 and pre-RC formation. CDK and DDK become active in late G1, activate (A) the MCM helicase and load on the replisome that contains the DNA polymerases. In addition, CDK inhibits any further licensing (L = 0, A = 1). Toward this end, CDK phosphorylates Sld2 and Sld3 proteins and DDK phosphorylates MCM proteins, which "pushes out" the "A" domain of Mcm5. Adapted from (Sclafani *et al.* 2007).

For its role in the initiation of DNA replication, ORC recruits Cdc6 to chromatin, which is rapidly followed by the binding of Cdt1. The final step in pre-RC formation during late G1 phase is the recruitment of the hexameric MCM2-7 complex to origins, which requires the ORC, the licensing cofactor Cdt1, and the ATPase Cdc6 (Figure 5B). The MCM (mini-chromosome-maintenance) proteins are highly related to each other, yet each one has a unique sequence which is conserved across eukaryotes (reviewed in (Bell et al. 2002)). Furthermore, deletion of a singe MCM subunit is lethal in yeast, suggesting that each of the six MCM proteins has a unique and important function (Dutta *et al.* 1997). The ring-shaped structure of the MCM complex encircles the DNA, which might be achieved by ATP-dependent clamp-loading by ORC and Cdc6. Loading of MCMs seems

to be the most important function of ORC and Cdc6, because once MCMs are loaded onto chromatin, replication can start even in the absence of ORC and Cdc6 (reviewed in (Machida *et al.* 2005)). The MCM proteins function not only in the initiation of DNA replication, but also during replication fork progression, where the MCM complex is believed to be the replicative helicase which moves along the replication fork and unwinds the DNA (Figure 5B). This is supported by ChIP analysis in *S. cerevisiae* showing that several MCMs associate with origins and origin-proximal sequences in a temporal manner similar to the replicative DNA polymerase (Aparicio *et al.* 1997). In addition, the MCMs form a hexameric complex around DNA, which shows ATPase and DNA helicase activity. While yeast MCMs are restricted to the sites of ORC binding, in mammalian cells they have been shown to spread out several kilobases (kb) from its initial loading site (reviewed in (Bell et al. 2002)). This might be a reason for the so often observed broad initiation zones in higher eukaryotes.

Only origins which have been licensed for replication by pre-RC binding can fire during the subsequent S phase. This licensing step is highly regulated, and separated from the actual replication initiation occurring in S phase (see below). Re-replication is prevented by inhibiting pre-RC assembly after the entry into S phase and until the end of mitosis. This is achieved through the action of CDKs (cyclin dependent kinases) and geminin, and by replication-dependent origin inactivation. Geminin, which is absent in yeast, inhibits pre-RC formation through interaction with Cdt1 (Figure 5B). It is degraded by anaphase-promoting-complex mediated ubiquitination during mitosis to allow pre-RC formation in G1. On the other hand, the activation of MCMs at the entry into S phase depends on increased CDK levels, therefore preventing origin firing during pre-RC assembly in G1 (reviewed in (Machida et al. 2005)). This tight regulation by multiple mechanisms allows each sequence of the genome to be replicated only once in each cell cycle, which is crucial for correct genome duplication ands cell division.

### 2.3.2. Replication fork formation and progression of DNA replication (S phase)

After pre-RC formation, the origin has to be activated through formation of a bi-directional replication fork. Origins can be activated at different times during S phase, and not all origins which were licensed for initiation by the pre-RC are finally activated. The excess licensed origins function only in case of replicative stress. When replication forks collapse, excess MCMs are required for the completion of DNA replication (Woodward *et al.* 2006). Those MCMs need to be put in place already in G1, because

origin licensing is inhibited in all other stages of the cell cycle to prevent re-replication (reviewed in (Machida et al. 2005)). The mechanism of origin activation is still poorly understood. Below I will give a summary of known steps and factors involved in this process.

Activation of origins occurs in S phase and requires Cdk2-Cyclin E and the Cbf7-Dbf4 kinase (DDK). The activation begins with the addition of Mcm10 to the pre-RC, which displaces Cdt1. Following this, DDK phosphorylates Mcm3-7, which activates the helicase. DDK and the CDK complex then recruits another protein called Cdc45 (Figure 5B). Loading of Cdc45 is an important step in origin activation, since it is required for activation of MCM2-7 on origins and chromosome unwinding at the replication forks (reviewed in (Masuda *et al.* 2003)). Together with the observation that Mcm2-7 does not perform helicase activity on its own and, like MCMs, Cdc45 travels across DNA with the replication fork, this implies Cdc45 as a helicase co-factor. In addition to MCM activation, de-repression of pre-RC components might be important as well. Indeed, pRB can suppress MCM or ORC activity (Bosco *et al.* 2001; Gladden *et al.* 2003), and acetylation of Mcm3 suppresses helicase activation until S phase (Takei *et al.* 2002). Ultimately, the DNA polymerases have to assembly at the origins. Cdc45 assembles many components of the replication machinery, including RPA, PCNA and DNA polymerases $\alpha$ and $\varepsilon$ at the site of initiation. At this stage the origin fires and DNA synthesis begins (reviewed in (Machida et al. 2005)).

At the start of replication, helicases unwind the DNA duplex, which results in short stretches of single-stranded DNA. This is stabilized by the single-strand binding protein RPA. Replication proceeds in both directions by a different mechanism. The leading strand is replicated continuously from 5' to 3' by DNA polymerases. Since the DNA polymerases only show 5'-3' processivity, the lagging strand is replicated also 5' to 3', but in a discontinuous fashion producing ~300bp Okazaki fragments (Figure 6). To initiate DNA synthesis, the primase lays down an RNA primer, from which the DNA polymerase $\alpha$ can synthesize a short stretch of DNA. Replication factor C (RFC) then binds and catalyzes the loading of PCNA, the ring-shaped replication factor that encircles DNA. Through this the replicative polymerases $\varepsilon$ or $\delta$ can bind and take over the DNA synthesis from polymerase $\alpha$ (Figure 6). Those enzymes carry out most of the DNA synthesis, and their processivity is enhanced by PCNA (reviewed in (Moldovan *et al.* 2007)). In addition, they contain a 3' to 5' exonuclease proofreading activity to reduce the incorporation of wrong nucleotides into the new DNA strand. The Okazaki fragments

on the lagging strand are finally fused together by DNA ligase, and topoisomerases take care of problems with DNA topology caused by the replication fork (Figure 6).

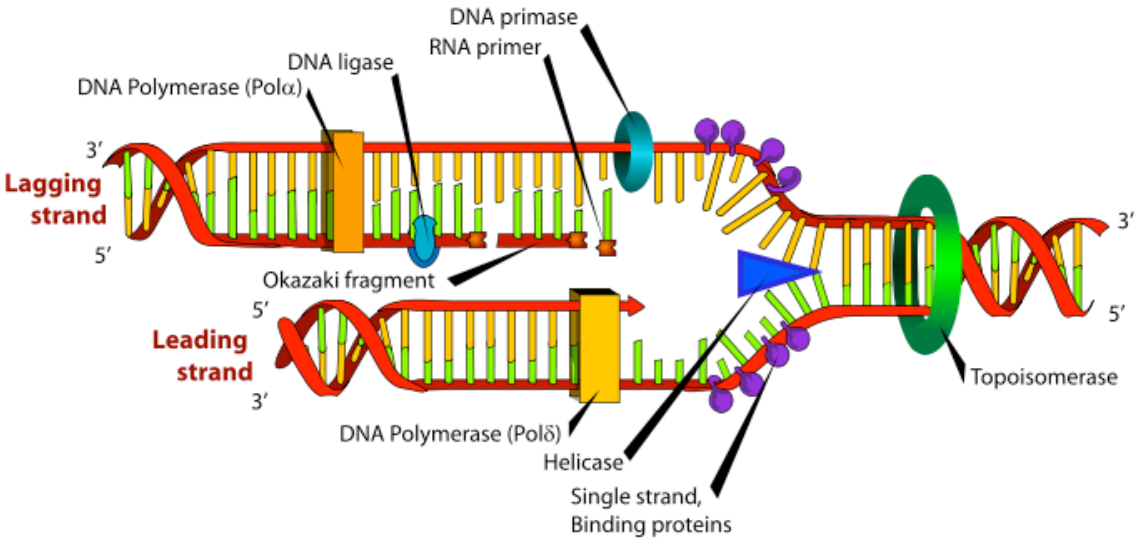



**Figure 6: Mechanism of DNA synthesis.** The leading strand is defined as the new DNA strand at the replication fork that is synthesized in the 5'→3' direction in a continuous manner. When the helicase unwinds DNA, two single stranded regions of DNA (the "replication fork") form. On the leading strand the DNA polymerase is able to synthesize DNA using the free 3' OH group donated by a single RNA primer and continuous synthesis occurs in the direction in which the replication fork is moving. The lagging strand is the DNA strand at the opposite side of the replication fork from the leading strand, running in the 3' to 5' direction. Because DNA polymerase cannot synthesize in the 3'→5' direction, the lagging strand is synthesized in Okazaki fragments. Along the lagging strand's template, primase builds RNA primers in short bursts. DNA polymerases are then able to use the free 3' OH groups on the RNA primers to synthesize DNA in the 5'→3' direction. The RNA fragments are then removed and new deoxyribonucleotides are added to fill the gaps where the RNA was present. DNA ligase then joins the deoxyribonucleotides together, completing the synthesis of the lagging strand. Adapted from http://commons.wikimedia.org/wiki/Image:DNA_replication.svg.

On average, a replication fork replicates ~1-3kb/minute, but fork velocity can vary between different regions of the genome depending on inter-origin distance (Conti *et al.* 2007). After replication forks have traveled a certain distance across the chromosome, they converge with incoming forks from neighboring origins. Forks can pause for some time at certain sites, and fork stalling can occur, in particular during replication of long stretches of repeated sequence (reviewed in (Rothstein *et al.* 2000)). This can cause fork collapse and the cease of DNA replication at that site, in which case it is important to have additional pre-RCs within the unreplicated region to finish DNA replication. The completion of DNA replication of the whole genome, even very late replicating sequences, in a timely fashion is crucial, especially since recent data in *S. cerevisiae* suggest the lack of a checkpoint for the completion of DNA replication at the end of S

phase (Torres-Rosell *et al.* 2007). Interestingly, a delay in replication timing in cancer cells can lead to delayed mitotic chromosome condensation and chromosomal instability (Smith *et al.* 2001b). This again demonstrates that DNA replication and also its correct timing, have to be tightly regulated for proper cell cycle progression and genome stability in dividing cells.

### 2.4. Connection of replication timing to chromatin and transcription

A potential coordination of gene transcription and timing of DNA duplication during S phase had been postulated for a long time as microscopical studies showed that highly compacted heterochromatin replicates late, while gene-rich euchromatin replicates earlier (reviewed in (Gilbert 2002)). Recent advances in microarray technology enabled researchers to investigate this phenomena genome-wide in a quantitative fashion and in various organisms (reviewed in (MacAlpine et al. 2005)). The emerging picture is that actively transcribed genes have a higher likelihood to replicate early in S phase (Schubeler *et al.* 2002; MacAlpine et al. 2004). Interestingly, this intimate connection between the fundamental processes of transcription and DNA duplication is only observed in the complex genomes of higher eukaryotes. Current models of a mechanistic link between both processes involve chromatin structure changes as a major determinant and several lines of evidence support this hypothesis.

#### 2.4.1. Genomic studies of DNA replication timing

The first microarray based genome-wide study of the timing of DNA replication was performed in budding yeast (Raghuraman *et al.* 2001). This fundamental work identified the temporal program of replication as well as the origins of replication and their time and frequency of firing. Rather unexpectedly however no global correlation between transcriptional activity and replication timing was observed even though changes in replication timing depending on SIR mediated transcriptional repression are well established (Stevenson *et al.* 1999; Zappulla *et al.* 2002). Evenly surprising, in *S. pombe,* centromeric heterochromatin replicates early (Kim *et al.* 2003).

In the larger and gene-poor genomes of higher eukaryotes however things appear different. The analysis of 5000 genes in *Drosophila* cells of embryonic origin (Kc) revealed a strong correlation between early replication timing of a gene and its likelihood of being actively transcribed (Schubeler et al. 2002). Interestingly this connection is not absolute. 30% of the earliest replicating genes were found to be inactive and 30% of the

latest replicating genes were actively transcribed suggesting that early replication, while positively correlated with active genes, is unlikely to be a general requirement for transcription (Figure 7A). This can be explained by the large size of replicons that contain several genes, not all of which are active or inactive at a given time (Figure 7B). A more extensive study of transcription and replication of chromosome 2L of *Drosophila* suggested that early replicating domains are defined by integrating the transcriptional status (measured as RNA polymerase binding by Chromatin-IP (ChIP) as well as mRNA levels) over estimated 180 kb (MacAlpine et al. 2004) (Figure 7B). Similar microarray studies in the human genome also revealed that early replicating regions contain more active genes and in mammals appear to overlap with GC rich isochors (White *et al.* 2004; Woodfine *et al.* 2004; Jeon *et al.* 2005). This correlation also holds for the abundant yet previously unnoticed regions that transcribe high levels of non-coding transcripts (White et al. 2004) indicating that early replication of regions of active transcription is a common feature in higher eukaryotes.

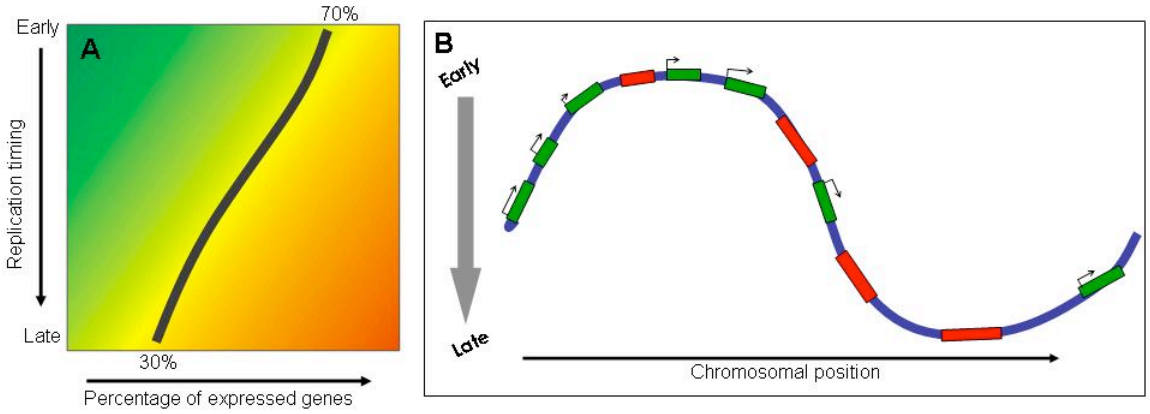



**Figure 7: Genome-wide relation of replication timing and transcription.** A) Microarray analysis revealed the activity and replication timing for a large set of *Drosophila* genes. About 70% of early replicating genes are expressed, while genes replicating later are less likely to be transcribed. Thus replication timing is positive correlated with the activation state of a gene. Yet 30% of the latest replicating genes are still active, showing that this correlation is not absolute (see text). The grey line represents the average percentage of active genes (x-axis) relative to replication timing (y-axis). B) Coordination of replication timing and transcription along a chromosomal region. The graph shows a schematic representation of replication timing (y-axis) along a part of a *Drosophila* chromosome (x-axis). Genes that are active (green boxes) or inactive (red boxes) are distributed differentially in early and late replicating regions. Thus integration of the transcriptional activity over large regions appears to mediate early replication timing.

### 2.4.2. Dynamic changes in replication timing

The replication of the genome appears as such a critical cellular process that one might assume that the process is hardwired and inflexible. However if transcription and

replication timing are linked the temporal program should behave dynamic reflecting development specific gene expression. Notably the imperfect correlation between transcription and early replication does not per se require that replication behaves highly dynamic during development as the observed global trend could be driven by housekeeping genes expressed in most tissues. Indeed a comparison of the gene-rich chromosome 22 in human fibroblastoid and lymphoblastoid cells did not reveal widespread differences in replication timing (White et al. 2004) and the analysis of several genes during lymphocyte differentiation only revealed a small number of genes that changed in replication timing (Azuara *et al.* 2003). However development specific changes in DNA replication timing that coincide with gene activation are well described as in case of the beta-globin (Cimbora *et al.* 2000) or immunoglobulin gene loci (Zhou *et al.* 2002b) suggesting that switches in replication could be frequent.

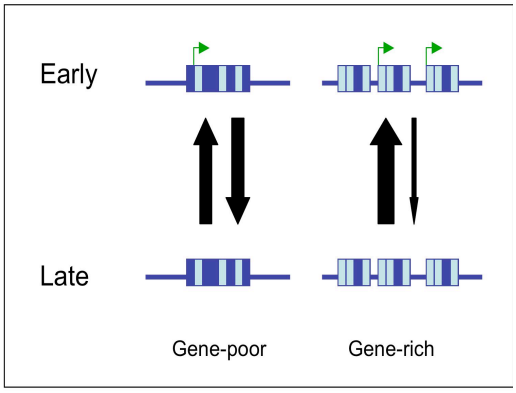



**Figure 8:** Transcriptional integration makes dynamic changes in replication timing more likely in gene-poor regions of the human genome. In a gene poor chromosomal region inactivation of a single gene reduces the transcriptional activity sufficiently to result in late replication. In gene-rich regions this would require the inactivation of several genes. Similarly activation of a single gene might be sufficient to result in transcriptional activity favoring early replication in gene-poor regions. (green arrow=active promoter, dark blue = intron, light blue = exon, black arrow indicates switch in replication timing, size of arrow indicates likelihood).

A recent study of 54 genes, which were selected based on their differential expression during differentiation of mouse embryonic stem cells to neuronal precursors showed that transcription-coupled dynamic changes in replication timing during differentiation were more frequent at genes residing in gene-poor chromosomal regions (Hiratani *et al.* 2004). A parallel study of 43 genes reported that about half of all genes show some dynamic changes in replication timing (Perry *et al.* 2004). As gene-poor isochors of the mouse genome are AT-rich and show a higher abundance of LINE elements it is possible that unique features of these regions contribute to their dynamic replication timing (Hiratani et al. 2004). However these results are also in agreement with a model that a threshold of transcriptional activity is required to mediate early replication. Consequently it is conceivable that in gene-poor regions the transcriptional silencing of a single gene is sufficient to result in late replication timing, whereas the lack of activity of a single transcript in a gene-rich region would be insufficient to mediate a switch to late

replication (Figure 8). To answer this question conclusively further comprehensive genome-wide measures at different developmental states would be required to determine the extent of dynamic replication timing and to define the chromosomal characteristics of regions that switch.

In many *Drosophila* tissues euchromatic regions or selected loci are multiplied during several rounds of endoreplication. A recent genome-wide analysis showed that regions that were reported to be late replicating in *Drosophila* Kc cells are largely overlapping with underreplicated regions in polytene chromosomes and appear to be enriched for tissue-specific genes (Belyakin *et al.* 2005). Underreplication does not seem to be a feature of gene density per se and thus endoreplication might share regulation with replication timing, which in both cases appears to involve the integration of transcriptional activity and chromatin structure over extended genomic regions.

### 2.4.3. Mechanisms and Molecules

Initiation of transcription and replication requires the formation of DNA bound multiprotein complexes in the context of chromatin. In the case of RNA polymerase recruitment the involvement of chromatin modifications and remodeling is well described (reviewed in (Sims *et al.* 2004)). Furthermore the process of transcription itself mediates defined histone tail modifications and even nucleosomal replacement (Ahmad *et al.* 2002; Wirbelauer et al. 2005), both of which might crosstalk to the initiation of replication. In *Xenopus* eggs induction of transcription from a plasmid leads to locally restricted elevated histone acetylation. In the same region a specific origin of replication is activated replacing the previously random initiation throughout the plasmid (Danis *et al.* 2004). In a first comprehensive study in higher eukaryotes early firing origins were mapped on *Drosophila* chromosome 2L (MacAlpine et al. 2004). The identified replication origins lacked a consensus sequence suggesting epigenetic determination of initiation sites. ORC binding was found to be excluded from AT poor regions, but enriched at RNA-polymerase bound but non-coding regions suggesting that replication initiates proximal to active promoters (MacAlpine et al. 2004). Direct interactions of ORC2 with transcriptional regulators, such as Rb and HP1, have been described and could be involved in regulating ORC activity in addition to a specific chromatin environment (Pak et al. 1997; Bosco et al. 2001). Similarly, it has been shown that chromatin factors can recruit ORC to human replication origins (Schepers *et al.* 2001; Thomae *et al.* 2008). This suggests that factors which normally regulate gene

expression could be used at the same time to attract proteins for replication initiation. Interestingly, some of the factors which interact with ORC, such as HP1, are typically involved in heterochromatin formation (Pak et al. 1997). ORC interaction with HP1 often occurs in a complex with the HP1/ORC-associated protein (HOAP) (Shareef *et al.* 2001). The role of this complex is largely unknown. It could reflect a chromatin-related function of ORC in heterochromatin, or an increased requirement of pre-initiation complexes in heterochromatin to accomplish complete replication of those densely packaged late replicating sequences.

A role of transcription factors for endoreplication and gene amplifications is well described. The Myb complex is required for chorion gene amplification (Beall *et al.* 2002) and the deletion of putative regulatory members leads to genome-wide amplifications (Bosco et al. 2001; Cayirlioglu *et al.* 2001; Beall et al. 2002; Lewis *et al.* 2004) suggesting that the complex is involved in specifying a region for amplification. However, the deletion of certain repressors seems to also lead to upregulation of preRC components resulting in increased amplification (Cayirlioglu *et al.* 2003). Recent knock-down studies and genome-wide profiles of proteins in the Myb-MuvB (MMB)/dREAM complex revealed that this complex locates in close proximity to many genes, some of which it can positively or negatively regulate (Georlette *et al.* 2007). It is possible that binding sites of this complex where it does not regulate genes reflect an additional function in the regulation of replication initiation. It is still poorly understood how gene amplification and transcription are connected yet it appears that transcription itself is not the main mediator. For example in *Sciara* embryos the specification of an initiation zone for amplification occurs at a developmental state prior to transcription at the locus, but subsequent of PolII binding (Lunyak *et al.* 2002).

Heterochromatic regions of the genome replicate late and thus late replication coincides with chromatin modifications that specify various forms of heterochromatin. Deletion of the Histone-deacetylase (HDAC) Rpd3 in *Drosophila* follicle cells leads to hyperacetylation and BrdU incorporation throughout the nucleus possibly reflecting global endoreplication (Aggarwal *et al.* 2004). Furthermore artificial tethering of Rpd3 to the chorion locus reduces gene amplification while recruitment of a Histone-Acetyltransferase (HAT) has the opposite effect indicating a role of repressive chromatin structure in restricting endoreplication (Aggarwal et al. 2004). Histones H3 and H4 are hyper-acetylated at origins during gene amplification (Hartl *et al.* 2007), and the location of those amplification origins and their enrichment for active chromatin modifications are

highly conserved between different *Drosophila* species (Calvi *et al.* 2007). Similarly, in a recent study in mammalian cells, artificial induction of histone acetylation at the β-globin origin resulted in advanced replication timing without an increase in transcription (Goren *et al.* 2008). These studies suggest that histone acetylation plays a crucial role not only in the activation of transcription, but also in early origin firing. This dual role of histone acetylation might reflect the requirement for an open chromatin state for any chromatin-templated event.

Endoreplication leading to *Drosophila* salvary gland polytene chromosomes is not uniform as intercalary heterochromatin is underreplicated. Underreplication depends on the 'Supressor of underreplication' (SuUR) protein (Belyaeva *et al.* 1998). SuUR is a putative SWI2/SNF2 chromatin remodeler (Makunin *et al.* 2002) and bound to underreplicated regions, which are largely overlapping with late replicating regions in *Drosophila* Kc cells (Schubeler et al. 2002; MacAlpine et al. 2004; Belyakin et al. 2005). Those underreplicated regions are also enriched in HP1 protein and repressive histone modifications. SuUR interacts with HP1, and SuUR targeting to chromatin could depend on this interaction (Koryakov *et al.* 2006; Pindyurin *et al.* 2008). Recently, SuUR has been shown to bind to late replicating regions in Kc cells, suggesting that this protein might be a general component of heterochromatin and possibly could be directly involved in regulating replication timing (Pindyurin *et al.* 2007).

The determinants of late replication in dividing cells are largely unknown. It has been shown that H3K9me3 in mouse cells is responsible for a delay in replication timing of pericentric heterochromatin. The replication timing of most of the genome however seems largely unaffected by a lack of H3K9me3 (Wu *et al.* 2006). Rpd3 deletion in yeast appears to advance the time of activity of some late firing origins (Vogelauer *et al.* 2002; Aparicio *et al.* 2004) and positioning of a silencer sequence from the mating type locus adjacent to an origin of replication can delay replication timing in cis in a SIR dependent manner (Zappulla et al. 2002), underscoring a tight link between chromatin structure and the time of firing of origins of replication. However, a genome wide correlation between active histone marks and origins of replication has not been observed in yeast (Nieduszynski et al. 2006). Such a correlation between replication timing and histone modifications exists in metazoa, where early replication correlates with gene activity. In *Drosophila*, active chromatin correlates with early replication (Schubeler et al. 2004), and a recent comparison of several histone modifications and replication timing for 1% of the human genome showed a correlation between H3K27me3 and late replication (Birney *et*

*al.* 2007). Together, these results suggest that chromatin might be a regulator of replication timing.

2.4.4. Nuclear organization and replication timing

The spatial organization of the genome in interphase nuclei is highly organized and replication and transcription occur in defined nuclear compartments. Current models suggest local foci of high transcriptional or replicative activity that consist of several active genes (Hassan *et al.* 1994) or replication forks (Hozak *et al.* 1993) with emerging evidence that genes present in one "transcriptional factory" are proximal in *cis* (Osborne *et al.* 2004). Moreover the spatial position and temporal activity of replication foci is mitotically inherited as shown by *in vivo* labeling of the replication machinery and newly replicated DNA (Sadoni *et al.* 2004). Previous elegant experiment using transfer of mammalian nuclei into *Xenopus* eggs had shown that replication timing is defined during early G1 around the same time when chromosomes are repositioned in the nucleus (Li *et al.* 2001). In *S.cerevisiae* the telomere-binding protein Ku is required for telomeric silencing, localization of telomeres to the nuclear periphery (Laroche *et al.* 1998) and their delayed replication timing (Cosgrove *et al.* 2002). However, excising a late firing origin from its chromosomal context results in loss of its peripheral localization, yet does not change its replication timing (Heun *et al.* 2001). This suggests that late replication is determined by certain chromatin modifications which also predispose a sequence in its chromosomal context to peripheral localization, or are established at the nuclear periphery. In metazoa the nuclear lamina seems to play an important role in chromatin regulation. It has been shown for several loci in mammalian cells that location close to the nuclear periphery correlates with transcriptional inactivation and late replication, while the same locus becomes early replicating and expressed in a cell type where it is located in the interior of the nucleus (Zhou et al. 2002b; Zhou *et al.* 2005). Recent genome wide determination of the regions in the genome which are associated with the nuclear lamina revealed that those regions are rather large, late replicating, gene-poor and often transcriptionally silent (Pickersgill *et al.* 2006; Guelen *et al.* 2008). Mutations in lamin A, which lead to premature aging, were shown to cause epigenetic changes in heterochromatin (Shumaker *et al.* 2006). Interestingly, cells expressing mutant lamins show defects in DNA replication, suggesting that lamins might play a role in genome duplication (Ellis *et al.* 1997; Spann *et al.* 1997; Moir *et al.* 2000). Together these studies suggest that replication timing and nuclear organization are linked however it remains to

be determined how this organization is mediated. While a direct role of the nuclear lamina in determining late replication cannot be excluded, it seems likely that the connection is mediated through the influence of nuclear structure on chromatin.

The nuclear periphery is not a uniform compartment, and not all parts of the nuclear envelope mediate silent chromatin structure (reviewed in (Akhtar *et al.* 2007)). On the contrary, nuclear pores, which mediate the transport of molecules between the nucleus and the cytoplasm, have been shown to enhance transcriptional activity of genes which are positioned in their proximity in budding yeast cells (Taddei *et al.* 2006). The hyperactivated *Drosophila* male X chromosome seems to localize to nuclear pores, as the dosage compensation complex has been shown to interact with nuclear pore proteins (Mendjan *et al.* 2006). It is possible that this activating compartment of the nucleus has a direct effect on the upregulation of transcription in male *Drosophila* cells. Furthermore, location at nuclear pores might correlate with early replication timing, similar to the late replication at the nuclear lamina described above.

### 2.4.5. Function of replication timing

Is differential replication timing just a downstream event of transcription without a function in gene regulation? While in some cases replication timing is changing upstream of transcription such as in the human beta-globin and immunoglobulin loci (Cimbora et al. 2000; Zhou et al. 2002b) in most cases the temporal order has not been defined. A focus on this “chicken or egg” problem appears justified but order does not necessarily predict hierarchy. In case of “chicken and egg” both are required for propagation and in analogy replication timing as well as nuclear organization might serve in stabilizing and maintaining transcriptional state through cell division even in those cases where they occur subsequent of transcriptional activation or repression.

Recent experiments using injection of plasmid DNA into cells at different times in S phase and analysis of the resulting chromatin structure suggested that differential chromatin gets deposited in early versus late S phase. Plasmids that were chromatinized in late S phase contained hypoacetylated nucleosomes and were less likely to be active than those chromatinized in early S (Zhang et al. 2002). Such replication timing dependent effects on chromatin structure support a model in which early replication can contribute to the stable propagation of an active state through mitosis by reducing the threshold required for subsequent gene activation (Figure 9).

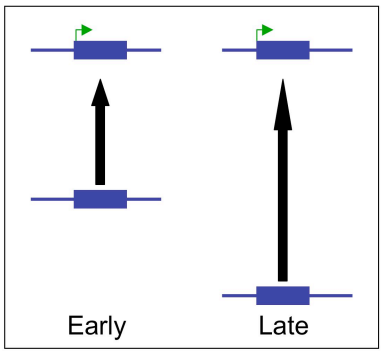



**Figure 9: Threshold model linking replication timing and chromatin state.** If late replication timing mediates a more compact chromatin (Zhang *et al.* 2002) late replicating genes (right) require higher activation "energy" (e.g. transcription factor binding) to overcome chromatin-mediated repression as do genes replicating early in S phase (left). Thus reestablishing of gene activity (green arrows) after mitosis could be more likely for early replicating genes and thereby stabilizing propagation of active states.

How chromatin marks are inherited during S phase is generally not well understood. Recent results indicate that RNA interference might play a critical role during the propagation of heterochromatin through S phase in *S. pombe* (Kloc et al. 2008), where early replication and gene activity are not so well correlated (Kim et al. 2003). Another mechanism of epigenetic memory is the incorporation of histone variants into actively transcribed regions, as demonstrated for H3.3, which confers transcriptional memory in the absence of transcription itself, in Xenopus embryos (Ng *et al.* 2008). The propagation of chromatin states through DNA replication requires that both daughter strands get a somewhat equal amount of old histones carrying a certain modification, and fresh histones. Little is known about the mechanisms involved in the redistribution of histones after replication, but recent results in human cells implicate the histone H3-H4 chaperone Asf1 as a key player in this process (Groth *et al.* 2007a). It is conceivable that the timing of replication influences the inheritance of chromatin states if certain modified histones themselves or their chaperones are available only during a certain part of S phase.

A separate explanation for early replication of active genes could be differential replication fidelity during the progression of S phase. Initial studies suggested that DNA duplication is less error-prone during early S phase (Wolfe *et al.* 1989). If true this would provide an evolutionary advantage for replicating coding regions early and could furthermore help to explain the clustering of genes in early replicating regions. Since coding regions have no intrinsic characteristic the process of transcription itself could be utilized as an indirect mark for genic regions, which mediates their early replication through chromatin and/or nuclear organization.

In higher eukaryotes the large and highly repetitive heterochromatic regions are duplicated in late S phase and a coinciding reduced replication fidelity or less sensitive repair machinery appear conceivable. Even though there is currently no molecular

evidence for differential activity of repair pathways during S phase progression this model would predict that silent mutations are more readily accumulating in late replicating, gene-poor regions in the genome. Indeed recent comparisons of the human and chimpanzee genomes support this hypothesis. Once it has been accounted for the prevalence of CpGs as mutational hotspots GC rich regions tend to have acquired less mutations than GC poor regions between both species arguing for differential mutation rates throughout the genome (Chimpanzee_Sequencing_and_Analysis_Consortium 2005; Hellmann *et al.* 2005). As GC rich regions of the human genome are gene-rich and have been shown to be both early replicating (Woodfine et al. 2004) and to reside in an open chromatin structure (Gilbert *et al.* 2004), both events could contribute to the observed differential divergence. This would provide an evolutionary benefit to link the processes of transcription and replication timing. Further comprehensive genomic analysis that takes sequence divergence between closely related species into account should allow to test if replication timing left a mark in the genome.

## 2.5. Scope of the thesis

At the start of my thesis in 2004, it had been postulated that active genes replicate early in S phase and inactive genes replicate late and thus that the regulation of transcription and replication are connected. However, this postulated link was not observed in a genome-wide study in *S.cerevisiae* (Raghuraman et al. 2001), while the first such study in *Drosophila melanogaster* revealed that early replicating genes are more likely to be active than those replicating later during S phase (Schubeler et al. 2002). It showed that about 70% of early replicating genes are active while 70% of late replicating genes are inactive. Since this interplay between replication and transcription was only observed in metazoans, it was hypothesized that a) it might reflect the more complex organization and regulation of the genomes of higher eukaryotes and that b) the replication timing of genes may behave dynamically depending on their activity. However, not all early replicating genes are active and therefore a change in expression does not have to coincide with a change in replication timing. My hypothesis was that replication is intimately linked to the control of gene expression and I predicted dynamic changes in replication timing between different epigenetic states. Since transcription is tightly linked to certain chromatin modifications (Schubeler et al. 2004), I also proposed that replication timing is linked to gene expression via chromatin.

To address those questions, I compared replication timing, transcription and histone acetylation between different cell lines on a genome-wide level using tiling arrays. In addition, I used RNA-interference to reduce the levels of chromatin proteins and tested the effect on replication timing.

The initiation of DNA replication defines the replication timing program, yet only very few origins of replication have been identified in meatazoans. Genome-wide studies of replication timing carried out at very high resolution allowed the identification of origins of replication throughout the yeast genome (Raghuraman et al. 2001), and in a recent effort early origins on a *Drosophila* chromosome were mapped (MacAlpine et al. 2004). I created a high resolution genome-wide map of replication timing, which allowed me to identify zones of replication initiation, and revealed the genomic organisation of replication timing.

# 3. Results and discussion

## 3.1. Chromatin structure marks cell-type and gender specific replication of the *Drosophila* genome

### 3.1.1. Summary

DNA replication begins at origins of replication which can fire at different time points during S phase. Despite recent efforts to map origins on a chromosome-wide scale (MacAlpine et al. 2004; Lucas *et al.* 2007), their location is still largely unknown in *Drosophila* or mammalian cells. The timing and efficiency of origin firing can vary between different cell types at certain tested loci (Zhou et al. 2002b; Hiratani et al. 2004; Azuara *et al.* 2006; Gregoire *et al.* 2006). It had been shown that the timing of DNA replication correlates with transcriptional activity on a genome-wide scale, where early replicating genes are more likely to be expressed than those replicating later during S phase (Schubeler et al. 2002). However, since this correlation is not absolute, it was not clear if replication timing can be dynamic for a large proportion of metazoan genomes or if only a few genes could change replication timing, while the general correlation is driven mostly by constitutively expressed genes. Furthermore, the mechanism of this correlation is still unknown. One possibility is that chromatin structure determines not only transcriptional activity, but also the activity of replication origins. Indeed, recent studies had identified histone acetylation as a crucial determinant for the firing of certain known origins (Vogelauer et al. 2002; Aggarwal et al. 2004; Goren et al. 2008).

To analyze this connection at the genomic level, I created high resolution replication timing profiles of *Drosophila* cells and compared it to chromatin and coding and non-coding transcription. This allowed the determination of the structure and size of replication domains and the approximation of the location and time of activity of initiation zones throughout S phase. I observed many distinct initiation zones in early S phase but less frequent initiation in mid S phase. Initiation peaks again in late S phase at multiple sites suggesting more promiscuous initiation at the end of S phase.

A comparison of different cell types revealed about 20% plasticity of replication timing on autosomes, most of which occur in large regions that coincide with local differences in transcription. The X chromosome shows a distinct behavior when hyperactivated in male cells: It replicates predominantly in early S phase and this switch to early replication occurs mostly at sites that are not hyperactivated transcriptionally but that show increased acetylation of lysine 16 of histone H4, suggesting a transcription independent yet chromosome-wide process related to chromatin. Interestingly H4K16ac is also enriched at initiation zones on autosomes, where it mostly coincides with transcription. Together these findings reveal a detailed blueprint on the organization of DNA

replication throughout the *Drosophila* genome and indicate a chromatin-dependent regulatory pathway of dynamic replication timing.

3.1.2. Submitted Manuscript

# Chromatin structure marks cell-type and gender specific replication of the *Drosophila* genome

Michaela Schwaiger[1], Michael B. Stadler[1],Hubertus Kohler[1], Edward J. Oakeley[1], and Dirk Schübeler[1*]

[1] Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, CH-4058 Basel, Switzerland

*Corresponding author
Email: dirk@fmi.ch

Running title: Plasticity and chromatin-dependence of replication timing in *Drosophila*

## Abstract

Duplication of eukaryotic genomes during S phase is coordinated in space and time. In order to identify zones of initiation and cell-type as well as gender-specific plasticity of DNA replication, we profiled replication timing, histone acetylation and transcription throughout the *Drosophila* genome. We observed two waves of replication initiation with many distinct zones firing in early and multiple, less defined peaks at the end of S phase, suggesting that initiation becomes more promiscuous at the end of S phase.

A comparison of different cell types revealed plasticity of replication timing at about 20% of autosomal sequences. Most of these occur in large regions but only half coincide with local differences in transcription. In contrast to confined autosomal differences, a global shift in replication timing occurs throughout the single male X chromosome. Unlike in females, the dosage compensated X chromosome replicates almost exclusively early. This difference occurs at sites which are not transcriptionally hyperactivated, but show increased acetylation of lysine 16 of histone H4. This suggests a transcription-independent, yet chromosome-wide process related to chromatin. Importantly, H4K16ac is also enriched at initiation zones as well as early replicating regions on autosomes.

Together, our data reveal organizational principles of DNA replication of the *Drosophila* genome and indicate that replication timing is dynamic and chromatin-dependent.

## Introduction

Duplication of the genome during S phase occurs in an ordered fashion as each sequence has to be replicated once per cell cycle. Most cell types have a temporally regulated program of genome duplication, where distinct chromosomal regions replicate at defined time points in S phase (Aladjem 2007). The coordinated completion of replication of all, even very late replicating sequences is crucial, especially since the existence of a checkpoint for genome duplication and completion of S phase in yeast has recently been put in question (Torres-Rosell *et al.* 2007).

The replication timing of a given sequence is defined by its distance to the closest origin, and by the time of firing of that origin. The measurement of replication timing can infer zones of initiation as regions that replicate earlier than their genomic neighborhood. This was utilized to identify origins throughout the *S. cerevisiae* genome (Raghuraman *et al.* 2001; Yabuki *et al.* 2002), where initiation occurs at defined sites that share consensus motifs (Nieduszynski *et al.* 2006). In higher eukaryotes only few sites of initiation have been mapped and many initiation events occur in broad initiation zones containing multiple origins (Dijkwel *et al.* 2000; Zhou *et al.* 2002a; Mesner *et al.* 2006). Recently, early firing origins on a *Drosophila* chromosome (MacAlpine *et al.* 2004), and several mammalian origins (Lucas *et al.* 2007) have been identified using microarray technology. This however revealed no consensus sequence predictive of metazoan origins of replication, even though the proteins that bind to origins are highly conserved between yeast and metazoans (Gilbert 2004). It has been speculated that epigenetic or structural features could determine the initiation of DNA replication, which in turn could explain dynamics in replication timing.

Early microscopic studies have shown that condensed heterochromatin replicates later during S phase than euchromatin (Gilbert 2002). The first genome-wide replication timing study in *S.cerevisiae* found no correlation between replication timing and active transcription (Raghuraman et al. 2001). On the other hand, early replication was correlated with sites of active transcription in metazoa, as revealed by similar genome-wide studies of replication timing and transcription in *Drosophila* (Schubeler *et al.* 2002; MacAlpine et al. 2004) and human cells (White *et al.* 2004; Woodfine *et al.* 2004; Jeon *et al.* 2005). Importantly, while highly significant, this correlation is not absolute. For example, in *Drosophila* 30% of late replicating genes are active and 30% of the early

replicating genes are inactive. This high frequency of exceptions makes it impossible to predict the extent of differential replication timing between cell-types.

Several single-gene examples of changes in replication timing or replication initiation with expression have been reported during the differentiation of mammalian cells (Hiratani *et al.* 2004; Perry *et al.* 2004; Gregoire *et al.* 2006) or for different cell types (Norio *et al.* 2005; Azuara *et al.* 2006). In addition, allelic differences in replication timing at imprinted genes and during X-inactivation in mammalian cells have been described (Gilbert 2002). However, a comparison of the gene rich human chromosome 22 between two cell types did not reveal widespread differences in replication timing (White et al. 2004). Thus the extent and determinants of dynamic changes in replication timing relative to development-specific expression is still unknown.

It is well established that active transcription always coincides with several characteristic histone modifications (Schubeler et al. 2004; Liu et al. 2005a; Rando 2007). Thus, a potential mechanistic link between changes in transcription and switches in replication timing may be chromatin structure (Schwaiger *et al.* 2006). Indeed, early replication has been linked to active chromatin modifications, such as histone acetyation (Vogelauer *et al.* 2002; Aparicio *et al.* 2004; Schubeler et al. 2004). On the other hand, late replicating regions are enriched in the repressive H3K27 methylation in human cells (Birney *et al.* 2007). For example the human β-globin locus replicates late in most cell types, but early in the cells where it is expressed. Upon deletion of the distal enhancer of the locus, open chromatin persists and the region still replicates early, even though it is not expressed (Cimbora *et al.* 2000; Schubeler *et al.* 2000). This particular case would imply that replication timing is more tightly linked to chromatin than to transcription. Furthermore, it suggests that an open chromatin conformation might poise an entire region for early replication, as has been suggested for other multigene loci (Chakalova *et al.* 2005).

In this study we have created high-resolution replication timing profiles of *Drosophila* cells and compared them to histone acetylation and transcription. This allowed us to determine the structure and size of replication domains and to approximate the location of initiation zones throughout S phase. Furthermore, the comparison of two cell types revealed a connection between replication timing differences that occur in large regions and localized differences in gene expression. In addition, we discovered a global shift in replication timing on the dosage compensated X chromosome in male *Drosophila* cells. This shift in replication timing to earlier replication is transcription-independent but reflects H4K16 acetylation, a chromatin modification characteristic of open, active

domains, which we find also enriched at initiation zones on autosomes. Together, these findings reveal a detailed picture of the organization of replication timing and imply chromatin structure as a key determinant of the replication program.

## Results

*High resolution replication timing analysis of the Drosophila genome*

A replication timing profile of sufficient temporal and spatial resolution is required to determine size and structure of replication domains. We applied an S phase fractionation assay to determine replication timing in *Drosophila* cells, where non-synchronized cells are pulse-labeled with BrdU and sorted into different S phase fractions (Gilbert *et al.* 1987). Replicating DNA in each fraction is isolated by immunoprecipitation with an antibody against BrdU (Hansen *et al.* 1993). To measure the temporal resolution of the assay, we sorted cells into three equal parts of S phase (early, mid and late) and quantified enrichments of replicating DNA in each fraction at a set of genes and a repetitive element (Fig. 1A-B). As previously demonstrated in mammalian cells (Cimbora et al. 2000; Azuara *et al.* 2003), we observe a distinct peak of enrichment for every sequence, indicative of a defined time of replication. Yet each sequence can also be detected at lower levels at other time-points in S phase. As a consequence an early replicating gene is about 10-fold more abundant in the earliest over the latest fraction. A mid-S phase replicating sequence such as the CG9743 gene (Fig. 1B) peaks in the central fraction but can also be detected at lower levels in early and late S phase. Its mid-S phase replication can thus also be inferred by its presence in the early and late fraction (Fig. 1C). Such broad distribution of replication timing around the peak results from combining the timing patterns of 60.000 non-synchronized cells in each sorted fraction. This integration of many cells creates a temporal resolution much higher than could be obtained measuring only a single cell with this assay. It furthermore allows us to infer the timing of all tested sequences by considering only the early and late fractions (Fig. 1B-C).

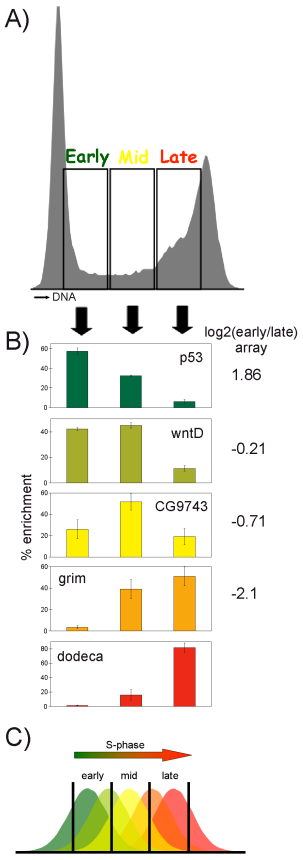



**Figure 1: High resolution replication timing analysis in *Drosophila* cells.** A) DNA content profile of Kc cells. The FACS sorting-gates for early, mid and late S phase are indicated. B) Enrichments of BrdU containing DNA in each sorted fraction as quantified by real-time PCR. Four genes and a pericentric repeat sequence are shown that replicate at different times. As expected, the heterochromatic repeat replicated later than all tested genes. Note that in all genes the relative replication timing can be correctly inferred from measuring only the early and late fraction by using PCR (bar graphs) or microarrays (log2 (early/late) array). This is also the case for mid-replicating sequences such as CG9743. Error bars represent the standard deviations between biological repeats. C) Graphical illustration of the broad distribution of replication timing patterns over the sorted regions allowing us to obtain continuous timing information by considering the early and late fractions.

To obtain high spatial resolution, we hybridized these two fractions enriched for early and late replicating DNA to arrays that cover the whole genome with 25-mer oligonucleotide probes separated by only 10bp. From the array measurements replication timing was calculated as the ratio of signal of the early versus the late replicating DNA as previously described (Schubeler et al. 2002). This global profile proved to be highly reproducible between biological replicates (Suppl. Fig. 1A and D) and independent of using either the Affymetrix or Nimblegen microarray platform for detection (Suppl. Fig. 1B). Figure 2a shows the timing values and profile for a representative region. As is evident from this section neighboring data points tend to have similar replication timing values. This spatial dependency is expected as replication timing does not change between proximal sequences. This similarity of neighboring probes can be quantified statistically as it results in very high autocorrelation, which extends up to 200 kb (Suppl. Fig. 1C). This finding confirms the previous observation that regions of similar replication timing tend to be large in *Drosophila* (MacAlpine et al. 2004). They are nonetheless much smaller than those in mammalian cells, which can extend over mega-basepairs (White et al. 2004; Woodfine *et al.* 2005).

*Structure and distribution of replication domains*

A zone where replication initiates replicates earlier than its up- and downstream neighboring regions and consequently replication timing can indicate where initiation occurs. This has already been shown for the yeast genome (Raghuraman et al. 2001) and at lower resolution for computationally predicted human origins (Woodfine et al. 2005; Audit *et al.* 2007). Indeed, zones of initiations appear as peaks in the replication timing profile, and the time of appearance of a peak during S phase reflects its time of initiation. Conversely, a valley in the replication timing profile represents a zone where replication forks converge and replication terminates. We used our 35bp resolution replication timing profile to define such initiation and termination zones genome-wide (Fig. 2A, Suppl. Fig. 2C, see methods for details).

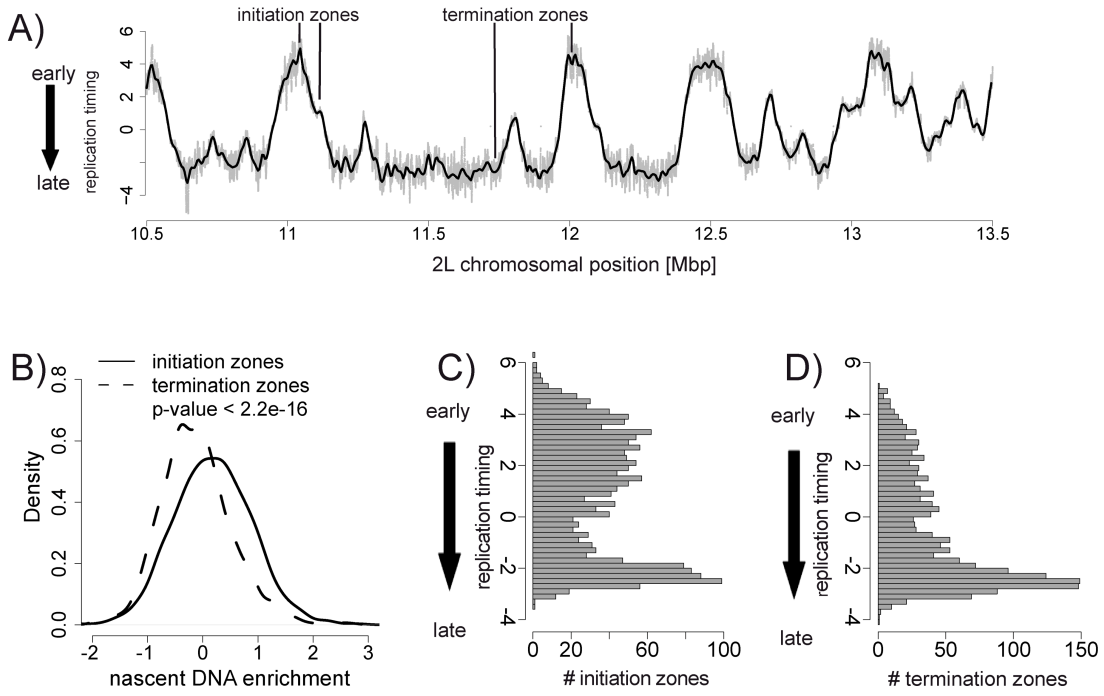



**Figure 2: Distribution of zones of replication initiation and replication fork convergence.** A) Shown are the replication timing values for a representative region of chromosome 2L in Kc cells. Initiation zones manifest as peaks, and regions of fork termination, as valleys. Several of these are indicated on this profile. Individual dots represent raw replication timing values and the black line represents the loess-smoothed replication timing profile (see methods for details). The y-axis denotes the replication timing ratio and the x-axis the chromosomal position in bp. B) Initiation zones are enriched for small nascent strand DNA. Shown is a density plot comparing enrichments of small nascent strands in timing-defined initiation versus termination zones. This analysis shows that initiation zones are enriched in small nascent DNA compared to termination zones ($p < 2.2e-16$). p-value was calculated using the Wilcoxon rank sum test. C) Histogram displaying the frequency of timing-defined initiation zones throughout S-phase, which reveals reduced

initiation events in mid S phase. D) Histogram displaying the frequency of timing-defined termination zones throughout S phase, revealing increased fork convergence events towards late S phase.

This analysis identified zones of initiation, which are expected to contain at least one active origin, at a resolution of 10 kb. To validate this approach we determined if peaks are enriched in small nascent strands that can only be detected proximal to active origins (Aladjem *et al.* 1998; Altman *et al.* 2001; Liu *et al.* 2003). We isolated 1-2kb long nascent DNA fragments from logarithmically growing cells by alkaline gel electrophoresis (Gray *et al.* 2007). Within this fraction we enrich for a control sequence of a previously described origin (Sasaki *et al.* 1999; MacAlpine et al. 2004) (Suppl. Fig. 2A). After hybridizing nascent DNA to tiling arrays we observed a significant enrichment of nascent DNA in initiation zones compared to termination zones (Fig. 2B). Moreover, we recovered 87% of all previously mapped replication origins on chromosome 2L (MacAlpine et al. 2004) (Suppl. Fig. 2B). Thus several lines of evidence confirm that our procedure correctly identifies regions, at a resolution of 10 kb, which contain sites of active initiation.

We next scored the frequency of such initiation zones throughout S phase. To this end, we quantified the abundance of initiation and termination zones at different time points (Fig. 2C-D). A high number of initiation zones are detected in early S phase, while their frequency declines towards mid-S phase. Interestingly, in late S phase, the number of initiation zones increased significantly (Fig. 2C). This suggests that, at the level of the genome, initiation of DNA replication occurs in a discontinuous manner, with many sites of initiation in early and late S phase, but fewer in between. The distribution of termination zones however is different (Fig. 2D). These are distributed in equally low amounts in early and mid-S phase but become very abundant in late S phase (Fig. 2D). This discrepancy in the temporal occurrence of initiation and termination suggests that early initiating forks progress over longer distances before they converge compared to late initiating forks.

To test this prediction we measured the length of replication domains as the distance between two convergence zones, which by definition contain an initiation zone (Suppl. Fig. 2C). In turn the resulting length of such a replication domain reflects the chromosomal distance replicated by forks originating from one initiation zone (Suppl. Fig. 2C-D). This analysis reveals that early-initiating replication domains are on average larger (80kb vs. 30 kb) than late-initiating domains (Suppl. Fig. 2D). Together with the

observation that the highest number of termination zones occurs in late S phase (Fig. 2D), this also means that in *Drosophila* many replication domains extend until late S phase (Fig. 2A).

Overall we find that late initiating domains are greater in number than early initiating domains yet each replicates a smaller region. Furthermore as shown in Figure 2a we identified large regions that replicate late, which consist of clustered late-replicating domains. In summary, replication timing across the *Drosophila* genome shows many zones of initiation in early S phase, and fewer in mid S phase. Replication from those initiation zones frequently continues until late S phase, where many replication forks converge. Interestingly, large late replicating regions contain a high number of late initiating zones that reside in close proximity.

*20% of autosomal sequences show cell-type specific differences in replication timing*

To define what fraction of the replication timing program differs between distinct cell types we compared replication timing between two different *Drosophila* cell lines. Kc cells are derived from embryos (Echalier 1997) and their transcriptome is similar to embryonic tissue (Greil *et al.* 2003), while Cl8 cells were isolated from wing imaginal discs of 3$^{rd}$ instar larvae (Peel *et al.* 1990). In addition we determined in both cell types transcription using 3'UTR as well as chromosomal tiling arrays (see methods). To determine regions of dynamic replication in an unbiased way we employed a three state Hidden Markov Model (HMM) to segment replication timing differences between both cell types (see methods). In order to focus on robust changes we excluded regional differences that are smaller than 20 kb or where the difference in timing extends over less than 25% of the total range (delta log2 < 1.74). These stringent criteria reveal 95 regions, corresponding to 12% of the genome on autosomes, that replicate earlier in Kc than in Cl8 cells (E:L) and 78 regions, corresponding to 9% of the genome on autosomes, that replicate earlier in Cl8 than in Kc cells (L:E) (Fig. 3A). These differentially replicating regions can be larger than 300kb, have an average size of 100kb (data not shown), and represent 21% of autosomal sequences. They often contain multiple replication domains suggesting that their initiation is regulated in a coordinated manner (Suppl. Fig. 2D).

*Dynamic replication timing on autosomes correlates with differential gene expression*

In both Kc and Cl8 cell types replication timing correlates with transcriptional activity (Suppl. Fig. 3A-B). In many cases, such as in the examples shown in Figure 3A, transcription in a differentially replicating region is higher in the cell type where this region replicates earlier. A genome-wide comparison of average transcription differences to replication timing (Fig. 3B) reveals that regions that replicate earlier in Cl8 cells (L:E) are also transcribed at a higher level there. Similarly, regions that replicate earlier in Kc cells (E:L) show increased transcription in Kc cells (Fig. 3B, Suppl. Fig. 4A). Notably, in many cases, these transcriptional changes only occur at a low percentage of genes in each differentially replicating region (Suppl. Fig. 4A).

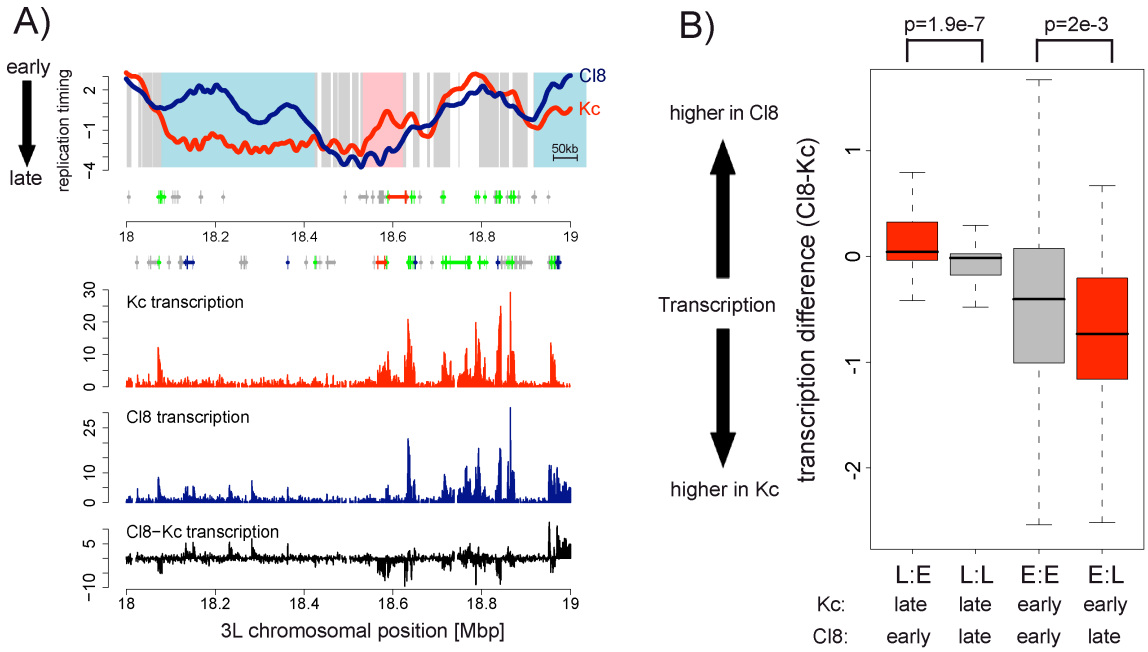



**Figure 3: Differences in replication timing correlate frequently with transcription differences.** A) Replication timing profiles of Kc (red) and Cl8 (blue) cells for a representative region on chromosome 3L. X-axis = 3L chromosomal position in mega-basepairs, y axis = log2 (early/late replication). Background coloring denotes regions that replicate earlier in Cl8 cells (L:E, blue), regions that replicate earlier in Kc cells (E:L, pink) and regions replicating similarly in both cell types (grey). Regions with small differences over short regions were not included in further analysis (white). Annotated genes are displayed below the profile (boxes=exons, lines=introns, small boxes=UTRs) and colored by their expression status (see methods for details, green= expressed in Kc and Cl8 cells, blue= expressed only in Cl8 cells, red=expressed only in Kc cells, grey= not expressed in Kc and Cl8 cells). Transcription levels of Kc (red) and Cl8 (blue) cells measured by tiling arrays are displayed on the same scale below, including transcription level differences (black). B) Distribution of transcription differences (Cl8-Kc transcription levels) for regions with differential replication timing on autosomes. The boxplots illustrate that on average differences in replication timing coincide with changes in transcription. L:E=regions replicating earlier in Cl8 cells, L:L=regions replicating late in both cell types, E:E=regions replicating early in both cell types, E:L=regions replicating earlier in Kc cells. p-values were calculated using the Wilcoxon rank sum test.

At the same time we do not detect such transcriptional changes in about half of differentially replicating regions and thus transcription is not the sole determining force (Suppl. Fig. 4A). Interestingly, genes that are transcriptionally inactive in both cell types, but which replicate earlier in Cl8 cells, are enriched for genes related to wing imaginal disc development (Table 1). Cl8 cells are derived from imaginal discs and thus early replication of genes involved in wing development, such as the wingless gene (wg, Suppl. Fig. 4B), might reflect an open chromatin state poised for subsequent activation. Taken together, this cell-type comparison of transcription and replication timing reveals evidence for both transcription dependent and transcription independent pathways that affect replication timing.

| Gene Ontology terms | p-value |
|---|---|
| Wnt receptor signaling pathway, calcium modulating pathway; ventral midline development; compartment specification; spiracle morphogenesis, open tracheal system | 0.0161 |
| sensory perception; dopamine metabolic process; catecholamine metabolic process; phenol metabolic process; indolalkylamine biosynthetic process; serotonin biosynthetic process; biogenic amine metabolic process; sensory perception of taste; indole derivative biosynthetic process | 0.0185 |
| leg disc development; leg morphogenesis; limb morphogenesis; imaginal disc-derived leg morphogenesis; leg segmentation; appendage segmentation; leg disc morphogenesis; imaginal disc-derived limb morphogenesis; imaginal disc-derived leg joint morphogenesis; leg joint morphogenesis; establishment of ommatidial polarity | 0.0265 |
| neurotransmitter metabolic process; amino acid derivative metabolic process | 0.0265 |
| gut development; hindgut morphogenesis; ectodermal gut morphogenesis; gut morphogenesis; digestive tract morphogenesis; fibroblast growth factor receptor signaling pathway | 0.0265 |
| wing disc pattern formation; imaginal disc pattern formation; neuroblast fate determination; neuroblast fate commitment; neuroblast differentiation; wing disc anterior/posterior pattern formation | 0.0281 |
| segmentation | 0.0348 |
| central nervous system development | 0.0392 |

Table 1: Gene ontology terms associated with 447 genes that replicate earlier in Cl8 cells compared to Kc cells but are not expressed in either cell type (based on Affymetrix expression arrays, see methods for details). GO terms and their p-values were calculated using GOstat ((Beissbarth et al. 2004), see methods for details).

*Gender-specific replication timing - absence of late replication on the male X chromosome*

A comparison of the timing profiles between both cell types revealed a remarkable difference for the X chromosomes. In female Kc cells, replication timing of the two X chromosomes was similar to autosomes, while replication of the single X in male Cl8 cells was dramatically advanced. Basically no late replication is detected on the X chromosome in male cells ($p < 2.2e-16$, Fig. 4A-B and Suppl. Fig. 5A).

Note that in *Drosophila* males, compensation of sex-specific differences in X-linked gene dosage is achieved by doubling the expression from the single X chromosome and involves the activity of a male specific dosage compensation complex (MSL complex) (Lucchesi *et al.* 2005; Mendjan *et al.* 2007; Straub *et al.* 2007). To test if the advanced replication timing is merely reflective of gender-specific transcriptional differences, we assessed replication timing differences in inactive genes or intergenic regions on the X chromosome. Surprisingly, we found that most of those regions show advanced replication timing in Cl8 cells (Suppl. Fig. 5B), even though they are not subject to transcriptional upregulation and/or bound by the MSL complex (Fig. 4A and (Alekseyenko *et al.* 2006)). This is in agreement with the fact that dosage compensation does not involve activation of additional male-specific genes (Gilfillan *et al.* 2006), but instead achieves upregulation of genes that are already active in females (Fig. 4A, (Straub *et al.* 2005)). Thus, the sites of transcription on the X chromosome are as similar between the two cell types as they are between autosomes. Furthermore, binding of the dosage compensation complex (DCC) is restricted to genes, which are already early replicating on the female X chromosomes (Fig. 4A, Suppl. Fig. 5C). Thus local DCC binding does not account for replication timing differences. We conclude that unlike many autosomal regions, switching of the X-chromosomal regions to early replication occurs outside of sites of active transcription and furthermore does not involve the recruitment of the MSL complex.

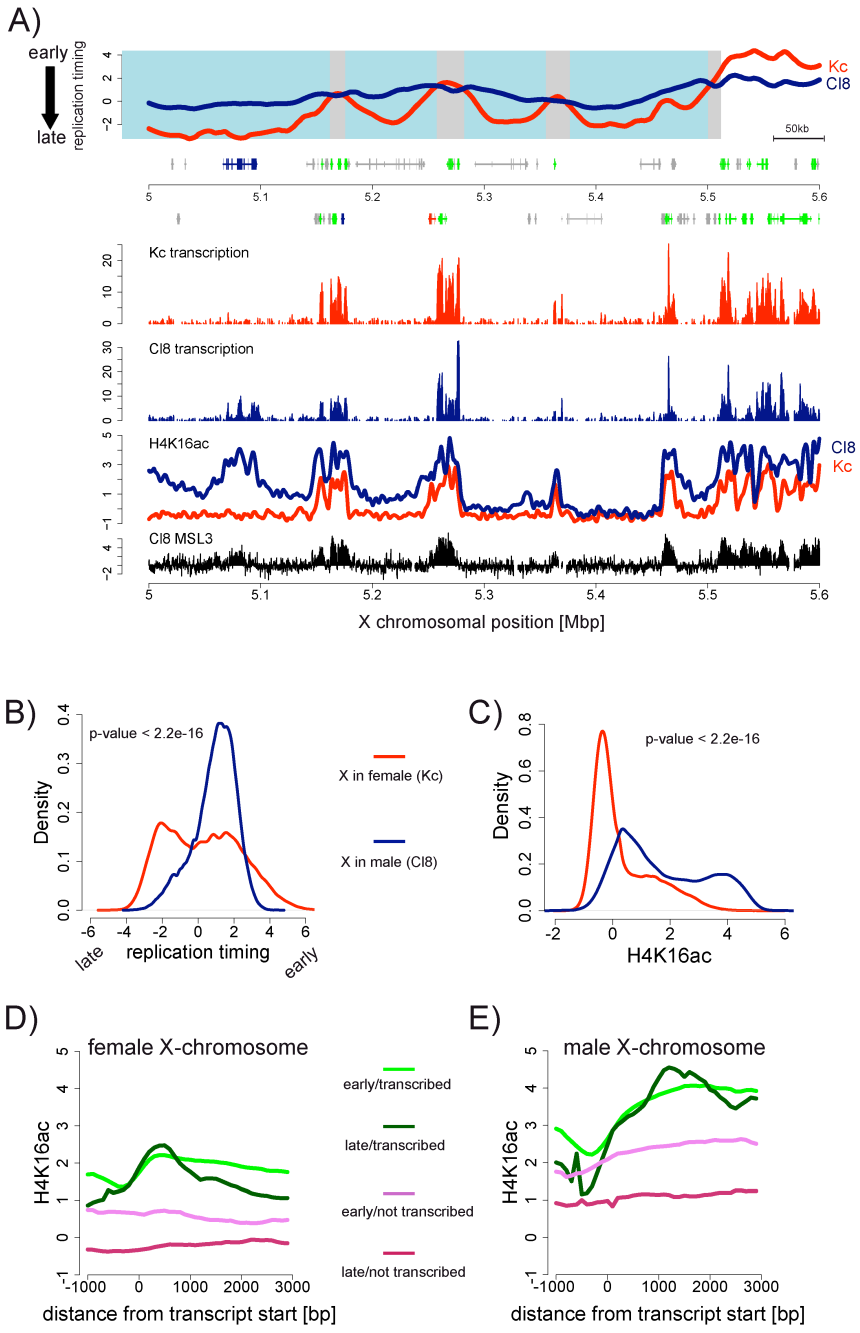



**Figure 4: Early replication timing of the male X chromosome.** A) Replication timing and transcription of representative regions on the X. Labeling is similar to Figure 3a. In addition H4K16ac profiles for Kc (red line) and Cl8 (blue line) cells are displayed and MSL3 binding data in Cl8 cells (black, (Alekseyenko et al. 2006)). B-C) Density plots of all raw replication timing (B) and H4K16ac values (C) of the X chromosome in Cl8 (male, blue lines) versus Kc (female, red lines) cells. p-values were calculated using the Wilcoxon rank sum test. D-E) Average signal for H4K16ac in Kc (D) and Cl8 (E) cells at genes relative to transcriptional activity (based on Affymetrix expression arrays, see methods for details) and replication timing (early= log2(early/late) > 0.2, late = log2(early/late) < -0.2). Genes were aligned at their start sites and signals were averaged (see methods for details). Active early replicating = green line, active late replicating = dark green line, inactive early replicating = violet line, inactive late replicating = red line.

*Acetylation of H4K16 marks early replicating regions on the male X that are not transcribed*

We next asked if early replication of non-transcribed sequences on the male X could be explained by changes in histone modifications that affect chromatin structure. The dosage compensated X chromosome is associated with highly elevated levels of acetylation of lysine 16 of histone H4 (H4K16ac) as measured by immunofluoresence, polytene staining, and chromatin-IP at selected genes (Turner et al. 1992; Smith et al. 2001a; Lucchesi et al. 2005; Gilfillan et al. 2006). Importantly, acetylation of this particular lysine residue has been shown to directly interfere with high-order chromatin compaction *in vitro* (Shogren-Knaak *et al.* 2006) and at the dosage compensated X chromosome in *Drosophila in vivo (Corona et al. 2002)* Moreover it is also present on autosomes in flies (Bell *et al.* 2007) and at accessible chromatin in human cells *in vivo* (Shogren-Knaak et al. 2006).

To address a possible link between H4K16ac and early replication, we generated genome-wide profiles by ChIP-CHIP in both Kc and Cl8 cells. We observed a global enrichment of this modification at the single male X chromosome (Fig. 4C), which is in agreement with previously published immunofluorescence data (Turner et al. 1992). Furthermore, this mark is not exclusive to the dosage compensated X as it is also present at active promoters on autosomal genes and on the female X (Fig.4D, Fig.5A, Suppl. Fig. 6C), consistent with a previous report (Bell et al. 2007).

At the dosage compensated male X chromosome, H4K16ac is mostly abundant throughout active genes (Fig. 4E) in agreement with single-gene studies (Smith et al. 2001a; Gilfillan et al. 2006; Bell et al. 2008). However, while most prevalent at active genes, the increased acetylation extends beyond the sites of dosage compensation and active transcription throughout the chromosome (Fig. 4A,C). Indeed, enrichment at inactive genes on the dosage-compensated X chromosome is as high as at active genes on autosomes. (Fig. 4A,E and Suppl. Fig. 6C). Surprisingly, H4K16ac at these transcriptionally inactive genes reflects precisely their timing of replication: Non-transcribed early replicating genes have significantly higher acetylation levels than the few inactive late replicating genes (p-value < 2.2e-16, Fig. 4E). This suggests a link between chromatin and replication timing independent of ongoing transcription. In summary, the dosage compensated X chromosome shows chromosome-wide advanced

replication, which relates most strongly with a local increase of H4K16ac outside of sites of dosage compensation.

*Acetylation of H4K16 marks early replicating regions and sites of initiation on autosomes*

An analysis of autosomal patterns of H4K16ac to transcription and replication reveals that H4K16ac is more abundant promoter-proximal at active genes in agreement with a function of this mark in transcriptional activation (Fig. 5A, Suppl. Fig. 6C). However, even at inactive autosomal genes, zones of early replication also bear higher levels of H4K16 acetylation (Fig. 5A, Suppl. Fig. 6C). For example, H4K16ac is substantially increased at a non-genic and non-transcribed site in the wingless locus in Cl8 cells and within an E:L differentially replicating region on chromosome 2L in Kc cells (Suppl. Fig. 4B and Suppl. Fig. 7A). Thus in analogy to the situation at the dosage-compensated X, early replicated regions on autosomes show increased levels of H4K16 acetylation even when not transcribed. Finally we note that the autosomal enrichment is particularly strong at zones of initiation. This is directly evident in a single chromosomal profile (Fig. 5B, Suppl. Fig. 8) and further supported by global analysis (p-value < 2.2e-16, Suppl. Fig. 7B). Importantly, this conclusion can also be drawn from the analysis of non-transcribed regions alone (Fig. 5C, p-value = 9.6e-12, Suppl. Fig 8). This argues that this histone modification, and not transcription per se, is the crucial mark for early initiation. It appears as if H4K16 acetylation on both autosomes as well as the dosage-compensated X chromosome is linked to early replication and sites of initiation.

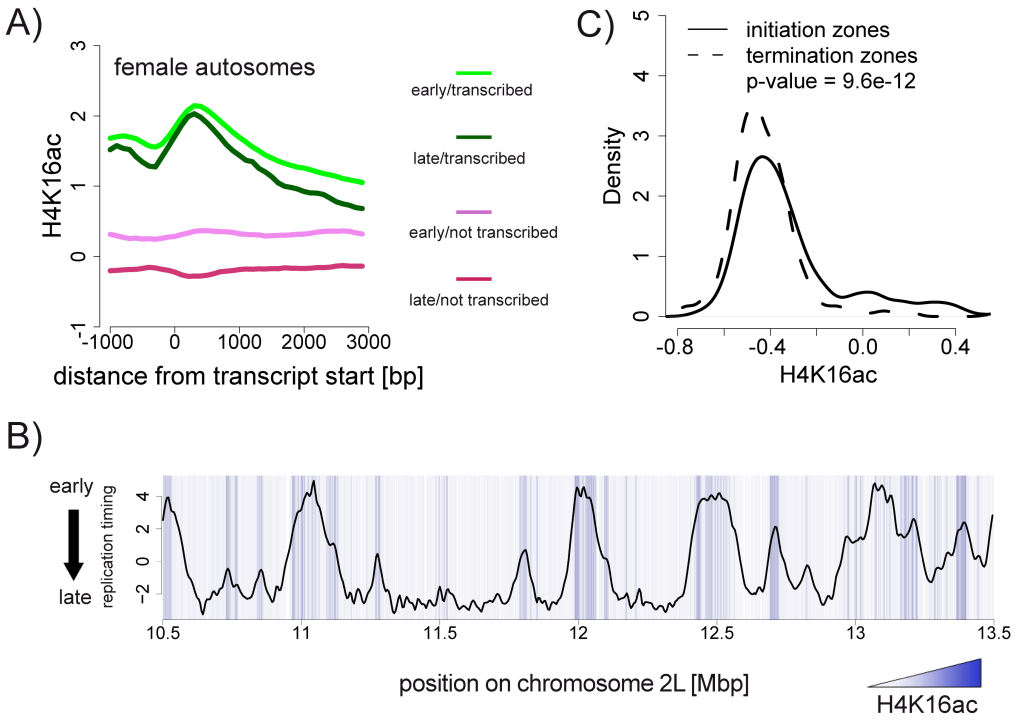



**Figure 5: Enrichment of H4K16ac at promoters and initiation zones on autosomes.** A) Average signal for H4K16ac at autosomal genes relative to transcription and replication timing. Note that among inactive genes early replicating ones have significantly higher signal. Genes were aligned at their start sites and signals were averaged (see methods for details). Active early replicating = green line, active late replicating = dark green line, inactive early replicating = violet line, inactive late replicating = red line. B) Shown is a Kc replication timing profile (line) with a heatmap illustrating H4K16 acetylation in the background. This reveals the presence of H4K16ac at sites of initiation regardless of their time of activity. White indicates absence of H4K16ac, darker colors indicate higher H4K16ac levels. C) Global abundance of H4K16ac in initiation and termination zones on autosomes. Shown is a density plot for H4K16ac in initiation or termination zones that are not transcribed (average transcription/zone < 0.5) revealing significant enrichment in initiation zones (p = 9.6e-12). p-value was calculated using the Wilcoxon rank sum test.

## Discussion

We provide a detailed analysis of replication timing of the *Drosophila* genome at 35bp resolution, which revealed insights into the chromosomal organization of replication, its tissue-dependency and its interplay with chromatin and transcription. We provide evidence that replication initiation occurs in a discontinuous manner throughout S phase. Tissue-specific replication timing manifests itself at the level of ~100kb large domains and in only half of all cases can be accounted for by differential transcription. Furthermore, we show that replication can be controlled chromosome-wide, because the X chromosome lacks late replication specifically in males. Finally we directly implicate

H4K16 acetylation as a transcription-independent mediator of early replication timing and as a chromatin mark highly abundant at sites of initiation.

*Genomic organization of replication timing and initiation*

The high resolution of our replication timing profiles allowed us to identify zones of replication initiation throughout S phase, which we confirmed in combination with measuring small nascent strand abundance. This revealed that sites of early initiation are rather distinct, which manifests in the timing profile as many large domains replicated from single initiation zones. Replication domains initiating in mid S phase form also highly symmetric peaks but tend to cover smaller regions. Late initiating domains are even smaller with an average size of 30 kb but often reside in close proximity to other late initiating domains.

Interestingly, the frequency of initiation appears discontinuous with high rates in early, a reduced frequency in mid-S and again increased appearance of initiation sites in late-S phase. The high frequency and proximity of late-firing initiation zones suggest that late regions are replicated by many proximal late firing origins of replication. This finding is particularly interesting in light of a recent report that suggested the absence of a checkpoint to control for the completion of DNA replication before mitosis (Torres-Rosell et al. 2007). This would in turn require a mechanism that mediates rapid replication of unreplicated regions in late S phase, which could be achieved by such promiscuous activation of many proximal origins. Furthermore, replicative stress that reduces replication fork progression leads to a decrease in inter-origin distance through activation of normally dormant origins (Anglana *et al.* 2003; Woodward *et al.* 2006). It is conceivable that a similar situation is encountered in late replicating regions.

*20% of the genome shows dynamic replication timing between distinct cell types*

Since the previously reported correlation between replication timing and transcription is not absolute, the percentage of the genome that replicates in a tissue-specific fashion remained to be tested quantitatively. For example, the general correlation could be driven by housekeeping genes that are active in most cells, resulting in a uniform replication timing program (Schwaiger et al. 2006). Here we show that dynamic replication timing differs significantly between two *Drosophila* cell types, affecting at least

20% of autosomal DNA. We also show by two different methodologies that this plasticity of DNA replication coincides with transcription differences in only half of all cases.

Early replication was shown previously to correlate with transcription levels over 180kb leading to the suggestion that replication timing integrates transcription over large regions (MacAlpine et al. 2004). Consistent with this model, we find that dynamic replication timing often occurs in large (~100kb) regions encompassing many genes. Some, but not all, genes in differentially replicating regions are differentially expressed between the two cell types. This implies that localized differences in gene expression of a fraction of genes in a large region relate to replication timing differences. In light of our finding that regions of initiation have high levels of H4K16 acetylation, replication differences could be accounted for by transcription-coupled and transcription-independent local changes in chromatin structure.

A change of replication timing in most cases will affect large regions covering many genes and genes with related function often cluster together in the *Drosophila* genome (Boutanaev *et al.* 2002). Furthermore, genes that are co-regulated and expressed at similar levels are clustered into, on average, 100kb long regions (Boutanaev et al. 2002; Spellman *et al.* 2002). In mammalian genomes this clustering appears functionally related to chromatin structure (Gierman *et al.* 2007), suggesting that widespread open chromatin at developmentally regulated mulitgene loci could lead to early replication or vice versa (Gilbert 2002; Chakalova et al. 2005; Schwaiger et al. 2006; Aladjem 2007). This in turn might increase the potential of gene expression over large regions as in the case of genes important for wing disc development in Cl8 cells (Sup.Fig.4b), where early replication could render the locus poised for activation.

*H4K16 acetylation links chromatin with early replication and initiation*

The relation between replication timing and chromatin structure has been controversial. Transcription itself involves an opening of chromatin structure (Chakalova et al. 2005) and thus early replication could in many situations be downstream of transcriptional activation (Gilbert 2002; Danis *et al.* 2004; Schwaiger et al. 2006; Aladjem 2007). However, previous work using injected plasmids suggested a role for early replication in mediating increased levels of histone acetylation (Zhang *et al.* 2002). This led to a model in which replication timing mediates an open chromatin structure required for transcription. This suggestion is compatible with our genome-wide analysis, where we

observe a preferential location of H4K16ac not only to active genes, but also to early replicating regions that are not transcribed (Fig. 4-5). In this study we focused on acetylation of H4K16, because this residue has been functionally linked to higher-order chromatin compaction, and chromatin opening on the dosage compensated X in *Drosophila* (Corona et al. 2002; Shogren-Knaak et al. 2006; Suganuma et al. 2008). It is possible that early replication and elevated H4K16ac at inactive genes will result in a more open chromatin confirmation compared to late replicating inactive genes. This might render them more responsive to downstream activating cues and thus replication timing could modulate the sensitivity to activators. This process could also function in maintenance of an active state through cell division (Schwaiger et al. 2006; Groth *et al.* 2007b). Importantly however, this mechanism does not override the parallel process of transcription-coupled acetylation, as late replicating genes that are actively transcribed are still hyperacetylated (Fig. 4D,E, Fig. 5A).

Interestingly, we also observe a strong abundance of H4K16ac at sites of initiation regardless of their time of activity (Fig. 5B). Several studies have suggested a positive function for histone acteylation for origin activity (Lin *et al.* 2003; Aggarwal *et al.* 2004; Danis et al. 2004; Norio 2006; Calvi *et al.* 2007; Hartl *et al.* 2007). Other reports however did not support this model (Prioleau *et al.* 2003; Dazy *et al.* 2006; Gregoire et al. 2006; Norio 2006). Furthermore, no genome-wide correlation between active chromatin marks and early origin firing was observed in *S. cerevisiae* (Nieduszynski et al. 2006), where specific sequences function as origins of replication. Here we identified a preferential localization of H4K16ac to initiation zones throughout the *Drosophila* genome compatible with a function of acetylation. It has been proposed that origins of replication lie frequently between promoters of active genes (MacAlpine *et al.* 2005; Huvet *et al.* 2007), which would make transcription and replication fork progression co-oriented (Huvet et al. 2007). Furthermore transcription and replication are thought to be coordinated in the nucleus (Chakalova et al. 2005; Aladjem 2007) to be spatially and temporally separated. It thus seems plausible that the enrichment of H4K16ac in initiation zones reflects location between highly acetylated, active promoters. Thus proximity to active promoters would result in an open chromatin confirmation through increased H4K16ac, which in turn enhances origin firing.

Most compelling is the fact that we score a similar enrichment for H4K16ac at initiation zones that are not proximal to active genes. This argues against a simple transcription-coupled process. Open chromatin structure, mediated by H4K16ac, could make DNA

more accessible for efficient initiation of DNA replication and thus provide a sequence independent component that could contribute to origin localization and activity. While these are testable models they do require a fine-mapping of actual origins at a resolution higher than our detection of zones of initiation at the level of 10 kb.

*Absence of late replication of the male X chromosome reflects transcription-independent changes in chromatin*

Our analysis reveals the almost complete loss of late replication on the single X chromosome in male *Drosophila* cells. About 90% of female late replicating regions replicate now early. Such chromosome-wide advance in replication timing has not been observed previously. In mammals, transcriptional inactivation of one of the female X chromosomes correlates with its late replication, which has been observed microscopically. This reflects the efficient silencing of this chromosome and increased chromatin compaction (Lucchesi et al. 2005). In contrast, dosage compensation in flies involves the two-fold upregulation of genes already active in females (Straub et al. 2005) and an open chromatin state mediated by H4K16ac (Corona et al. 2002). Interestingly, we show that the advanced replication of the dosage-compensated X occurs mostly outside of transcriptionally activated regions and thus is unlikely to be accounted for by transcriptional changes. Importantly the local increases in H4K16ac, which we report throughout the male X chromosome, can be directly related to the loss of late replication. This suggests a transcription-independent, chromatin-dependent process, which leads to early replication chromosome-wide. While this likely reflects a different chromatin compaction it is tempting to speculate that it also reflects a particular nuclear organization as the dosage compensated X chromosome has been shown to associate directly with nuclear pores (Mendjan *et al.* 2006; Akhtar *et al.* 2007).

Concluding remarks

Together our findings provide new principles of the replication timing program of the *Drosophila* genome and its dynamics relative to histone acetylation and transcription. Our data further support a model where open chromatin structure is a general feature of early replication timing and sites of initiation, which frequently coincides with, but does not require transcription.

## Materials and Methods

*Tissue culture and BrdU labeling*

*Drosophila* Kc cells were kept in HyQ-SFX (Hyclone). Clone8 (Cl.8+) cells were kept in Shields and Sang medium (Sigma), supplemented with 12.5IU/100ml insulin, 2% heat inactivated FCS and 2.5% fly extract (Peel et al. 1990). For labeling of newly synthesized DNA, we added BrdU (Sigma) to a logarithmically growing culture at a final concentration of 50 μM. After 60 min of incorporation time, we added Triton X-100 to a final concentration of 0.1% and DAPI to a final concentration of 3ug/ml for 5 minutes at room temperature. Cells were immediately used for fractionation.

*Cell cycle fractionation and isolation of BrdU labeled DNA*

We sorted cells into S-phase fractions on the basis of DNA content using Fluorescent Activated Cell Sorting (FACS). We collected 60,000 cells from each fraction directly into lysis buffer. DNA was purified, sonicated, denatured and immunoprecipitated with an antibody specific for BrdU (Becton Dickinson) as described (Schubeler et al. 2002), but with two consecutive rounds of immunoprecipitation to increase specificity. Enrichments of five control genes in early and late replicating DNA fractions were analyzed by real-time PCR (using a mitochondrial DNA sequence as control, primer sequences available upon request).

*Transcription analysis*

Total RNA was isolated from cells using Trizol (Invitrogen) and subsequently purified using an RNeasy kit (QIAGEN). For hybridization to Affymetrix tiling arrays, we made double-stranded cDNA by performing two rounds of cDNA synthesis using random primers and addition of 2 mM dUTPs using the GeneChip® WT Double-Stranded cDNA Synthesis Kit (Affymetrix). This cDNA was fragmented and end-labeled using the GeneChip® WT Double-Stranded DNA Terminal labeling Kit (Affymetrix) and hybridized to GeneChip® *Drosophila* Tiling 1.0R arrays (Affymetrix) according to manufacturer's instructions. For hybridization to expression arrays, cDNA synthesis and hybridizations were carried out according to standard Affymetrix procedures.

*Nascent strand analysis*

Nascent strand DNA in a size range of 1000 to 2000 bp was isolated from logarithmically growing Kc cells by alkaline gel electrophoresis as described (Gray et al. 2007). Genomic DNA from Kc cells in G2 phase (isolated by FACS) was used as a control that does not contain nascent strands. The enrichments of nascent DNA compared to genomic DNA for sequences within and distal to the DNA-Polα origin were analyzed using real-time PCR (Sup.Fig.2a).

*Chromatin Immunoprecipitation (ChIP)*

ChIP for H4K16 acetylation was carried out as described (Bell et al. 2007).

*Target sequence amplification and microarray processing*

To obtain sufficient target DNA for microarray hybridization, we amplified the denatured and immunoprecipitated DNA as described (Schubeler et al. 2002). Nascent DNA was amplified similarly, but with only one round of 3' degenerate primer annealing for subsequent linear amplification. The enrichments of several control genes were verified to ensure correct amplification before labeling and hybridization to arrays. For use with Affymetrix tiling arrays, dUTP was incorporated into the amplification reaction at 2 mM, the amplified DNA was fragmented and end-labeled using the GeneChip® WT Double-Stranded DNA Terminal labeling Kit (Affymetrix) and hybridized to GeneChip® *Drosophila* Tiling 1.0R arrays (Affymetrix) according to manufacturer's instructions. For Nimblegen arrays (Steve Henikoff custom design, (Mito *et al.* 2005), 2005-04-20_Dros_60mer_ChIP) labeling, hybridization and data extraction was performed by NimbleGen systems, including dye swaps. All microarray experiments were carried out in at least two biological replicates.
Affymetrix tiling arrays were analyzed using MAT (Model-based Analysis of Tiling-array) software (Johnson *et al.* 2006). The bandwith was set to 1000bp for replication timing and nascent DNA analysis, and 300bp for H4K16ac and transcription data. MAT scores were extracted from the BAR files generated using the Python script 'Bar2Wig.py' kindly provided by Wei Li (Harvard University). Data from the Wiggle files were reformatted using Perl for subsequent analysis in R and the ratio of early/late replicating DNA,

bound/input DNA and nascent/genomic DNA was calculated. Signal values for Affymetrix expression arrays were estimated using the GC-RMA module from Genedata's Refiner 4.5 package (Genedata, Basel, Switzerland). Expression data analysis was performed in Genedata's Analyst 4.5 package. Only those genes which had an Affymetrix detection p-value < 0.05 and signal values > 4 were assumed to be expressed. All analytical procedures were based on Flybase release 4.3 of the *Drosophila* genome. Raw-data files can be accessed at: http://www.fmi.ch/groups/schubeler.d/Gen_Dev_08/Schwaiger_supplemental.htm

*Replication timing profile analysis*

All analytical procedures were done using R (R_Development_Core_Team 2006) and all custom scripts are available upon request. To define zones of initiation and termination, five replication timing profiles generated in Kc cells on Affymetrix tiling arrays were first smoothed by local polynomial regression fitting (using R's loess function). Then we sampled a data point from the smoothed timing profile every 500bp, and defined peaks in the profile as points of initiation. To define replication domains, we extended the domains starting at initiation points in parallel towards both sides according to the following algorithm: Starting with the highest (earliest) replication timing value and advancing to lower (later) values, create a new initiation point at each chromosomal location with this value if it is at least 20kb away from earlier initiation points, or grow earlier replication domains that are adjacent to this location. Termination points were defined as the endpoints on either side of the replication domains. Finally, we extended the points of initiation/termination to either side until the replication timing value differed by more than 0.1. This resulted in zones of initiation/ termination which were on average about 10kb in size. Crossvalidation with several different smoothing parameters and different sizes of zones yielded similar results.

*Segmentation of replication timing profiles by Hidden Markov Models (HMMs)*

To define replication timing regions in Kc cells (Suppl. Fig. 2E), we segmented replication timing data using HMMs, as described (Birney et al. 2007). The basic premise of HMMs is that observed data are generated stochastically from a pre-determined

number of hidden background probability distributions, or states. We used three states, to distinguish early, mid and late replication. The parameters of the HMMs (emission probabilities, here modeled as normal distributions, and the transition probabilities between states), are estimated via unsupervised learning (Baum-Welch algorithm) from the replication timing profile. In our case, model parameters describe the range of replication timing values that are typical for early, mid and late replication states, and the patterns of changing state along chromosomal positions. In order to rule out the possibility that trained models correspond to suboptimal local minima in parameter space, the training procedure was repeated several times using varying initial parameters, all resulting in highly similar trained models. "Late", “mid” and "early" replication states were assigned to genomic positions according to the most probably path through the trained model states given the observed data (Viterbi algorithm). Consecutive genomic positions with identical replication states were merged, and to eliminate likely outlier values in the replication timing data, we further merged regions of identical states that were separated by less than 5kb (Sup.Fig.2e). The segmentation algorithm was implemented in Python using the GHMM library (Schliep *et al.* 2004).

To identify which parts of the genome show differences in replication timing in an unbiased fashion, we used the same three state HMM approach on a profile of replication timing differences (Kc-Cl8 replication timing). Thereby we can define regions that replicate earlier in Kc cells (E:L), those that do not show differences in replication timing, and regions that replicate earlier in Cl8 cells (L:E), in analogous way to how we defined early, mid and late replicating regions in one cell type. Many regions showed timing differences of various degrees, and about half of the genome showed no replication timing differences. For further analysis of the most prominent differences in replication timing, regions larger than 20kb with an average timing difference higher than the median of timing difference of regions which the HMM predicted to have no timing differences plus (E:L) or minus (L:E) $1/4^{th}$ of the range of all average timing differences were selected. However, all observed correlations of replication timing with transcription are also observed using different cutoffs and even with all differentially replicating domains as predicted by the Hidden Markov Model.

*Analysis of the distribution of H4K16 acetylation across genes*

To plot the average enrichment of H4K16ac across genes, we separated genes into active and inactive based on Affymetrix expression arrays (see above), and into early and late replicating based on average replication timing value per gene (early=>0.2, late=<-0.2). We aligned genes by calculating 100bp sliding windows from 1000bp before until 3000bp after the transcription start site of those genes. Genes which overlapped with another transcript on the other strand were excluded, and only genes between 3000 and 5000bp length were used, to make the enrichments comparable.

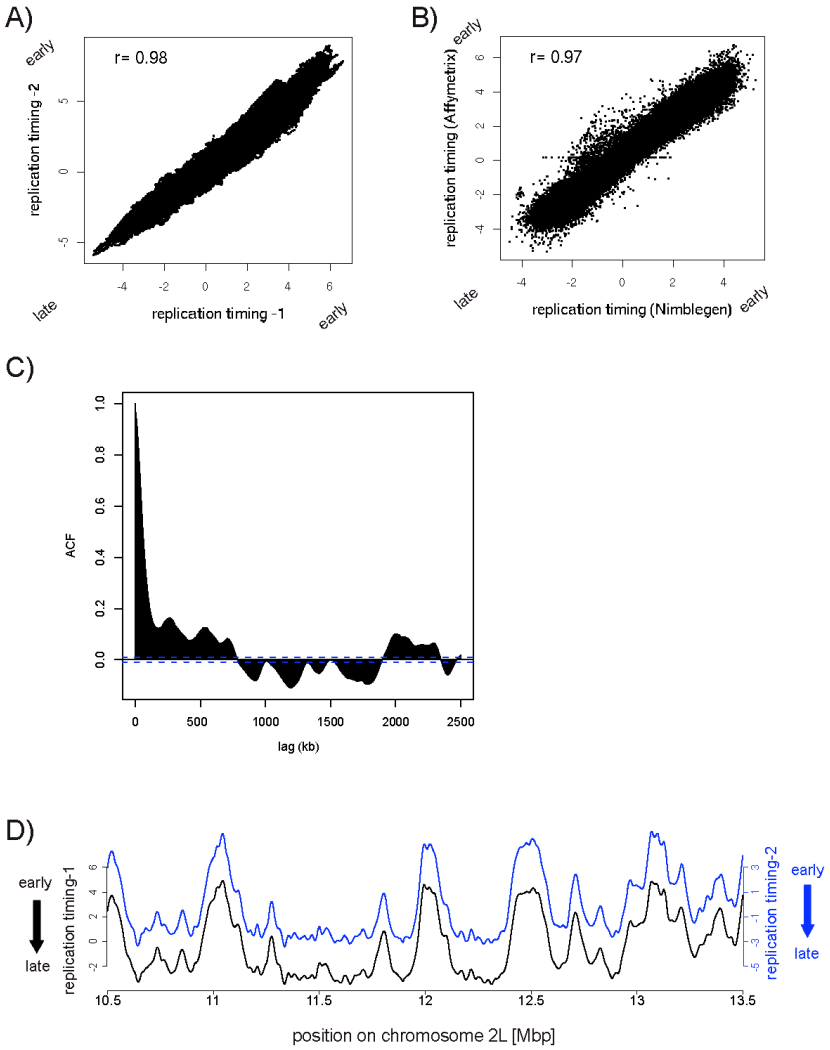
### *Gene Ontology (GO) term analysis*

We used all autosomal genes that are not expressed based on Affymetrix expression arrays (see above) and lie within L:E DRRs (regions which replicate earlier in Cl8 cells based on HMM segmentation of Kc-Cl8 replication timing profiles). We then searched for significantly enriched GO terms of the category “biological process” within each of those gene-lists using GOstat (Beissbarth *et al.* 2004). GO terms were clustered if indicating gene lists are inclusions or differ by less than one gene, and only the best 30 over- or underrepresented GO terms with a p-value below 0.05 were considered.

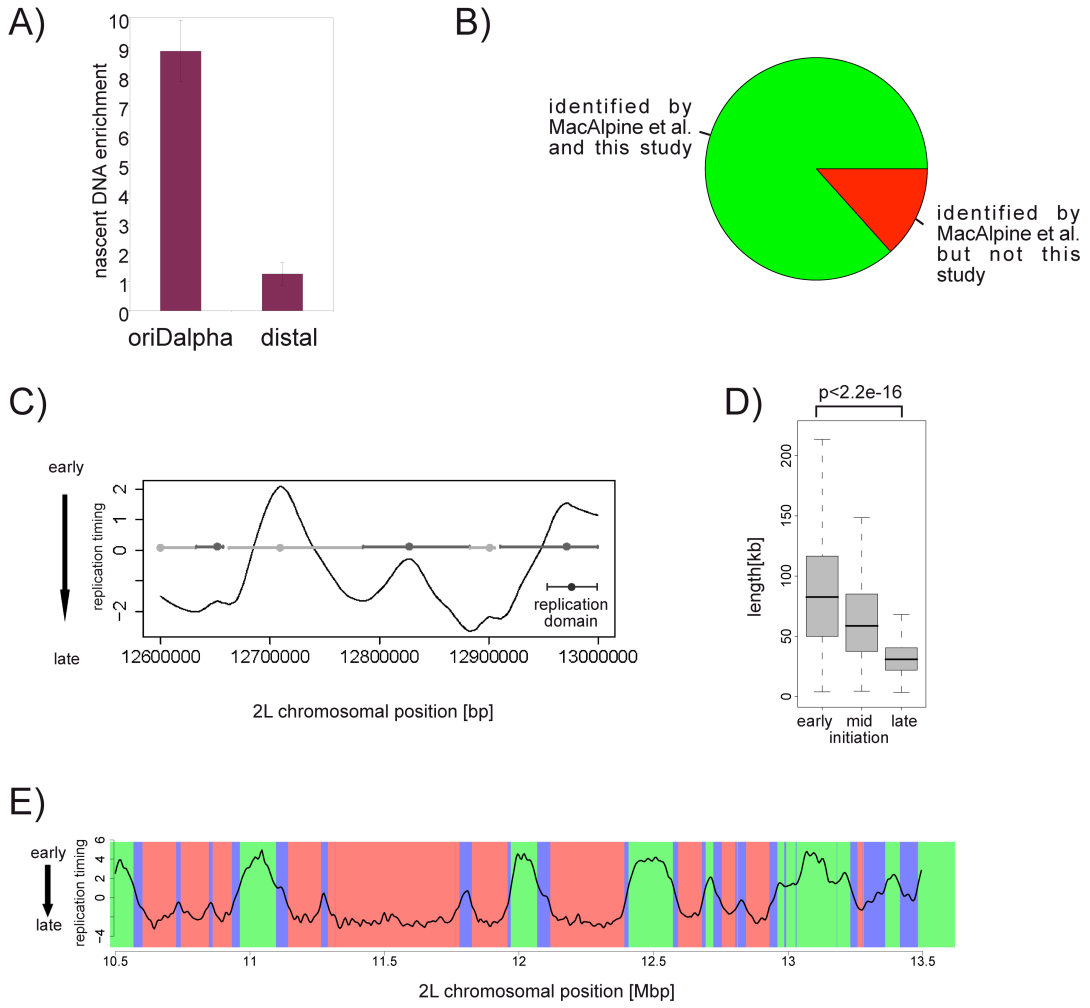
## Acknowledgements

We thank Antoine Peters, Susan Gasser, Kenji Shimada and members of the laboratory for helpful comments on the manuscript, and Steven Gray and members of the laboratory of David Gilbert for advice on nascent DNA and replication timing analysis. We are grateful to Deborah Cottam for providing the Cl8 cell line, and to Wei Li for providing scripts for MAT analysis. M.S. was supported by a predoctoral fellowship of the Boehringer Ingelheim Fonds. Research in the laboratory of D.S is supported by the Novartis Research Foundation, by the European Union (NoE “The Epigenome” (LSHG-CT-2004-503433), and LSHG-CT-2006-037415) and the EMBO Young Investigator program.

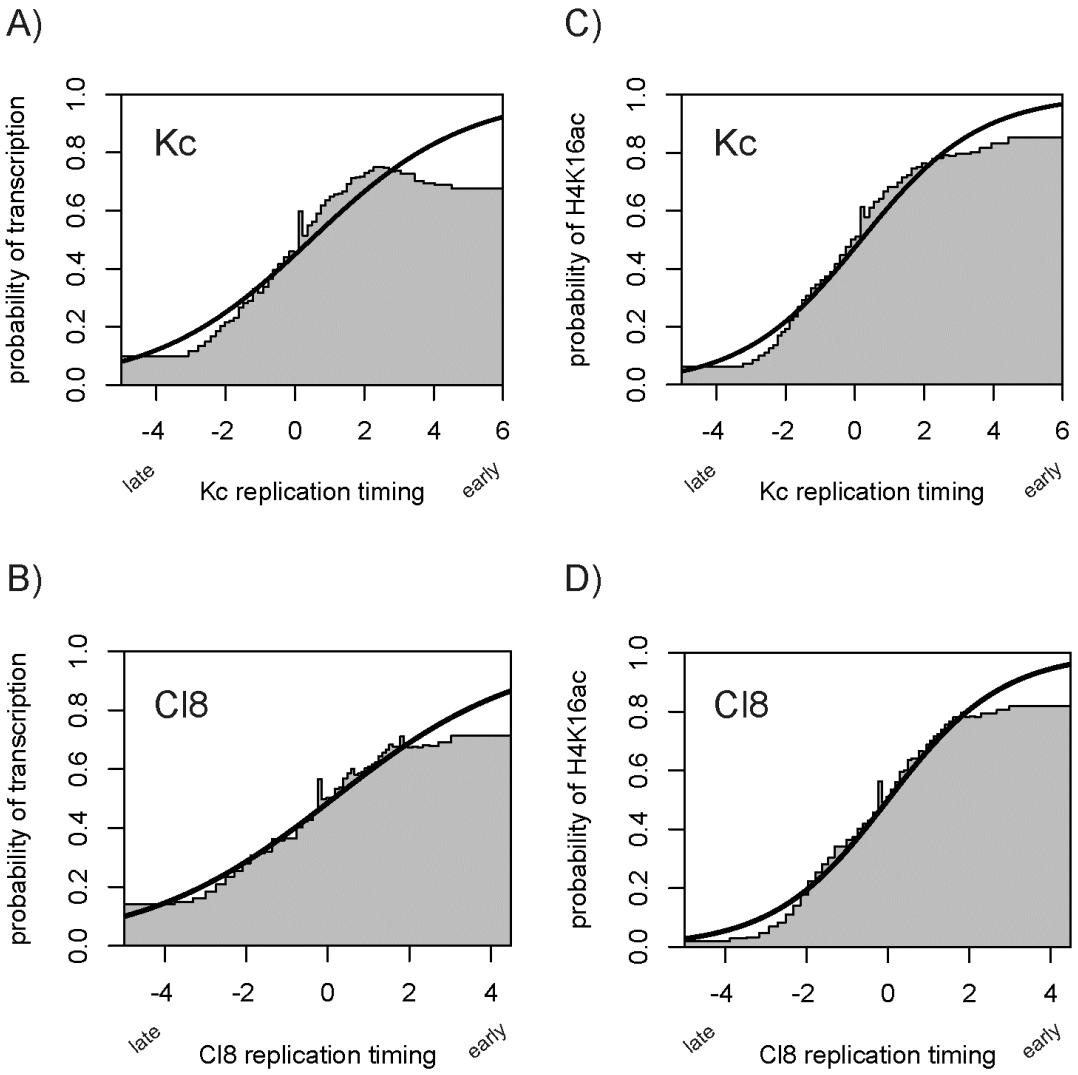
## 3.1.3. Supplementary Data



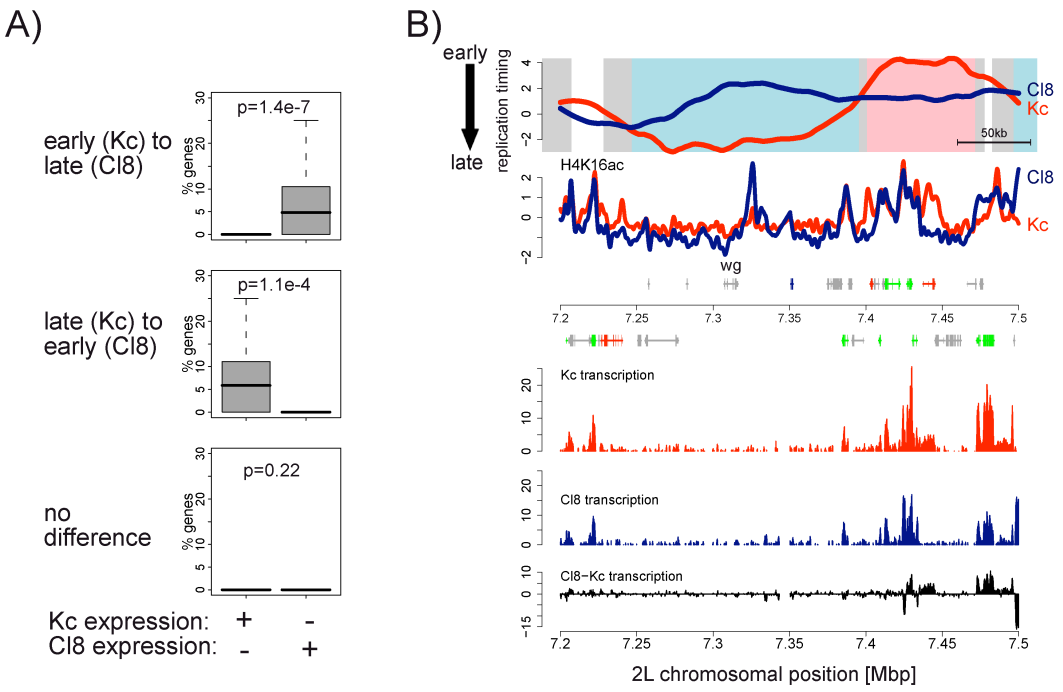

**Supplementary Figure 1: Reproducibility of replication timing analysis.** A) Pairwise correlation between two biological replicates of replication timing of Kc cells on Affymetrix tiling arrays illustrating high reproducibility between individual experiments. B) Correlation between detection on Affymetrix arrays (where early and late replicating DNA are hybridized separately) and detection on a NimbleGen array (where early and late replicating DNA were labeled with different dyes and hybridized simultaneously). This comparison illustrates the robustness of assay and detection as it is largely platform and labeling independent. Data from both platforms were summarized in sliding windows of 1000bp (stepsize = 500bp) for direct comparison. C) Autocorrelation analysis of replication timing in Kc cells (1000bp sliding windows, stepsize = 500bp) on chromosome 2L. The x-axis shows the distance on the chromosome in kb as calculated from the number of sliding windows. The autocorrelation function (ACF) indicates the degree of similarity between neighboring data-points and illustrates that replication timing extends over large regions. Interestingly, we observe periodic negative (early vs. late) and positive (both data-points early or both data-points late) autocorrelation, which suggests a certain degree of periodicity in the replication timing across a chromosome. D) Comparison of replication timing profiles from two biological replicates illustrating high reproducibility between repeats at the level of replication timing profiles and the location of initiation zones. The x-axis represents the chromosomal position (in mega basepairs), the left y-axis the first replication timing measurement (black line) and the right y-axis the second replication timing measurement (blue line), which was plotted on a shifted scale for easier comparison.



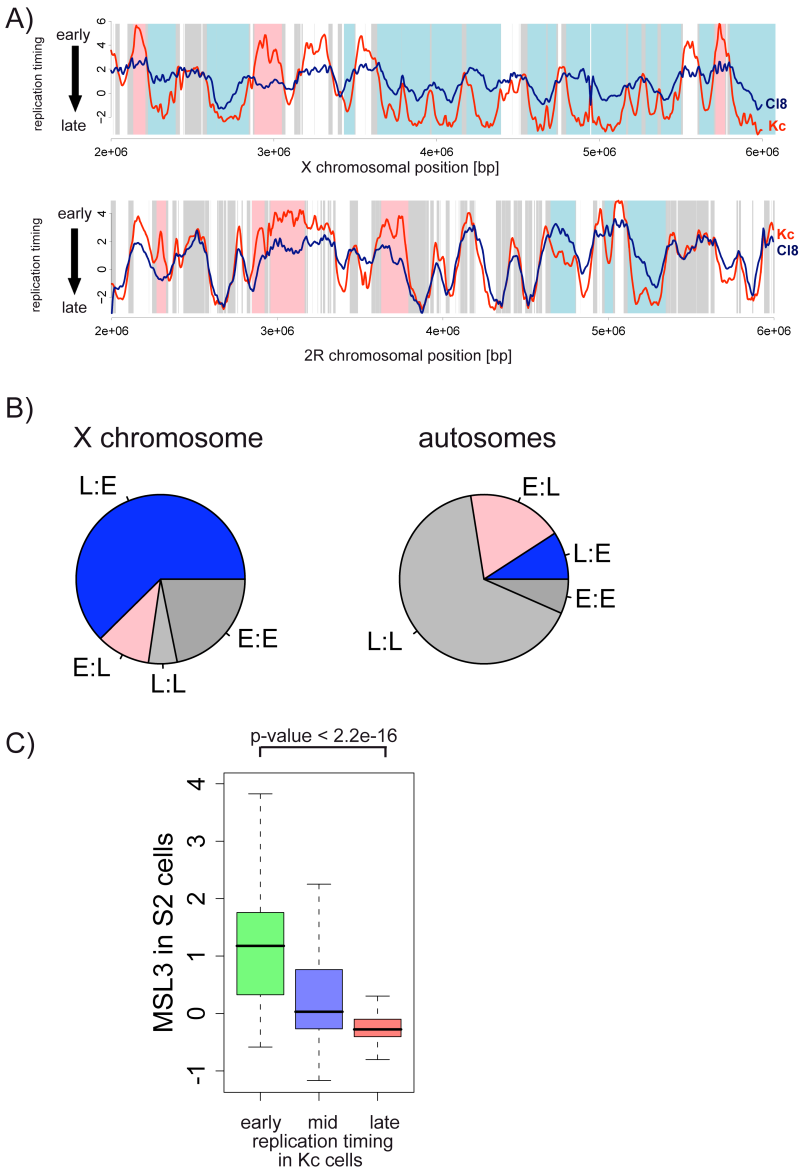

**Supplementary Figure 2: Definition and analysis of initiation zones.** A) Experimental enrichments for small nascent strand DNA purified from Kc cells was verified at a previously studied origin of replication compared to a site 10 kb distal (oriDalpha, (Sasaki et al. 1999; MacAlpine et al. 2004)). Small nascent strand abundance was quantified by real-time PCR and normalized to genomic DNA from G2 cells. B) 87% of previously defined early origins on chromosome 2L (MacAlpine et al. 2004) overlap with initiation zones identified in our analysis (green, 39/45; red = no overlap, 6/41), illustrating that most predicted origins by MacAlpine *et al.* can be identified by high resolution replication timing. C) Depicted are "replication domains" as regions between two termination zones. The circle labels the initiation zones and the horizontal grey lines the extend of domains, termination zones are located at both ends of each domain (small vertical bars). The black line represents the replication timing profile. D) The size of replication domains (regions likely replicated by forks initiating from one initiation zone) is a function of initiation time revealing that early initiation results in larger replication domains providing evidence for fork progression throughout S phase. The boxplot illustrates the size distributions of domains with initiation zones in early (replication timing >1), mid (replication timing between 1 and -1) and late (replication timing <1) S phase. The black bar indicates the median of each distribution. p-value was calculated using the Wilcoxon rank sum test. E) We defined replication timing regions as sequences that replicate either early, mid or late regardless of time of initiation. Segmentation into these groups was done using a three state Hidden Markov Model (see methods for details). Shown is a representative regions with the segmentation results shown as colored background behind a replication timing profile (green = early, blue = mid, red = late).



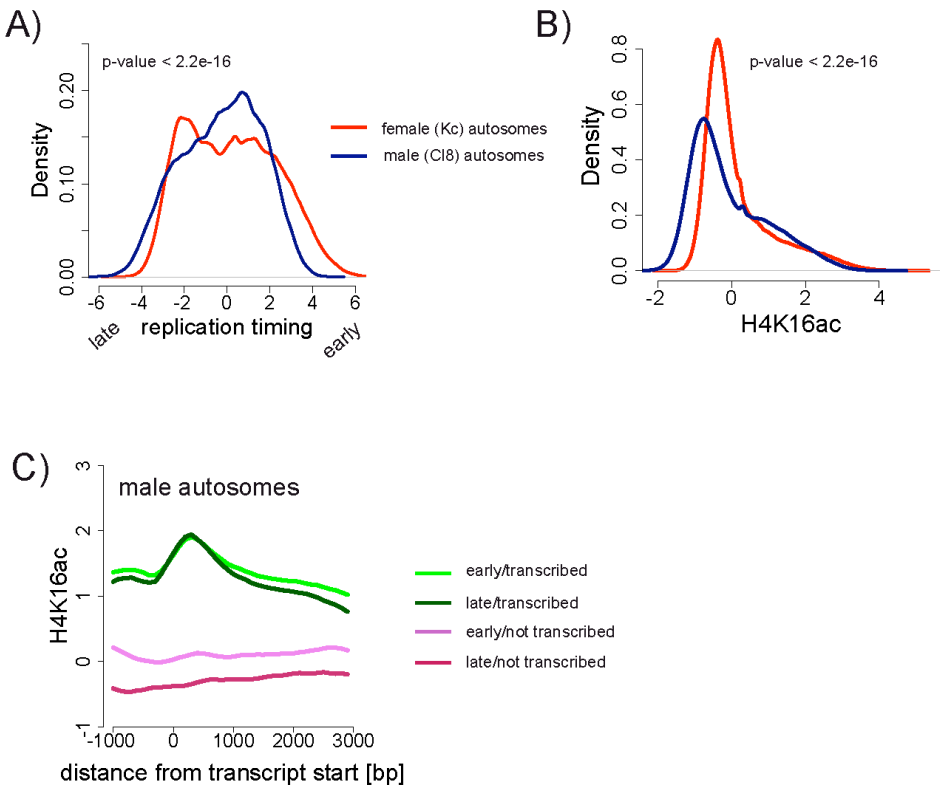

**Supplementary Figure 3: Logistic regression analysis for transcriptional state and H4K16 acetylation state versus replication timing.** The relationship between the probability of expression and replication timing and the probability of H4K16 acetylation and replication timing was analyzed using logistic regression as in (Schubeler et al. 2002), which is a statistical method used if there are only two potential outcomes for one of the two variables. A-B) Chromosomal tiles of 1000 bp were defined as transcribed or non-transcribed based on signal intensity. Then we rank chromosomal tiles based on their replication timing (x-axis) and define the percentage of transcribed windows, which is displayed as the likelihood of expression (y-axis). This logistic regression analysis confirms the strong correlation for transcription (thick black line = logistic regression curve, thin black lines = 95% confidence interval, grey bars represent individual bins (out of 50)). C-D) Logistic regression analysis for H4K6ac versus replication timing. We divided chromosomal tiles into H4K16 acetylated versus non-acetylated based on signal above or below the median and determined the percentage of tiles in this chromatin state (y-axis) relative to replication timing (x-axis). Notably this reveals an even stronger correlation between H4K16 acetylation (C-D) and early replication compared to transcription and early replication (A-B).



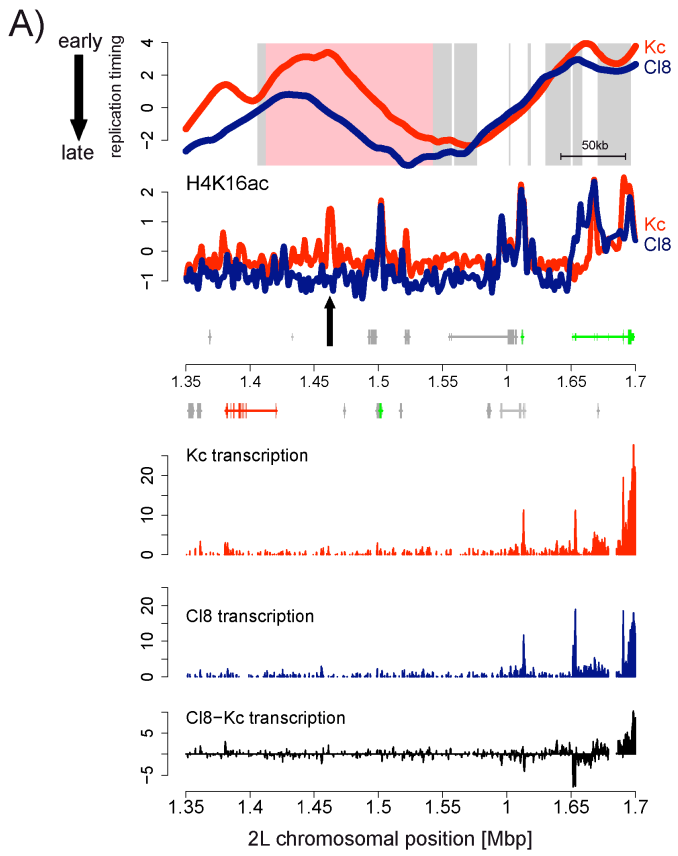

**Supplementary Figure 4: Gene expression versus replication timing differences.** A) To illustrate how differences in replication timing relate to on/off states of genes we identified genes that change expression status between Kc and Cl8 cells (genes expressed only in Kc or only in Cl8 cells). Then we divided the genome into regions that replicate earlier in Kc cells (top plot), those that replicate earlier in Cl8 cells (middle plot), and regions which show no difference in replication timing (bottom plot) based on HMM segmentation (see methods for details). For every single region, we determined the number of genes for each expression status, and the percentage of genes that are expressed only in Kc cells (left row of boxplots), or only in Cl8 cells (right row of boxplots). The distribution of these percentages for all regions within each of the three groups is shown as boxplots, with the black line representing the median. This analysis shows that transcription changes are more frequent in regions with differential replication timing and that they have the same directionality as replication changes, with a switch to early replication entailing gene activation. Such a trend is not present at regions that do not change timing ("no difference"). On average, less than 10% of genes in differentially replicating regions show differential expression. p-values were calculated using the Wilcoxon rank sum test. B) Replication timing, transcription and H4K16 acetylation for a region on chromosome 2L that replicates earlier in Cl8 cells containing genes important for development. wg=wingless gene. Labeling is similar to Figure 3A of the main text.

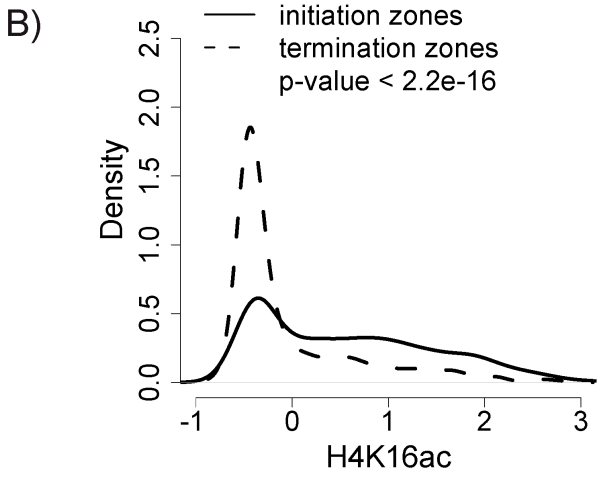



**Supplementary Figure 5: X-linked replication timing differences.** A) Replication timing profiles for a part of chromosome X and 2R showing the almost exlcusive early replication (blue background) of the male X which is not present on autosomes. Background coloring denotes regions that replicate earlier in Cl8 cells (L:E, blue), regions that replicate earlier in Kc cells (E:L, pink) and regions that replicating similarly in both cell types (grey). B) Switches to early replication on the X occur in regions that lack transcription. Percentage of the genome on the X chromosome (left) and autosomes (right) in regions which contain no active genes (based on Affymetrix expression arrays, see methods for details) in any cell type. Regions that replicate earlier in Cl8 cells (L:E, blue), earlier in Kc cells (E:L, pink), early in both cell types (E:E, darkgrey) or late in both cell types (L:L, grey) are compared. This illustrates that 62% of X-linked, non-transcribed sequences switch from late to early replication, leaving only 5% of non-transcribed sequences replicating late in both cell types. However, most non-transcribed sequences on autosomes replicate late in both cell types (66%). C) MSL3 binding in male embryonic S2 cells (Alekseyenko et al. 2006) occurs predominantly at regions that replicate early in female embryonic Kc cells (this study). Replication timing in female Kc cells was grouped in early, mid and late according to the HMM model (see methods). For each region the binding of MSL3 in male S2 cells is calculated. For each replication timing group the distribution of MSL3 binding is plotted, with the black line representing the median of the distribution.

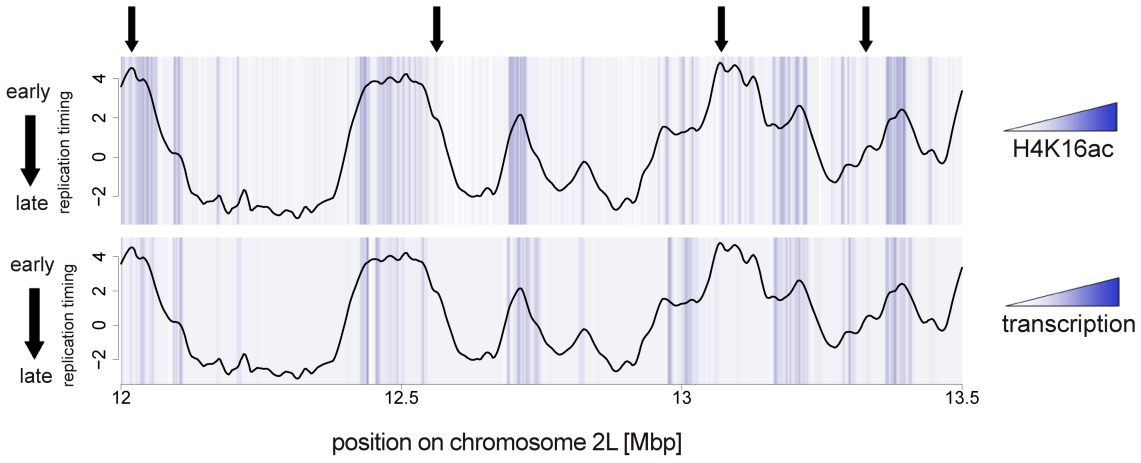



**Supplementary Figure 6: Distribution of H4K16ac on autosomes.** A-B) Comparison of replication timing (A) and H4K16ac (B) of autosomes in Cl8 (male, blue lines) and Kc (female, red lines) cells. p-values were calculated using the Wilcoxon rank sum test. C) Distribution of the enrichment for H4K16ac in Cl8 cells throughout active early replicating (green line), active late replicating (darkgreen line), inactive early replicating (violet line) and inactive late replicating (violetred line) autosomal genes aligned at their transcription start site (see methods for details).




**Supplementary Figure 7: Acetylated H4K16 is enriched at initiation zones.** A) Replication timing, H4K16ac and transcription comparison of a region on chromosome 2L. Labeling is similar to Suppl. Fig. 4B. Note the strong enrichment for H4K16ac only in Kc below an initiation zone, which is active only in Kc cells (marked by an arrow). B) Density plot of the enrichment of H4K16ac averaged within initiation or termination zones on autosomes. p-value ($p < 2.2e-16$) was calculated using the Wilcoxon rank sum test.




**Supplementary Figure 8: Chromatin and transcription mark zones of replication initiation.** Shown is a Kc replication timing profile (line) with a heatmap for H4K16 acetylation (top graph) or transcription (bottom graph) in the background. H4K16ac is enriched at initiation zones, which frequently correlates with high levels of transcription. Interestingly however, while H4K16ac is always present at sites of transcription, it also occurs in non-transcribed regions, many of which are initiation zones (some examples are indicated with black arrows on the top of the graph). White indicates absence of H4K16ac/transcription, darker colors indicate higher H4K16ac/transcription levels. x-axis= chromosomal position in mega basepairs, y-axis= replication timing (log2).

## 3.2. Connection of differential replication timing to nuclear organization

### 3.2.1. Summary

The control of replication timing and transcription might be intimately linked to the organization of the genome within the nucleus (reviewed in (Chakalova et al. 2005)). Especially the nuclear periphery has been implicated in gene silencing and is often the site of late replication, as has recently been demonstrated for the *Drosophila* genome (Pickersgill et al. 2006). Single gene studies in mammalian cells have shown that the location of a gene relative to the nuclear periphery can be dynamic and correlates with both replication timing and gene activity (reviewed in (Zink 2006)).

Having generated replication timing profiles of the *Drosophila* genome in two cell lines representing different differentiation states, I aimed at comparing these data to nuclear organization. I determined the location of four differentially replicating regions relative to the nuclear envelope in *Drosophila* Kc and Cl8 cells by DNA FISH. Three of those regions locate closer to the nuclear periphery when late then when early replicating. Those differences in replication timing and nuclear position also correlate with differences in transcription levels of several genes within the analyzed regions of the genome. This suggests that the position of a sequence within the nucleus of *Drosophila* cells can be dynamic in relation to replication timing and gene expression.

### 3.2.2. Introduction

The location of a sequence within the nucleus seems to be important for transcription and DNA replication in eukaryotic cells. In *S. cerevisiae*, the nuclear periphery is a repressive compartment (reviewed in (Akhtar et al. 2007)), but can also be activating if a gene locates close to the nuclear pores (Taddei et al. 2006). In animal cells, the nuclear envelope is coated by the nuclear lamina, which is made up of a meshwork of lamin and lamin associated proteins (reviewed in (Akhtar et al. 2007)). Several reports have suggested a strong correlation between replication timing and nuclear organization. Replication domains might even consist of replication factories, stable chromosomal units occupying defined sites within the nucleus at specific times of the cell cycle (reviewed in (Zink 2006)). The location of late replicating, inactive genes to the nuclear periphery has been reported in mammalian cells, where the same sequence can move towards the interior of the nucleus when active and early replicating (Zhou et al. 2002a; Zhou et al. 2002b). However, it is not clear what establishes peripheral location of a sequence, and if it is cause or consequence of late replication. Differential nuclear location and replication timing of imprinted loci seems to be independent of DNA methylation and imprinted gene expression (Gribnau *et al.* 2003), suggesting that chromatin modifications might establish differential nuclear localization and replication timing. Recently, a very strong correlation was described between late replication and association with the nuclear lamina in *Drosophila* cells (Pickersgill et al. 2006). In this genome-wide Dam ID study using cDNA arrays, about 300 genes in *Drosophila* Kc cells bind to B-type lamin. Those genes often cluster together within large gene-poor regions and are repressed and late replicating (Pickersgill et al. 2006). However, it was not clear if those lamin bound late replicating sequences could show dynamic replication timing, and if possible differences in replication timing would correlate with differences in the location of a sequence within the nucleus.

To investigate the correlation between replication timing differences and position within the nucleus in *Drosophila* cells, I assayed the position of four differentially replication regions in Kc and Cl8 cells using DNA FISH. I chose four regions which I identified to show significant differences in replication timing between Kc and Cl8 cells based on previous analysis (see chapter 3.1). Three of four regions located closer to the nuclear periphery when replicating late than when replicating early, suggesting a strong

correlation between replication timing differences and differences in location of a sequence relative to the nuclear periphery in *Drosophila* cells.

### 3.2.3. Results and discussion

To investigate the position within the nucleus of genes which show dynamic replication timing in *Drosophila* cells, I generated DNA FISH probes of about 5kb within four large regions of differential replication timing between Kc and Cl8 cells. The location of those probes in the genome relative to previously identified differentially replicating regions and transcription differences is shown in Figure 1. These probes were hybridized to nuclei of the two cell lines and the resulting DNA FISH stainings were analyzed using confocal microscopy to obtain three-dimensional images of the FISH signals within the cells.

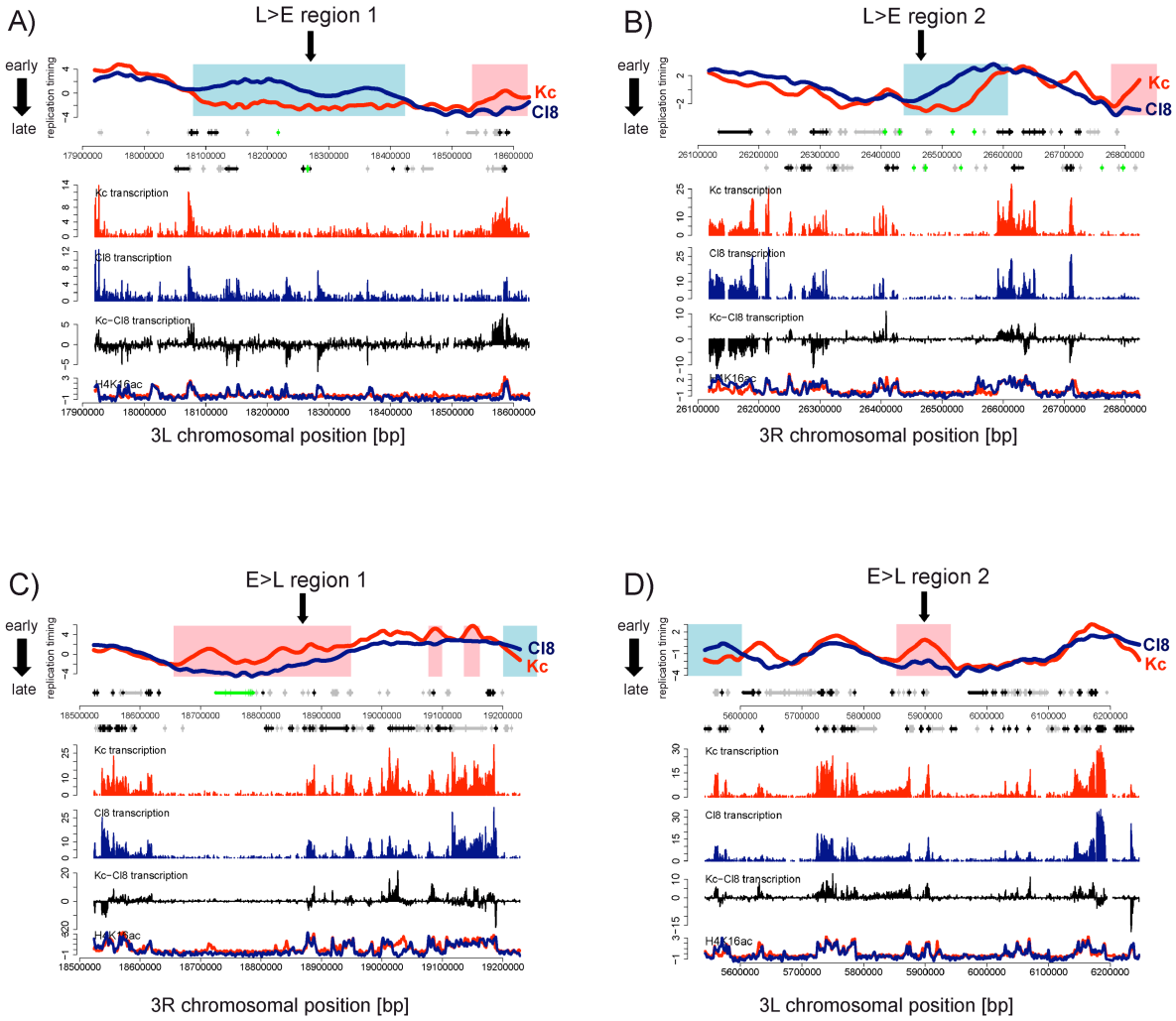


**Figure 1: Replication timing, transcription, and H4K16ac in Kc and Cl8 cells around four differentially replicating regions.** Replication timing profiles of Kc (red) and Cl8 (blue) cells for four regions on chromosome 3 containing L>E region 1 (A), L>E region 2 (B), E>L region 1 (C), or E>L region 2 (D). X-axis = chromosomal position in basepairs, y axis = log2 (early/late replication). Background coloring denotes regions that replicate earlier in Cl8 cells (blue), regions that replicate earlier in Kc cells (pink) and regions replicating similarly in both cell types (white). The black arrows mark the exact positions of the probes used for DNA FISH within each differentially replicating region. Annotated genes are displayed below the profile (boxes=exons, lines=introns, small boxes=UTRs) and colored by their lamin binding status in Kc cells as determined by (Pickersgill et al. 2006) (green= lamin target gene, black= not a lamin target gene, grey= lamin binding not analyzed). Transcription levels of Kc (red) and Cl8 (blue) cells measured by tiling arrays are displayed on the same scale below, including transcription level differences (Kc-Cl8 transcription, black). At the bottom, H4K16ac profiles for Kc (red line) and Cl8 (blue line) cells are displayed, y-axis = log2 (H4K16ac antibody bound DNA/Input DNA).

The first probe (L>E region 1) is in an about 300kb large differentially replicating region which replicates late in *Drosophila* Kc cells, but displays strongly advanced replication timing in Cl8 cells. This correlates with the transcriptional activation of several genes and non-annotated transcripts within L>E region 1 in Cl8 cells (Figure 1A). 46% of the FISH signals for this region are located at the periphery of the nucleus in Kc cells (Figure 2A and C). This rather frequent association with the nuclear periphery is consistent with the late replication and lack of transcription in this region in Kc cells. To test if the frequency of association with the periphery is dynamic in this region, I performed DNA FISH using the same probe in Cl8 cells. There L>E region 1 replicates early (Figure 2B), and only 31% of FISH signals are located close to the nuclear periphery (Figure 2D), which indicates a random distribution within the nucleus relative to its periphery. These results demonstrate that the position of a sequence relative to the nuclear periphery can by dynamic in *Drosophila* cells in relation to differences in replication timing and transcription.

Another region replicating earlier in Cl8 cells (L>E region 2, Figure 1B) showed very strong association with the nuclear periphery in Kc cells (Figure 2E), consistent with the high number of lamin target genes observed in this region in a previous study (Pickersgill et al. 2006). Surprisingly, more than half of all FISH signals located close to the nuclear periphery not only in Kc cells, but also in Cl8 cells where the region replicates early (Figure 2E). Interestingly, this differentially replicating region did not contain any differences in transcription between the two cell types (Figure 1B). This suggests that changes in position of a sequence relative to the nuclear periphery are less likely to occur when replication timing is advanced without transcriptional activation of any genes within the differentially replicating region.

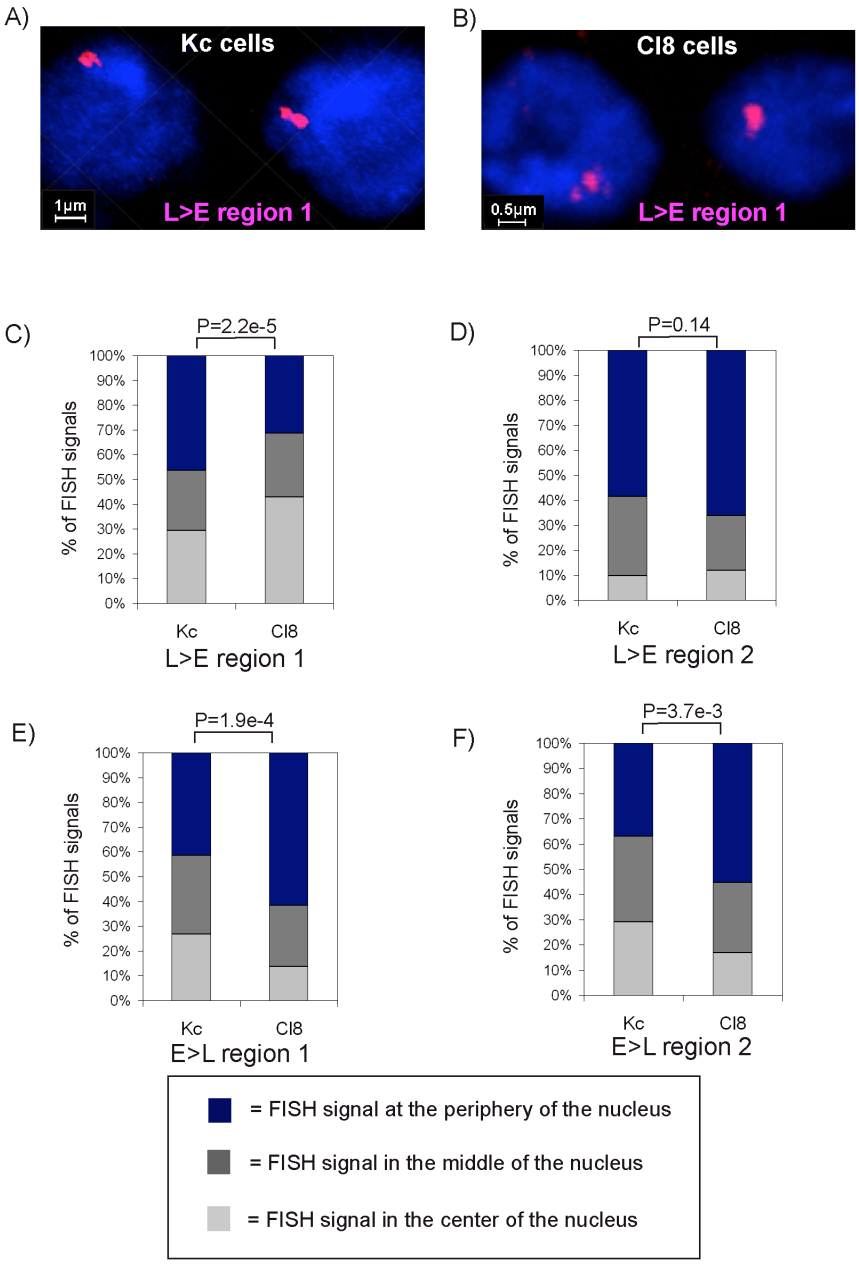



**Figure 2: Three of four replication timing differences coincide with differences in position relative to the nuclear periphery.** A) DNA FISH with a probe detecting L>E region 1, which replicates earlier in Cl8 than Kc cells, in Kc cells. Blue=DAPI, Red=FISH signal. B) DNA FISH with a probe detecting L>E region 1, which replicates earlier in Cl8 than Kc cells, in Cl8 cells. Blue=DAPI, Red=FISH signal. C) Quantification of FISH signals for L>E region 1 (replicating earlier in Cl8 cells) in Kc (left, number of total signals counted (n) = 394) and Cl8 (right, number of total signals counted (n) = 314) cells. 46% of FISH signals in Kc cells lie close to the nuclear periphery (left, blue bar), while only 31% of FISH signals in Cl8 cells are found associated with the nuclear periphery (right, blue bar). P-value=2.2e-5. D) Quantification of FISH signals for L>E region 2 (replicating earlier in Cl8 cells) in Kc (left, n = 310) and Cl8 (right, n = 250) cells. 58% of FISH signals in Kc cells lie close to the nuclear periphery (left, blue bar), and 66% of FISH signals in Cl8 cells are found associated with the nuclear periphery (right, blue bar). P-value=0.14. E) Quantification of FISH signals for E>L region 1 (replicating earlier in Kc cells) in Kc (left, n = 171) and Cl8 (right, n = 131) cells. 42% of FISH signals in Kc cells lie close to the nuclear periphery (left, blue bar),

and 62% of FISH signals in Cl8 cells are found associated with the nuclear periphery (right, blue bar). P-value=1.9e-4. F) Quantification of FISH signals for E>L region 2 (replicating earlier in Kc cells) in Kc (left, n = 165) and Cl8 (right, n = 125) cells. 37% of FISH signals in Kc cells lie close to the nuclear periphery (left, blue bar), and 55% of FISH signals in Cl8 cells are found associated with the nuclear periphery (right, blue bar). P-value=3.7e-3. p-values were calculated using the Wilcoxon rank sum test.

Two regions which replicate earlier in Kc compared to Cl8 cells (Figure 1C-D) were analyzed by DNA FISH in both cell types. 42% of FISH signals detecting E>L region 1 locate close to the periphery in Kc cells. This suggests a preference for peripheral location of this sequence already in Kc cells where it replicates early (Figure 2E, left). However, upon late replication in Cl8 cells, 62% of FISH signals are detected at the periphery (Figure 2E, right). This suggests that despite a preference for peripheral location of this region when replicating early, a shift to late replication in a different cell type still correlates with an increased association of this region with the nuclear periphery. E>L region 2 is part of differentially replication region which is only about 90kb in size (Figure 1D). It shows a random distribution within the nucleus relative to its periphery in Kc cells where it replicates early (Figure 2F, left). However, when it replicates late in Cl8 cells, 55% of FISH signals for E>L region 2 lie at the nuclear periphery (Figure 2F, right). This suggests that a delay in replication timing of even a small region can correlate with increased association with the nuclear periphery. Furthermore, I could detect higher levels of transcription of some genes in Kc cells compared to Cl8 cells within both differentially replication regions which replicate earlier in Kc cells (Figure 1C-D). Together with the lack of position differences of L>E region 2 described above, this could suggest that both, replication timing and transcription differences together correlate better with differences in position relative to the nuclear periphery than do differences in replication timing alone.

By comparing the position within the nucleus of four differentially replicating regions in two *Drosophila* cell types, I show that three sequences which associate with the nuclear periphery to a different degree locate further away from the periphery when replicating early than when replicating late. These differences in replication timing and nuclear position also correlate with differences in transcription levels between the cell types. I propose that in *Drosophila*, like in mammalian cells, most differences in replication timing correlate with a reorganization of sequences in the nucleus.

One region, which is most tightly associated with the nuclear periphery, and does not show transcription in any cell type, still shows a shift to earlier replication in Cl8 cells

compared to Kc cells, where it replicates late. This suggests that some replication timing differences can occur independently of transcription, and even despite a location close to the nuclear periphery a region can replicate early. It is possible that this early replication correlates with differences in chromatin structure. Interestingly, one side of the differentially replicating region, in which L>E region 2 lies, shows increased acetylation of Lysine 16 on histone H4 in Cl8 cells (Figure 1B, right side of the blue box). This mark might induce early replication even at a peripheral location in the nucleus. Recently, it has been shown that the dosage compensated male X chromosome, which shows strong H4K16 hyperacetylation, associates with nuclear pores (Mendjan et al. 2006). It is possible that L>E region 2 changes its position within the nucleus upon early replication in Cl8 cells towards the nuclear pores, and not towards the interior of the nucleus. This difference in position would not be detected by DNA FISH analysis. Genome-wide lamin binding profiles in Kc and Cl8 cells would help to elucidate this question and would be required to determine the global extend of the correlation between replication timing differences and differences in association with the nuclear lamina.

In summary, I have shown that three of four tested genomic regions show dynamic position in the nucleus relative to the nuclear periphery in relation to dynamic replication timing and gene expression. This suggests that the location of a sequence relative to the nuclear periphery can be dynamic in *Drosophila* cells and correlates strongly with dynamic replication timing and transcription.

## 3.2.4. Materials and Methods

### DNA FISH probe preparation.

FISH probes were generated by labeling PCR products of ~5kb in length with digoxigenin using the DIG Nick translation kit (Roche).

L>E region 1 (covering gene CG4345):
Position: 18269842 bp to 18275309 bp on chromosome 3L
Sense primer: 5'- GAACCACCAACCACAGGAAC -3'
Antisense primer: 5'- GCATCAGGAGCAGAGGAAAC -3'

L>E region 2 (covering gene CG31015):

Position: 26468533 bp to 26473470 bp on chromosome 3R

Sense primer: 5'- GTTCGTGGTTCATTCCCTTG -3'
Antisense primer: 5'- TCAGTCCGTTGTATTGATGGAG -3'

E>L region 1 (covering gene p53):

Position: 18873720 bp to 18878247 bp on chromosome 3R

Sense primer: 5'- GTTGTCGTTGCCGTCTCC-3'

Antisense primer: 5'- GTAGAAACCCACCGTTATTCC-3'

E>L region 2 (covering gene CG32407):

Position: 5892798 bp to 5897065 bp on chromosome 3L

Sense primer: 5'- GTAGCAGGACGAGAAGATGG-3'

Antisense primer: 5'- GCTCAATCCAGAAGACAAGG-3'

**DNA Fluorescent In Situ Hybridization (FISH).**

Cells were grown on polylisine-coated glass coverslips, fixed with permeabilized in 4%PFA+0.3%Triton X-100 for 12min, RNAse treated for one hour at 37˚C, and denatured in 50% formamide / 2X SSC (pH 7.2) for 30 min at 80°C. Then cells were hybridized with the probe overnight at 42°C, blocked in 4X SSC / 0.1% Tween / 5% BSA (Gibco) for 15 min at RT, and incubated with a sheep anti-digoxigenin-rhodamine antibody (Roche). After incubation with the secondary Alexa 546 Donkey anti-Sheep antibody (Molecular Probes), DNA was counterstained with 0.04 ug/ml DAPI.

**Microscopy and image analysis.**

DNA FISH stainings were analyzed using a laser scanning confocal microscope LSM510 META (Zeiss) and LSM510 software. We recorded a *z* series of either 0.2µm or 0.3µm slices for approximately 10-50 cells at a time. The slices were imported into the IMARIS 5.9.0. alpha software and displayed in 3D. The distance of each FISH signal to the nuclear periphery (as determined by DAPI staining) was measured in IMARIS 5.9.0. alpha by hand, and normalized to the diameter of the nucleus it resides in. A p-value was calculated between all signal to periphery distances in Kc and Cl8 cells for each probe. For display in bar charts, the signal to periphery distances of each probe in Kc or Cl8 cells were divided into three distance-to-periphery groups based on equal volume.

## 3.2. Heterochromatin protein 1 (HP1) modulates replication timing of *Drosophila* heterochromatin

### 3.3.1. Summary

The results discussed in chapter 3.1 showed that replication timing can be highly dynamic between two epigenetic states with different gene expression and chromatin modification profiles. However, it is not clear if this dynamic replication timing is driven by changes in chromatin modifications or chromatin binding proteins, or if replication timing itself can affect the association of chromatin modifiers with the DNA. A candidate protein for regulating replication timing is heterochromatin protein 1 (HP1). It binds to DNA throughout heterochromatin and as well within some euchromatic genes (Greil et al. 2003; de Wit et al. 2007). While many roles of HP1 have been described, its most important role seems to be silencing genes and repeats. This role of HP1 might be mediated through its interaction with H3K9me2, which is associated with chromatin compaction (reviewed in (Peters et al. 2005)). Since transcription and replication timing might be connected via chromatin, it is possible that HP1 also delays replication timing of heterochromatin. Interestingly, HP1 has been shown to interact with ORC2, which binds to replication origins (Pak et al. 1997), thereby suggesting a connection of HP1 with the initiation of DNA replication.

To determine if HP1 plays a role in modulating replication timing, I used RNAi to reduce HP1 protein levels in *Drosophila* Kc cells. This led to an advance in replication timing of centromeric repeats, suggesting that the late replication of those regions depends on high HP1 and H3K9me2 levels. Chromosome 4 and pericentric heterochromatin also showed strong differences in replication timing after HP1 knockdown. Surprisingly, most of those differentially replicating regions replicated later in HP1 RNAi cells. This suggests that HP1 might facilitate origin activation within those regions.

Since early replication correlates with gene expression, it is not clear if this role of HP1 is mediated by changes in transcription after HP1 knockdown, or independent of gene expression. I measured transcription level changes during HP1 RNAi and detected transcription differences in some, but not all regions with delayed replication timing. However, I cannot measure the transcription of repeats, since they are not represented on the tiling array. A possible mechanism of how HP1 allows the expression of heterochromatic genes is that it might silence nearby repeats, whose transcription would otherwise interfere with gene expression. It is conceivable that transcription of repetitive elements might also interfere with early origin firing, thereby leading to a delay in replication timing. To answer this question, genome-wide maps of the exact location of replication origins will be required.

### 3.3.2. Introduction

The replication of eukaryotic genomes starts at many origins of replication, which are distributed throughout chromosomes and can fire at different times during S phase. This process can be dynamic between different cell types representing different epigenetic states (reviewed in (Schwaiger et al. 2006)). Replication origins have been mapped at high resolution throughout the yeast genome, and share a consensus sequence motif (MacAlpine et al. 2005; Nieduszynski et al. 2006). However, metazoan origins of replication have been difficult to identify. The few initiation start sites which have been mapped lack a consensus sequence (MacAlpine et al. 2004). Together with the observation that the initiation of DNA replication can be dynamic between cell types, this suggests that replication initiation in higher eukaryotes is determined by epigenetic features rather than a certain DNA sequence.

Initiation of DNA replication at different times and places results in a highly cell type specific replication timing program. Early microscopic studies have shown that in general, euchromatin replicates before heterochromatin. Euchromatin refers to the part of the genome which is rich in active genes, and contains almost no repetitive elements. The late replicating heterochromatin consists of centromeric and telomeric repeats and pericentric and intercalary heterochromatin (reviewed in (Zhimulev *et al.* 2003)). There is a gradual transition between euchromatin and heterochromatin. However, for the sake of simplicity, I will distinguish between euchromatin and two forms of heterochromatin in *Drosophila*: Centromeric repeats and pericentric heterochromatin. Centromeric repeats are defined as the regions around centromers containing no known genes, but very specific repetitive elements. I refer to pericentric heterochromatin as regions of the genome between euchromatin and centromeric repeats. These regions contain many active genes, but at the same time are still repeat rich and display certain properties of heterochromatin.

Within euchromatin, early replication correlates with high gene expression levels and histone acetylation (see chapter 3.1). There is evidence for a role of hyperacetylated chromatin in facilitating origin activation and thereby early replication timing (Aggarwal et al. 2004; Calvi et al. 2007; Goren et al. 2008). However, histone acetylation or other chromatin modifications do not seem to be the exclusive determinants of replication timing. It is possible that other chromatin binding proteins also play a role in facilitating or inhibiting origin activation. Indeed, heterochromatin protein 1 (HP1) interacts with ORC2 (Pak et al. 1997), which is a member of the origin recognition complex and important for

origin activation (reviewed in (Bell et al. 2002)). This might implicate a role for HP1 in recruiting the origin recognition complex to chromatin. HP1 has initially been described as a protein binding to heterochromatin, where it helps to spread the repressive H3K9me2 histone modification (reviewed in (Hediger et al. 2006)). HP1 could facilitate origin firing within heterochromatin, but at the same time also delay replication timing by supporting heterochromatin formation.

The HP1 gene is also called SuVAR3-5, since it was first discovered as a suppressor of position effect variegation (PEV) in *Drosophila* (Eissenberg *et al.* 1990). PEV describes the observation, that euchromatic genes acquire a variegated expression pattern when positioned close to or within heterochromatin (Muller *et al.* 1930). This silencing is suppressed when certain genes are mutated, for example HP1 or SuVAR3-9 (Wustmann *et al.* 1989), the histone-methyltransferase which methylates Lysine 9 on Histone H3 (Schotta et al. 2002). This suggests that their function is to inactivate genes or repetitive elements.

HP1 seems to bind to H3K9 methylation independent of the enzyme which established the modification, since HP1 can bind many, mostly non-pericentric regions of the genome independent of SuVAR3-9 (Greil et al. 2003; de Wit et al. 2005; de Wit et al. 2007). Furthermore, on chromosome 4, HP1 binding depends on dSETDB1 (Seum et al. 2007; Tzeng et al. 2007). However, it is also possible that in some places HP1 binds chromatin independently of H3K9 methylation. The binding patterns of HP1 in *Drosophila* cells have recently been studied in great detail using DamID and microarray technology. While pericentric HP1 binding is stable throughout development, HP1 binds many, mostly long, genes in a developmental stage specific way (Greil et al. 2003). HP1 target genes which lie in pericentric heterochromatin (also termed “heterochromatic genes”) tend to be highly expressed and also have a different HP1 binding pattern than euchromatic HP1 target genes, which show average expression levels (de Wit et al. 2007). Heterochromatic genes tend to be surrounded by repetitive sequences, which can also lie within the introns of those genes (Devlin *et al.* 1990; Schulze *et al.* 2005). Upon loss of HP1 or translocation to euchromatin, the expression of heterochromatic genes is reduced in *Drosophila* (Wakimoto *et al.* 1990; Clegg *et al.* 1998; Lu *et al.* 2000; Schulze et al. 2005). It is possible that this is due to the loss of heterochromatin-mediated silencing of associated repeats.

Besides covering pericentric heterochromatin, HP1 is highly enriched on chromosome 4, and also has a preference for binding to the male X chromosome (de Wit et al. 2005). In

addition, knock-down of HP1 in flies results in preferential lethality of male flies (Liu *et al.* 2005b). HP1 bound regions often cover large, up to 100kb long regions in the genome, which never overlap with large polycomb regions (de Wit et al. 2007). Interestingly, in *Drosophila* polytene chromosomes and Kc cells, HP1 target regions often overlap with regions bound by the Suppressor of Underreplication (SuUR) protein (Koryakov et al. 2006; Pindyurin et al. 2007). SuUR binds to underreplicated and late replicating regions on *Drosophila* polytene chromosomes (Makunin et al. 2002), and to late replicating regions in Kc cells (Pindyurin et al. 2007). SuUR might play a role in suppressing origin firing since it reduces the endo-replication of heterochromatic regions on polytene chromosomes (Belyaeva et al. 1998). HP1 has been shown to interact with SuUR and could be crucial for mediating chromatin binding of the SuUR protein (Pindyurin et al. 2008). Therefore it is conceivable that loss of HP1 from chromatin might induce an advance in replication timing of some heterochromatic regions.

To determine possible effects of HP1 on replication timing, we used RNA interference (RNAi) to reduce HP1 protein levels in *Drosophila* Kc cells. This resulted in advanced replication timing of pericentric repeats as measured by immunofluorescence and qPCR. Using high resolution tiling arrays, we found that about 10% of the genome changed replication timing after HP1 knockdown. The biggest part of those changes was a delay in replication timing on chromosome 4 and heterochromatic regions enriched in HP1 binding, most of which lie close to the centromeres. Finally, transcription analysis of control and HP1 RNAi cells showed that some, but not all, replication timing changes correlate with transcription changes.

### 3.3.3. Results

To reduce the levels of HP1 protein in *Drosophila* Kc cells by RNA interference (RNAi) I added double-stranded RNA (dsRNA) specific for the HP1 coding sequence to the cells for 8 days. I observed a strong depletion of HP1 protein by western blot in dsRNA treated cells (Figure 1A) and also by immunofluorescence (Figure 1B). To determine if replication timing of heterochromatin, which in *Drosophila* cells clusters in the chromocenter, was affected by HP1 depletion, I pulse-labeled cells with BrdU for one hour and detected replicating DNA by immunofluorescence with an antibody recognizing BrdU (Figure 1C). I detected BrdU in about 30% of all nuclei. Among those I observed three different patterns of BrdU incorporation, characteristic of early (only euchromatin shows BrdU), mid (the whole nucleus shows BrdU), and late (BrdU is detected only in the chromocenter) replication (Figure 1C) in both, control and HP1 knockdown, cells.

The percentage of BrdU stained cells showing each pattern was about equal in control cells (Figure 1D). However, I observed a reduced number of cells in late S phase after HP1 RNAi (Figure1D). This coincided with an increased number of cells showing the mid S phase pattern (Figure 1D). These results suggest that some cells might be blocked in their replication before late S phase, or that heterochromatin shows an advance in replication timing, replicating at the same time as euchromatin and therefore resulting in BrdU staining throughout the whole nucleus.

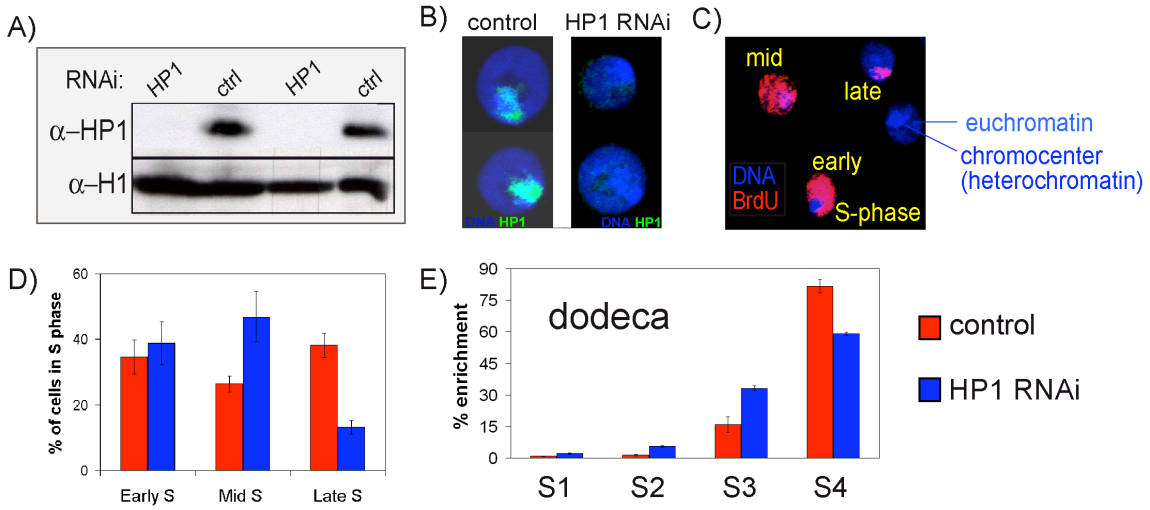



**Figure 1: Knockdown of HP1 in Kc cells by RNA interference.** A) Western blot detecting HP1 protein in untreated control cells (ctrl) and cells treated with dsRNA directed against HP1 (HP1) for 7 days. H1 shows equal loading of total protein. B) Cytological analysis of HP1 localization by immunofluorescence with an antibody recognizing HP1 protein in Kc cells. HP1 localizes mainly to the chromocenter in control cells (ctrl, left), but is detected only at very low levels in HP1 knockdown cells (HP1, right). C) Cytological analysis of replication timing. Kc cell nuclei with three different patterns of BrdU incorporation after pulse-labeling are shown. D) Quantification of the percentage of BrdU positive nuclei in early, mid and late S phase based on their BrdU staining pattern in control and HP1 knockdown cells. E) Enrichments of BrdU containing DNA in four FACS sorted fractions (S1-S4) as quantified by real-time PCR. S1 represents the earliest and S4 the latest S phase fraction as measured by DNA content. The dodeca pericentric repeat sequence is shown. Error bars represent the standard deviations between three biological repeats.

To study the possibly advanced replication timing at higher resolution, I measured replication timing by labeling HP1 knockdown cells and control cells with BrdU for one hour. Cells where then sorted into four different S phase fractions based on DNA content using FACS. Replicating DNA from each fraction was isolated by immunoprecipitation with an antibody detecting BrdU, and the abundance of replicating DNA in each S phase fraction was compared using qPCR.

Since the chromocenter represents mostly repetitive sequences within heterochromatin, I tested if the replication timing of centromeric repeats on chromosome 3 is advanced

after HP1 RNAi. Primers specific for the dodeca repeat sequence located in pericentric heterochromatin of chromosome 3 were used (Abad *et al.* 1992). I detected the highest enrichment of replicating DNA in S4, representing the latest S phase fraction (Figure 1E). However, in HP1 RNAi cells, the enrichment in late S phase is reduced, but increased in the two fractions representing mid S phase (S2 and S3, Figure 1E). This suggests that the dodeca repeat shows advanced replication timing after HP1 knockdown. This is in accordance with a global advance of centromeric heterochromatin replication timing as implicated by BrdU immunofluorescence.

Since HP1 has been shown to bind to many genes located within heterochromatin (Greil et al. 2003; de Wit et al. 2007), I tested if replication timing was altered after HP1 knockdown at heterochromatic genes. Transcription of those genes often depends on their heterochromatic environment, since translocating them into euchromatin or reducing HP1 protein levels can eliminate their expression (reviewed in (Yasuhara et al. 2006)). Two heterochromatic genes on chromosome 2 (light and rolled) did not show a difference in enrichments of BrdU'IPd DNA in the four S phase fractions between HP1 knockdown and control cells (data not shown). However, another region on chromosome 3, containing the Dbp80 and RpL15 genes showed a strong effect on replication timing after HP1 knockdown. The large Dbp80 gene replicates late in Kc cells, but showed advanced replication timing in HP1 knockdown cells (Figure 2A), while the early replicating RpL15 gene replicated later after HP1 knockdown (Figure 2B). This shows that modulating HP1 levels can have a strong effect on replication timing of some heterochromatic genes.

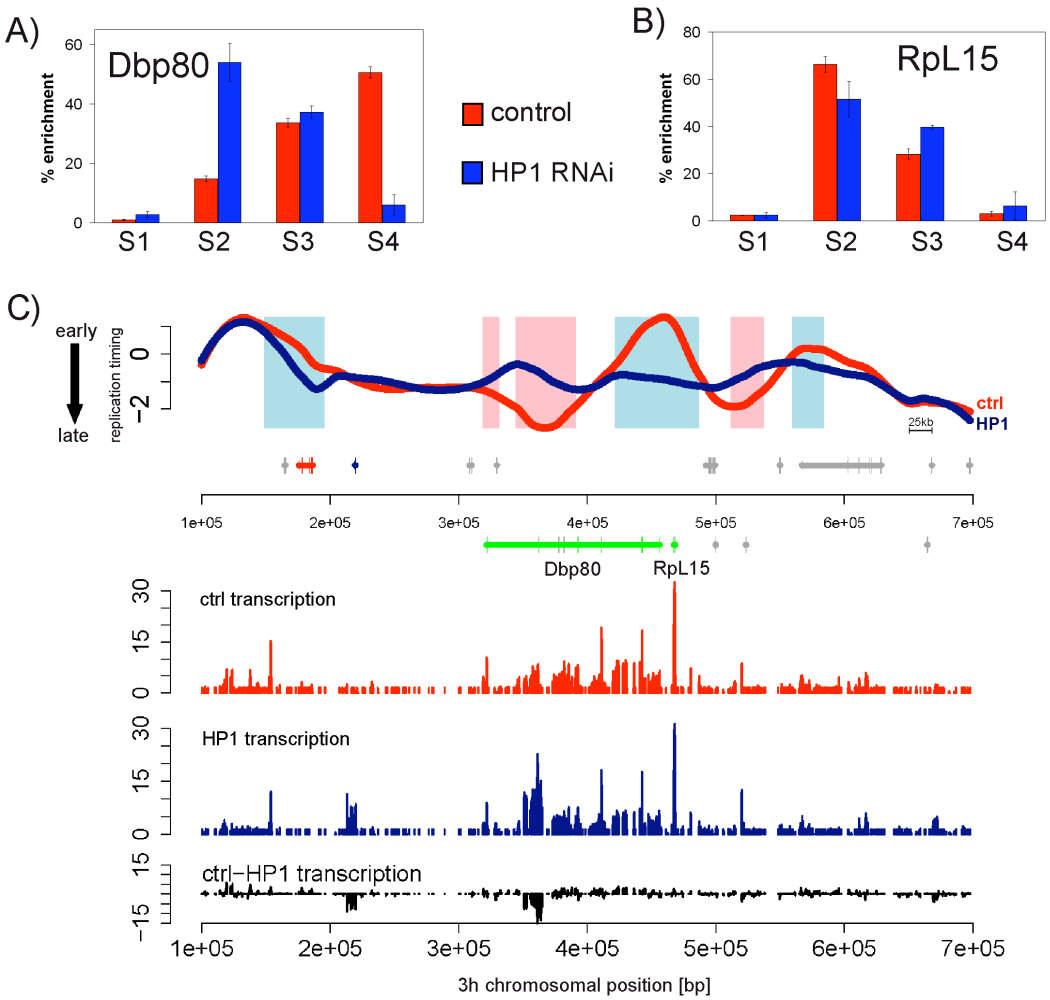



**Figure 2: Differences in replication timing at heterochromatic genes.** A) Enrichments of BrdU containing DNA in four FACS sorted fractions (S1-S4) as quantified by real-time PCR with a primers specific for the 3' region of the large Dbp80 gene. S1 represents the earliest and S4 the latest S phase fraction as measured by DNA content. Error bars represent the standard deviations between three biological repeats. B) Enrichments of BrdU containing DNA in four FACS sorted fractions (S1-S4) as quantified by real-time PCR with a primers specific for the RpL15 gene. S1 represents the earliest and S4 the latest S phase fraction as measured by DNA content. Error bars represent the standard deviations between three biological repeats. C) Replication timing profiles of control (red, ctrl) and HP1 knockdown (blue, HP1) Kc cells for a representative region on chromosome 3h. X-axis = 3h chromosomal position in basepairs, y axis = log2 (early/late replication). Background coloring denotes regions that replicate earlier in HP1 cells (L:E, pink), regions that replicate earlier in ctrl cells (E:L, blue) and regions replicating similarly in both cell types (white). Annotated genes are displayed below the profile (boxes=exons, lines=introns, small boxes=UTRs) and colored by their expression status (see methods for details, green= expressed in ctrl and HP1 cells, blue= expressed only in HP1 cells, red=expressed only in ctrl cells, grey= not expressed in ctrl and HP1 cells). Transcription levels of ctrl (red) and HP1 (blue) cells measured by tiling arrays are displayed on the same scale below, including transcription level differences (black).

To analyze replication timing changes on a genome-wide scale, I obtained early and late replicating DNA by sorting BrdU labeled cells into early and late S phase, followed by BrdU-IP, and hybridized it to whole genome tiling arrays. The resulting replication timing

profiles were highly reproducible (see chapter 3.1/Supplementary Figure 1). To determine regions of dynamic replication in an unbiased way I employed a three state Hidden Markov Model (HMM) to segment replication timing differences between control and knockdown cells (see methods). In order to focus on robust changes I excluded regional differences that are smaller than 5 kb or where the difference in timing extends over less than 12.5% of the total range (delta log2 < 0.8). These stringent criteria reveal 271 regions, corresponding to 7% of the genome, that replicate earlier in control than in HP1 knockdown cells (E:L) and 107 regions, corresponding to 3% of the genome, that replicate earlier in HP1 RNAi than in control cells (L:E). These differentially replicating regions can be larger than 100kb, have an average size of 30kb (data not shown), and represent 10% of all sequences.

Figure 2C shows a region within pericentric heterochromatin of chromosome 3 (chromosome 3h) containing the Dbp80 and RpL15 genes which are differentially replicating according to qPCR analysis. In agreement with the qPCR data I detect differences in replication timing using tiling arrays (Figure 2C). The 3' region of advanced replication timing in HP1 knockdown cells overlaps with the Dbp80 gene, while the RpL15 gene shows delayed replication timing after knockdown of HP1 (Figure 2C). To determine if these changes in replication timing correlate with changes in gene expression, I hybridized RNA from control and HP1 knockdown cells to 3'UTR as well as chromosomal tiling arrays. Both, the Dbp80 and RpL15 genes are active in Kc cells, even after depletion of HP1 (Figure 2C). Tiling array analysis shows increased transcription in part of the Dbp80 gene in HP1 knockdown cells (Figure 2C). Therefore, HP1 depletion could have a direct effect on the replication timing of this region, or transcriptional activation of parts of the Dbp80 gene could lead to earlier replication of the Dbp80 gene. Interestingly, the delay in replication timing around the RpL15 gene does not seem to coincide with reduced transcript levels (Figure 2C), suggesting that HP1 could have a more direct role in the firing of some early origins within heterochromatin.

To test if replication timing changes occur in the same regions where HP1 is bound, I compared differential replication timing to published HP1 binding data on chromosome 2 and 4 (de Wit et al. 2007) and found that regions replicating earlier in HP1 RNAi cells did not show significantly elevated HP1 binding levels (Figure 3A). Chromosome 4 does not contain regions replicating earlier in HP1 RNAi cells, therefore this suggests that the advanced replication timing of 47 regions on chromosome 2 does not take place in

regions normally bound by HP1. Interestingly, regions on chromosome 2 and 4 which replicate later after HP1 knockdown were more often bound by HP1 than regions replicating at the same time (Figure 3A), suggesting that those differences might depend directly on HP1 binding to differentially replicating regions. HP1 also shows a particularly strong enrichment on chromosome 4 ((de Wit et al. 2007) and Figure 3C). The correlation of HP1 binding with delayed replication timing was mostly due to delayed replication timing of chromosome 4 and pericentric regions on chromosome 2 (Figure 3B). We detected a global delay of replicating timing of the early replicating chromosome 4 in HP1 depleted cells (Figure 3B and 3D). This suggests that high levels of HP1 present on chromosome 4 might be important for maintaining its early replication. We also observed a weak reduction in transcription levels throughout the chromosome (Figure 3C). Interestingly however, wile many genes showed a very small decrease in transcript levels, only very few genes showed a complete loss of gene expression. Furthermore, not all differentially replicating regions also showed a reduction in transcription (Figure 3D). This shows that depletion of HP1 leads to a global delay of replication timing and reduction of transcription on chromosome 4. These data suggest that the effect of HP1 on the timing of chromosome 4 replication is independent from its effect on transcription.

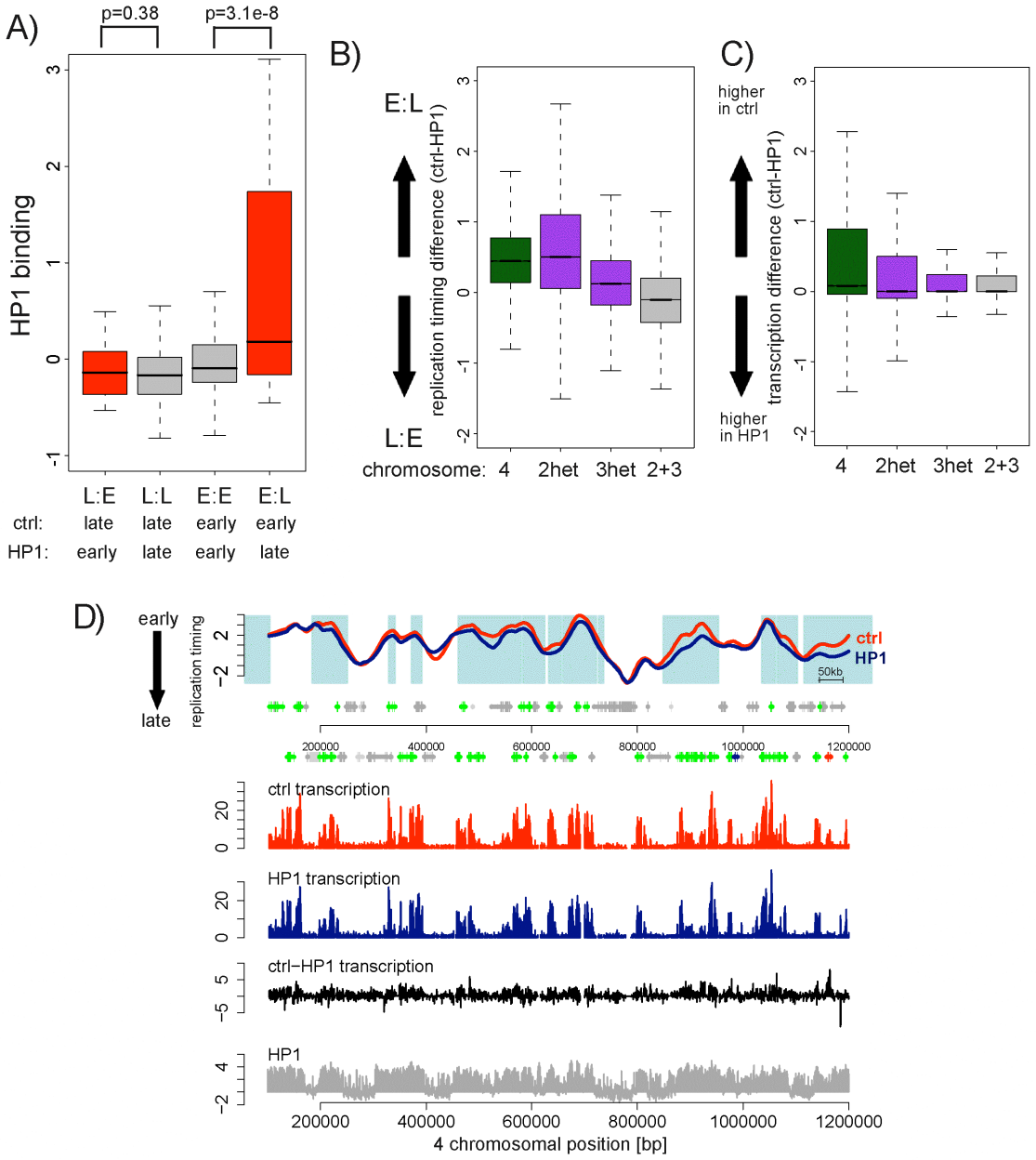



**Figure 3: Delayed replication timing of chromosome 4 after knockdown of HP1.**
A) Distribution of HP1 binding levels (de Wit et al. 2007) for regions with differential replication timing. The boxplots illustrate that regions replication later after HP1 knockdown (E:L) show high levels of HP1 binding. L:E=regions replicating earlier in HP1 cells, L:L=regions replicating late in both, E:E=regions replicating early in both, E:L=regions replicating earlier in ctrl cells. p-values were calculated using the Wilcoxon rank sum test.
B) Distribution of replication timing differences (control-HP1 RNAi replication timing, log2 scale) within different chromosomal regions. The boxplots illustrate that chromosome 4 (4, green boxplot) and pericentric regions on chromosome 2 (2het, purple boxplot) and to a lesser extend on chromosome 3 (3het, purple boxplot) show delayed replication timing (E:L) more often than euchromatin on chromosome 2 and 3 (2+3, grey boxplot). C) Distribution of transcription differences (control-HP1 RNAi transcription, log2 scale) within different chromosomal regions. The boxplots illustrate that chromosome 4 (4, green boxplot) and pericentric regions on chromosome 2 (2het, purple boxplot) but not on chromosome 3 (3het, purple boxplot) show reduced transcription levels slightly more often than euchromatin on chromosome 2 and 3

(2+3, grey boxplot). D) Replication timing profiles of control (red, ctrl) and HP1 knockdown (blue, HP1) Kc cells for a representative region on chromosome 4. X-axis = 4 chromosomal position in basepairs, y axis = log2 (early/late replication). Background coloring denotes regions that replicate earlier in HP1 cells (L:E, pink), regions that replicate earlier in ctrl cells (E:L, blue) and regions replicating similarly in both cell types (white). Annotated genes are displayed below the profile (boxes=exons, lines=introns, small boxes=UTRs) and colored by their expression status (see methods for details, green= expressed in ctrl and HP1 cells, blue= expressed only in HP1 cells, red=expressed only in ctrl cells, grey= not expressed in ctrl and HP1 cells). Transcription levels of ctrl (red) and HP1 (blue) cells measured by tiling arrays are displayed on the same scale below, including transcription level differences (black).

Several regions on chromosome 2 and 3 are differentially replicating after depletion of HP1. Most of these regions lie within pericentric heterochromatin (Figure3B), which also shows high levels of HP1 binding (de Wit et al. 2007). Figure 4A displays a region spanning part of the pericentric heterochromatin and part of euchromatin on chromosome 2R. The region close to the centromer shows high levels of HP1 binding throughout (Figure 4A) and many regions with strong replication timing differences, mostly replicating later in HP1 RNAi cells than in control cells (Figure 4A, left). The overall transcription differences within chromosome 2 heterochromatin tend towards reduced expression after HP1 knockdown (Figure 3C). However, some of those differentially replicating regions do not contain any differences in transcription levels (Figure 4A).

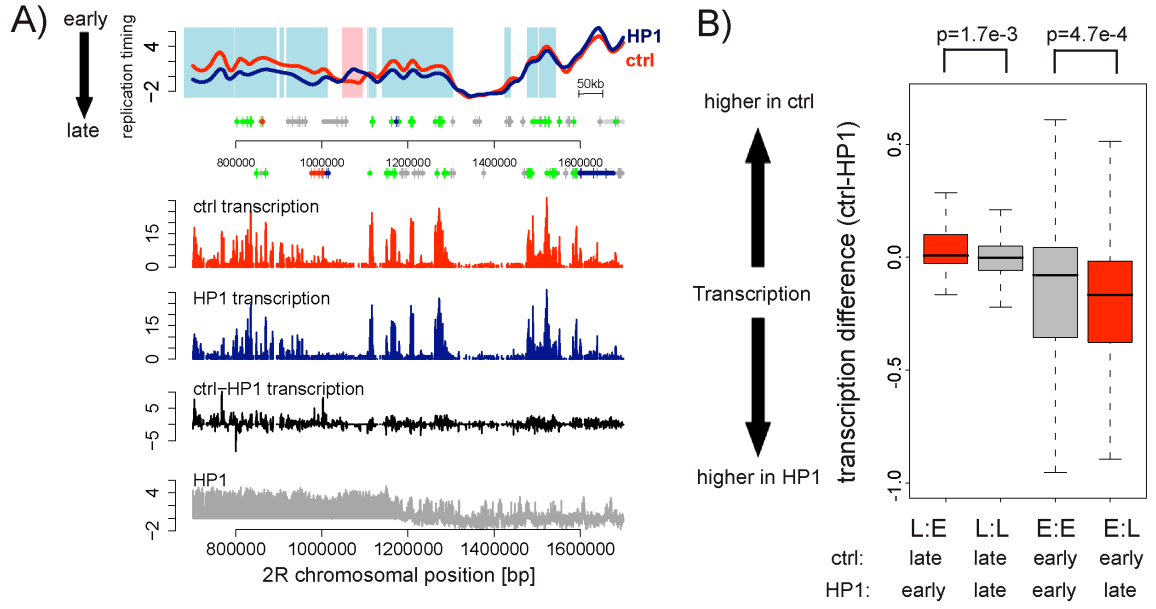



**Figure 4: Differences in replication timing do not always coincide with transcription differences.** A) Replication timing profiles of control (red, ctrl) and HP1 knockdown (blue, HP1) Kc cells for a representative region on chromosome 3h. X-axis = 3h chromosomal position in basepairs, y axis = log2 (early/late replication). Background coloring denotes regions that replicate earlier in HP1 cells (L:E, pink), regions that replicate earlier in ctrl cells (E:L, blue) and regions replicating similarly in both cell types (white). Annotated genes are displayed below the profile (boxes=exons, lines=introns, small boxes=UTRs)

and colored by their expression status (see methods for details, green= expressed in ctrl and HP1 cells, blue= expressed only in HP1 cells, red=expressed only in ctrl cells, grey= not expressed in ctrl and HP1 cells). Transcription levels of ctrl (red) and HP1 (blue) cells measured by tiling arrays are displayed on the same scale below, including transcription level differences (black). B) Distribution of transcription differences (ctrl-HP1 transcription levels) for regions with differential replication timing on autosomes. The boxplots illustrate that on average differences in replication timing coincide with changes in transcription. L:E=regions replicating earlier in HP1 cells, L:L=regions replicating late in both, E:E=regions replicating early in both, E:L=regions replicating earlier in ctrl cells. p-values were calculated using the Wilcoxon rank sum test.

A genome-wide comparison of average transcription differences within differentially replicating regions to replication timing (Figure 4B) reveals that regions that replicate earlier in HP1 knockdown cells (L:E) are also transcribed at a higher level there. Similarly, regions that replicate earlier in control cells (E:L) show increased transcription in control cells (Figure 4B). Notably, in many cases, these transcriptional changes only occur at a low percentage of genes in each differentially replicating region (data not shown). At the same time only approximately 20% of differentially replicating regions show transcription differences, which are significantly higher than in regions replicating at the same time (Figure 4B). This suggests that transcription is not the sole determining force, but HP1 can affect replication timing and transcription independently of each other on several regions within euchromatin and heterochromatin. However, I cannot exclude the possibility of differences in the transcription of repeats, which are not represented on the tiling array.

In summary, I show that depletion of HP1 from *Drosophila* tissue culture cells results in changes in replication timing, most of which lie within heterochromatin and on the 4$^{th}$ chromosome. While many replication timing changes correlate with transcriptional differences, some regions of the genome seem to change replication timing independent of transcription. Taken together, this provides evidence for a role of HP1 in modulating replication timing.

### 3.3.4.Discussion

HP1 is important for heterochromatic silencing. It binds the repressive H3K9me2 histone modification, and loss of HP1 results in suppression of PEV, while an increase in HP1 protein can enhance PEV in *Drosophila*. On the other hand, HP1 is required for expression of some heterochromatic genes (reviewed in (Hediger et al. 2006)). I demonstrate that a reduction of HP1 levels in *Drosophila* Kc cells results in an advance of replication timing of late replicating heterochromatic centromeric repeats (Figure 1). Interaction of HP1 with CAF-1 is required for replication of heterochromatin in mouse

cells, suggesting that compact chromatin mediated by HP1 needs to be alleviated to allow replication fork progression through S phase (Quivy *et al.* 2008). Depletion of HP1 could result in the loss of the requirement for this mechanism and therefore allow slightly earlier replication of centromeric repetitive heterochromatin. This advanced replication timing of repeats is consistent with observations in mammalian cells, where replication timing of pericentric repeats is advanced in cells depleted of SuVAR 3-9 (Wu et al. 2006), the histone methylase which sets the H3K9me2 mark. Furthermore, studies of endoreplication of *Drosophila* polytene chromosomes have suggested a role for HP1 in the maintenance of under-replicated regions in connection with the SuUR (Suppressor of Underreplication) protein (Makunin et al. 2002; Koryakov et al. 2006; Pindyurin et al. 2007; Pindyurin et al. 2008). Heterochromatin has to be propagated through cell division to maintain its epigenetic state. It has been suggested that next to histone modifications, heterochromatin binding proteins like HP1 play a crucial role in this process. Loss of HP1 might also affect replication timing by reducing the efficiency of heterochromatin maintenance through the cell cycle.

I studied the replication timing of the non-repetitive part of the genome in HP1 RNAi cells at high resolution using tiling arrays. Cells with reduced HP1 protein levels display differences in replication timing. These differences mostly occur in HP1-bound heterochromatin and the 4$^{th}$ chromosome, suggesting that replication timing changes are directly dependent on HP1. The lower number of replication timing differences within euchromatin might not be directly linked to HP1, but are more likely to be caused by downstream gene expression changes. In contrast to centromeric repeats, pericentric heterochromatin is delayed in replication timing after HP1 knockdown. This is surprising, since heterochromatin generally replicates late during S phase, and HP1 is thought to maintain a stable heterochromatic environment, which might inhibit DNA replication. HP1 is highly enriched on chromosome 4, where I observe a global replication delay. Chromosome 4 is entirely heterochromatic, however, in the distal 1.2 mega-basepairs the gene density is typical of euchromatin (reviewed in (Riddle *et al.* 2006)). In Kc cells, most genes within this distal region are active, and it replicates in early S phase (Figure 3D).

It is possible, that through its interaction with ORC2 (Pak et al. 1997), HP1 supports origin firing on chromosome 4 and possibly other regions of the genome with elevated HP1 levels, including heterochromatic genes. A reduction of HP1 levels in the cell would then result in delayed or less efficient origin firing. This might then lead to the observed

replication timing delay in HP1 RNAi cells. However it is also conceivable that some of the observed delays in replication timing after HP1 knockdown are a result of down-regulated transcription and concomitant changes in histone modifications. While many regions of delayed replication timing indeed also show slightly down-regulated transcription, many do not show any changes in gene expression levels (Figure 4B). This suggests that HP1 might support not only the transcription, but also the replication of some heterochromatic genes.

Many different models have been put forward for why pericentric genes depend on their heterochromatic environment for expression (Yasuhara et al. 2006). Another model is that repetitive sequences within pericentric heterochromatin are transcriptionally up-regulated after loss of HP1. Since I cannot measure the transcription of repeats using tiling arrays, it would not be feasible for me to detect such a global up-regulation of repeat transcription. Yet, such an effect on repetitive elements could be a possible reason for the reduction of transcription of heterochromatic genes after loss of HP1 or translocation to euchromatin. In this model, transcription of repetitive elements would interfere with the proper expression of genes. The role of HP1 would then be to silence these repeats, while genes can remain active. Similarly, such an increase in transcriptional activity within repeats could lead to a perturbation of origin firing within pericentric heterochromatin. Indeed, transcription can have a negative influence on origin selection in mammalian cells (Sasaki *et al.* 2006). It has been suggested that eukaryotic replication origins tend to be located in intergenic, non-transcribed regions, yet close to actively transcribed genes (MacAlpine et al. 2005). This would be consistent with a model where aberrant transcription of intergenic, repetitive regions interferes with origin selection as well as gene expression. While the zones of replication initiation seem to be similar in control and HP1 knockdown cells (Figure 4A), only a detailed characterization of the location of origins of replication would allow to distinguish between a delay in origin firing and the use of different origins after HP1 knockdown.

In summary, I report a dual role of HP1 in the control of replication timing of repetitive and unique heterochromatin sequences in *Drosophila* cells. Centromeric repeats replicate earlier after HP1 knockdown. Surprisingly, pericentric heterochromatin and the 4$^{th}$ chromosome display a replication delay. The exact mechanism of the effect of HP1 RNAi on replication timing remains to be determined. Yet, the results presented here suggest that proper chromatin structure is important for the regulation of DNA replication.

### 3.3.5. Materials and Methods

**RNA interference (RNAi ) in Kc cells.**

Double stranded RNA (dsRNA) was prepared from a PCR product spanning the entire HP1 coding region, generated with primers containing a T7 RNA Polymerase binding site using the MEGASCRIPT T7 *In Vitro* Transcription Kit (Ambion, Cat.No. AM1334). The RNA was purified and heated to 70°C for 10 minutes and slowly cooled down to room temperature for about 30 minutes to enhance annealing. 50µg of dsRNA was added to $10^6$ cells every 2nd day and 8 days after initial addition of dsRNA cells were harvested and the efficiency of HP1 reduction was estimated by western blot analysis using a monoclonal α-HP1 antibody.

**Immunofluorescence analysis.**

Immunofluorescence staining was carried out as described (Wirbelauer et al. 2005), using a polyclonal rabbit α-HP1 antibody. For replication timing analysis by immunofluorescence, cells were labeled with BrdU for 1 hour, and BrdU was detected using the 5'-Bromo-2'-deoxy-uridine Labeling and Detection kit I (Roche, Cat.No.11296736001). DNA was counterstained with 0.04 ug/ml DAPI. Stainings were analyzed using a laser scanning confocal microscope LSM510 META (Zeiss) and LSM510 software.

**Analysis of replication timing and transcription.**

Replication timing and transcription measurements were performed as described in chapter 3.1. Real-time PCR analysis was performed in triplicates and array experiments were done in two biological replicates. The definition of differentially replicating regions based on replication timing profiles was done using Hidden Markov Models with adding cutoff where a replication timing difference has to span more than 1/8th of the distribution of control-HP1 replication timing difference data points. This analysis was also performed on replication timing differences between two control samples. Thereby I found that 0.5% of the genome showed differences in replication timing, compared to 10% between control and HP1 knockdown cells. This suggests that most of the replication timing changes we detect after HP1 RNAi are significant.

# 4. Bibliography


Abad, J.P., Carmena, M., Baars, S., Saunders, R.D., Glover, D.M., Ludena, P., Sentis, C., Tyler-Smith, C., and Villasante, A. 1992. Dodeca satellite: a conserved G+C-rich satellite from the centromeric heterochromatin of Drosophila melanogaster. *Proc Natl Acad Sci U S A* **89**: 4663-4667.

Aggarwal, B.D. and Calvi, B.R. 2004. Chromatin regulates origin activity in Drosophila follicle cells. *Nature* **430**: 372-376.

Ahmad, K. and Henikoff, S. 2002. The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. *Mol Cell* **9**: 1191-1200.

Akhtar, A. and Gasser, S.M. 2007. The nuclear envelope and transcriptional control. *Nat Rev Genet* **8**: 507-517.

Aladjem, M.I. 2007. Replication in context: dynamic regulation of DNA replication patterns in metazoans. *Nat Rev Genet* **8**: 588-600.

Aladjem, M.I., Rodewald, L.W., Kolman, J.L., and Wahl, G.M. 1998. Genetic dissection of a mammalian replicator in the human beta-globin locus. *Science* **281**: 1005-1009.

Albert, I., Mavrich, T.N., Tomsho, L.P., Qi, J., Zanton, S.J., Schuster, S.C., and Pugh, B.F. 2007. Translational and rotational settings of H2A.Z nucleosomes across the Saccharomyces cerevisiae genome. *Nature* **446**: 572-576.

Alekseyenko, A.A., Larschan, E., Lai, W.R., Park, P.J., and Kuroda, M.I. 2006. High-resolution ChIP-chip analysis reveals that the Drosophila MSL complex selectively identifies active genes on the male X chromosome. *Genes Dev* **20**: 848-857.

Altman, A.L. and Fanning, E. 2001. The Chinese hamster dihydrofolate reductase replication origin beta is active at multiple ectopic chromosomal locations and requires specific DNA sequence elements for activity. *Mol Cell Biol* **21**: 1098-1110.

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

Aparicio, J.G., Viggiani, C.J., Gibson, D.G., and Aparicio, O.M. 2004. The Rpd3-Sin3 histone deacetylase regulates replication timing and enables intra-S origin control in Saccharomyces cerevisiae. *Mol Cell Biol* **24**: 4769-4780.

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

Audit, B., Nicolay, S., Huvet, M., Touchon, M., d'Aubenton-Carafa, Y., Thermes, C., and Arneodo, A. 2007. DNA replication timing data corroborate in silico human replication origin predictions. *Phys Rev Lett* **99**: 248102.

Azuara, V., Brown, K.E., Williams, R.R., Webb, N., Dillon, N., Festenstein, R., Buckle, V., Merkenschlager, M., and Fisher, A.G. 2003. Heritable gene silencing in lymphocytes delays chromatid resolution without affecting the timing of DNA replication. *Nat Cell Biol* **5**: 668-674.

Azuara, V., Perry, P., Sauer, S., Spivakov, M., Jorgensen, H.F., John, R.M., Gouti, M., Casanova, M., Warnes, G., Merkenschlager, M. et al. 2006. Chromatin signatures of pluripotent cell lines. *Nat Cell Biol* **8**: 532-538.

Balasov, M., Huijbregts, R.P., and Chesnokov, I. 2007. Role of the Orc6 protein in origin recognition complex-dependent DNA binding and replication in Drosophila melanogaster. *Mol Cell Biol* **27**: 3143-3153.

Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. 2007. High-resolution profiling of histone methylations in the human genome. *Cell* **129**: 823-837.

Beall, E.L., Manak, J.R., Zhou, S., Bell, M., Lipsick, J.S., and Botchan, M.R. 2002. Role for a Drosophila Myb-containing protein complex in site-specific DNA replication. *Nature* **420**: 833-837.

Beissbarth, T. and Speed, T.P. 2004. GOstat: find statistically overrepresented Gene Ontologies within a group of genes. *Bioinformatics* **20**: 1464-1465.

Bell, O., Conrad, T., Kind, J., Wirbelauer, C., Akhtar, A., and Schubeler, D. 2008. Transcription-coupled methylation of histone H3 at lysine 36 regulates dosage compensation by enhancing recruitment of the MSL complex in Drosophila melanogaster. *Mol Cell Biol* **28**: 3401-3409.

Bell, O., Wirbelauer, C., Hild, M., Scharf, A.N., Schwaiger, M., MacAlpine, D.M., Zilbermann, F., van Leeuwen, F., Bell, S.P., Imhof, A. et al. 2007. Localized H3K36 methylation states define histone H4K16 acetylation during transcriptional elongation in Drosophila. *Embo J* **26**: 4974-4984.

Bell, S.P. and Dutta, A. 2002. DNA replication in eukaryotic cells. *Annu Rev Biochem* **71**: 333-374.

Bell, S.P., Kobayashi, R., and Stillman, B. 1993. Yeast origin recognition complex functions in transcription silencing and DNA replication. *Science* **262**: 1844-1849.

Bell, S.P. and Stillman, B. 1992. ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. *Nature* **357**: 128-134.

Belyaeva, E.S., Zhimulev, I.F., Volkova, E.I., Alekseyenko, A.A., Moshkin, Y.M., and Koryakov, D.E. 1998. Su(UR)ES: a gene suppressing DNA underreplication in intercalary and pericentric heterochromatin of Drosophila melanogaster polytene chromosomes. *Proc Natl Acad Sci U S A* **95**: 7532-7537.

Belyakin, S.N., Christophides, G.K., Alekseyenko, A.A., Kriventseva, E.V., Belyaeva, E.S., Nanayev, R.A., Makunin, I.V., Kafatos, F.C., and Zhimulev, I.F. 2005. Genomic analysis of Drosophila chromosome underreplication reveals a link between replication control and transcriptional territories. *Proc Natl Acad Sci U S A* **102**: 8269-8274.

Birney, E. Stamatoyannopoulos, J.A. Dutta, A. Guigo, R. Gingeras, T.R. Margulies, E.H. Weng, Z. Snyder, M. Dermitzakis, E.T. Thurman, R.E. et al. 2007. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* **447**: 799-816.

Bosco, G., Du, W., and Orr-Weaver, T.L. 2001. DNA replication control through interaction of E2F-RB and the origin recognition complex. *Nat Cell Biol* **3**: 289-295.

Boutanaev, A.M., Kalmykova, A.I., Shevelyov, Y.Y., and Nurminsky, D.I. 2002. Large clusters of co-expressed genes in the Drosophila genome. *Nature* **420**: 666-669.

Cairns, B.R. 2007. Chromatin remodeling: insights and intrigue from single-molecule studies. *Nat Struct Mol Biol* **14**: 989-996.

Calvi, B.R., Byrnes, B.A., and Kolpakas, A.J. 2007. Conservation of Epigenetic Regulation, ORC Binding and Developmental Timing of DNA Replication Origins in the Genus Drosophila. *Genetics* **177**: 1291-1301.

Carey, M., Li, B., and Workman, J.L. 2006. RSC exploits histone acetylation to abrogate the nucleosomal block to RNA polymerase II elongation. *Mol Cell* **24**: 481-487.

Carrozza, M.J., Li, B., Florens, L., Suganuma, T., Swanson, S.K., Lee, K.K., Shia, W.J., Anderson, S., Yates, J., Washburn, M.P. et al. 2005. Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* **123**: 581-592.

Cayirlioglu, P., Bonnette, P.C., Dickson, M.R., and Duronio, R.J. 2001. Drosophila E2f2 promotes the conversion from genomic DNA replication to gene amplification in ovarian follicle cells. *Development* **128**: 5085-5098.

Cayirlioglu, P., Ward, W.O., Silver Key, S.C., and Duronio, R.J. 2003. Transcriptional repressor functions of Drosophila E2F1 and E2F2 cooperate to inhibit genomic DNA synthesis in ovarian follicle cells. *Mol Cell Biol* **23**: 2123-2134.

Chakalova, L., Debrand, E., Mitchell, J.A., Osborne, C.S., and Fraser, P. 2005. Replication and transcription: shaping the landscape of the genome. *Nat Rev Genet* **6**: 669-677.

Chimpanzee_Sequencing_and_Analysis_Consortium. 2005. Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* **437**: 69-87.

Cimbora, D.M., Schubeler, D., Reik, A., Hamilton, J., Francastel, C., Epner, E.M., and Groudine, M. 2000. Long-distance control of origin choice and replication timing in the human beta-globin locus are independent of the locus control region. *Mol Cell Biol* **20**: 5581-5591.

Clegg, N.J., Honda, B.M., Whitehead, I.P., Grigliatti, T.A., Wakimoto, B., Brock, H.W., Lloyd, V.K., and Sinclair, D.A. 1998. Suppressors of position-effect variegation in Drosophila melanogaster affect expression of the heterochromatic gene light in the absence of a chromosome rearrangement. *Genome* **41**: 495-503.

Conti, C., Sacca, B., Herrick, J., Lalou, C., Pommier, Y., and Bensimon, A. 2007. Replication fork velocities at adjacent replication origins are coordinately modified during DNA replication in human cells. *Mol Biol Cell* **18**: 3059-3067.

Corona, D.F., Clapier, C.R., Becker, P.B., and Tamkun, J.W. 2002. Modulation of ISWI function by site-specific histone acetylation. *EMBO Rep* **3**: 242-247.

Cosgrove, A.J., Nieduszynski, C.A., and Donaldson, A.D. 2002. Ku complex controls the replication time of DNA in telomere regions. *Genes Dev* **16**: 2485-2490.

Cote, J., Quinn, J., Workman, J.L., and Peterson, C.L. 1994. Stimulation of GAL4 derivative binding to nucleosomal DNA by the yeast SWI/SNF complex. *Science* **265**: 53-60.

Danis, E., Brodolin, K., Menut, S., Maiorano, D., Girard-Reydet, C., and Mechali, M. 2004. Specification of a DNA replication origin by a transcription complex. *Nat Cell Biol* **6**: 721-730.

Dazy, S., Gandrillon, O., Hyrien, O., and Prioleau, M.N. 2006. Broadening of DNA replication origin usage during metazoan cell differentiation. *EMBO Rep* **7**: 806-811.

de Wit, E., Greil, F., and van Steensel, B. 2005. Genome-wide HP1 binding in Drosophila: developmental plasticity and genomic targeting signals. *Genome Res* **15**: 1265-1273.

-. 2007. High-resolution mapping reveals links of HP1 with active and inactive chromatin components. *PLoS Genet* **3**: e38.

Deuring, R., Fanti, L., Armstrong, J.A., Sarte, M., Papoulas, O., Prestel, M., Daubresse, G., Verardo, M., Moseley, S.L., Berloco, M. et al. 2000. The ISWI chromatin-remodeling protein is required for gene expression and the maintenance of higher order chromatin structure in vivo. *Mol Cell* **5**: 355-365.

Devlin, R.H., Bingham, B., and Wakimoto, B.T. 1990. The organization and expression of the light gene, a heterochromatic gene of Drosophila melanogaster. *Genetics* **125**: 129-140.

Dijkwel, P.A., Mesner, L.D., Levenson, V.V., d'Anna, J., and Hamlin, J.L. 2000. Dispersive initiation of replication in the Chinese hamster rhodopsin locus. *Exp Cell Res* **256**: 150-157.

Dillon, N. and Festenstein, R. 2002. Unravelling heterochromatin: competition between positive and negative factors regulates accessibility. *Trends Genet* **18**: 252-258.

Dion, M.F., Altschuler, S.J., Wu, L.F., and Rando, O.J. 2005. Genomic characterization reveals a simple histone H4 acetylation code. *Proc Natl Acad Sci U S A* **102**: 5501-5506.

Dorigo, B., Schalch, T., Bystricky, K., and Richmond, T.J. 2003. Chromatin fiber folding: requirement for the histone H4 N-terminal tail. *J Mol Biol* **327**: 85-96.

Dutta, A. and Bell, S.P. 1997. Initiation of DNA replication in eukaryotic cells. *Annu Rev Cell Dev Biol* **13**: 293-332.

Echalier, G. 1997. *Drosophila cells in culture*. Academic Press, New York, NY.

Eissenberg, J.C., James, T.C., Foster-Hartnett, D.M., Hartnett, T., Ngan, V., and Elgin, S.C. 1990. Mutation in a heterochromatin-specific chromosomal protein is associated with suppression of position-effect variegation in Drosophila melanogaster. *Proc Natl Acad Sci U S A* **87**: 9923-9927.

Ellis, D.J., Jenkins, H., Whitfield, W.G., and Hutchison, C.J. 1997. GST-lamin fusion proteins act as dominant negative mutants in Xenopus egg extract and reveal the function of the lamina in DNA replication. *J Cell Sci* **110 ( Pt 20)**: 2507-2518.

Felsenfeld, G. and Groudine, M. 2003. Controlling the double helix. *Nature* **421**: 448-453.

Fischle, W., Tseng, B.S., Dormann, H.L., Ueberheide, B.M., Garcia, B.A., Shabanowitz, J., Hunt, D.F., Funabiki, H., and Allis, C.D. 2005. Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. *Nature* **438**: 1116-1122.

Flaus, A. and Owen-Hughes, T. 2004. Mechanisms for ATP-dependent chromatin remodelling: farewell to the tuna-can octamer? *Curr Opin Genet Dev* **14**: 165-173.

Georlette, D., Ahn, S., MacAlpine, D.M., Cheung, E., Lewis, P.W., Beall, E.L., Bell, S.P., Speed, T., Manak, J.R., and Botchan, M.R. 2007. Genomic profiling and expression studies reveal both positive and negative activities for the Drosophila Myb MuvB/dREAM complex in proliferating cells. *Genes Dev* **21**: 2880-2896.

Gierman, H.J., Indemans, M.H., Koster, J., Goetze, S., Seppen, J., Geerts, D., van Driel, R., and Versteeg, R. 2007. Domain-wide regulation of gene expression in the human genome. *Genome Res* **17**: 1286-1295.

Gilbert, D.M. 2002. Replication timing and transcriptional control: beyond cause and effect. *Curr Opin Cell Biol* **14**: 377-383.

-. 2004. In search of the holy replicator. *Nat Rev Mol Cell Biol* **5**: 848-855.

Gilbert, D.M. and Cohen, S.N. 1987. Bovine papilloma virus plasmids replicate randomly in mouse fibroblasts throughout S phase of the cell cycle. *Cell* **50**: 59-68.

Gilbert, N., Boyle, S., Fiegler, H., Woodfine, K., Carter, N.P., and Bickmore, W.A. 2004. Chromatin architecture of the human genome: gene-rich domains are enriched in open chromatin fibers. *Cell* **118**: 555-566.

Gilfillan, G.D., Straub, T., de Wit, E., Greil, F., Lamm, R., van Steensel, B., and Becker, P.B. 2006. Chromosome-wide gene-specific targeting of the Drosophila dosage compensation complex. *Genes Dev* **20**: 858-870.

Gladden, A.B. and Diehl, J.A. 2003. The cyclin D1-dependent kinase associates with the pre-replication complex and modulates RB.MCM7 binding. *J Biol Chem* **278**: 9754-9760.

Goren, A., Tabib, A., Hecht, M., and Cedar, H. 2008. DNA replication timing of the human {beta}-globin domain is controlled by histone modification at the origin. *Genes Dev* **22**: 1319-1324.

Gray, S.J., Gerhardt, J., Doerfler, W., Small, L.E., and Fanning, E. 2007. An origin of DNA replication in the promoter region of the human fragile X mental retardation (FMR1) gene. *Mol Cell Biol* **27**: 426-437.

Gregoire, D., Brodolin, K., and Mechali, M. 2006. HoxB domain induction silences DNA replication origins in the locus and specifies a single origin at its boundary. *EMBO Rep* **7**: 812-816.

Greil, F., de Wit, E., Bussemaker, H.J., and van Steensel, B. 2007. HP1 controls genomic targeting of four novel heterochromatin proteins in Drosophila. *Embo J* **26**: 741-751.

Greil, F., van der Kraan, I., Delrow, J., Smothers, J.F., de Wit, E., Bussemaker, H.J., van Driel, R., Henikoff, S., and van Steensel, B. 2003. Distinct HP1 and Su(var)3-9 complexes bind to sets of developmentally coexpressed genes depending on chromosomal location. *Genes Dev* **17**: 2825-2838.

Gribnau, J., Hochedlinger, K., Hata, K., Li, E., and Jaenisch, R. 2003. Asynchronous replication timing of imprinted loci is independent of DNA methylation, but consistent with differential subnuclear localization. *Genes Dev* **17**: 759-773.

Groth, A., Corpet, A., Cook, A.J., Roche, D., Bartek, J., Lukas, J., and Almouzni, G. 2007a. Regulation of replication fork progression through histone supply and demand. *Science* **318**: 1928-1931.

Groth, A., Rocha, W., Verreault, A., and Almouzni, G. 2007b. Chromatin challenges during DNA replication and repair. *Cell* **128**: 721-733.

Gu, W., Szauter, P., and Lucchesi, J.C. 1998. Targeting of MOF, a putative histone acetyl transferase, to the X chromosome of Drosophila melanogaster. *Dev Genet* **22**: 56-64.

Guelen, L., Pagie, L., Brasset, E., Meuleman, W., Faza, M.B., Talhout, W., Eussen, B.H., de Klein, A., Wessels, L., de Laat, W. et al. 2008. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature*.

Hansen, R.S., Canfield, T.K., Lamb, M.M., Gartler, S.M., and Laird, C.D. 1993. Association of fragile X syndrome with delayed replication of the FMR1 gene. *Cell* **73**: 1403-1409.

Hartl, T., Boswell, C., Orr-Weaver, T.L., and Bosco, G. 2007. Developmentally regulated histone modifications in Drosophila follicle cells: initiation of gene amplification is associated with histone H3 and H4 hyperacetylation and H1 phosphorylation. *Chromosoma* **116**: 197-214.

Hassan, A.B., Errington, R.J., White, N.S., Jackson, D.A., and Cook, P.R. 1994. Replication and transcription sites are colocalized in human cells. *J Cell Sci* **107 ( Pt 2)**: 425-434.

Hediger, F. and Gasser, S.M. 2006. Heterochromatin protein 1: don't judge the book by its cover! *Curr Opin Genet Dev* **16**: 143-150.

Heitz, E. 1928. Das heterochromatin der moose. *I Jahrb Wiss Botanik* 762–818.

Hellmann, I., Prufer, K., Ji, H., Zody, M.C., Paabo, S., and Ptak, S.E. 2005. Why do human diversity levels vary at a megabase scale? *Genome Res* **15**: 1222-1231.

Henikoff, S. 2008. Nucleosome destabilization in the epigenetic regulation of gene expression. *Nat Rev Genet* **9**: 15-26.

Heun, P., Laroche, T., Raghuraman, M.K., and Gasser, S.M. 2001. The positioning and dynamics of origins of replication in the budding yeast nucleus. *J Cell Biol* **152**: 385-400.

Hiratani, I., Leskovar, A., and Gilbert, D.M. 2004. Differentiation-induced replication-timing changes are restricted to AT-rich/long interspersed nuclear element (LINE)-rich isochores. *Proc Natl Acad Sci U S A* **101**: 16861-16866.

Hirota, T., Lipp, J.J., Toh, B.H., and Peters, J.M. 2005. Histone H3 serine 10 phosphorylation by Aurora B causes HP1 dissociation from heterochromatin. *Nature* **438**: 1176-1180.

Hou, Z., Bernstein, D.A., Fox, C.A., and Keck, J.L. 2005. Structural basis of the Sir1-origin recognition complex interaction in transcriptional silencing. *Proc Natl Acad Sci U S A* **102**: 8489-8494.

Hozak, P., Hassan, A.B., Jackson, D.A., and Cook, P.R. 1993. Visualization of replication factories attached to nucleoskeleton. *Cell* **73**: 361-373.

Hsu, H.C., Stillman, B., and Xu, R.M. 2005. Structural basis for origin recognition complex 1 protein-silence information regulator 1 protein interaction in epigenetic silencing. *Proc Natl Acad Sci U S A* **102**: 8519-8524.

Huvet, M., Nicolay, S., Touchon, M., Audit, B., d'Aubenton-Carafa, Y., Arneodo, A., and Thermes, C. 2007. Human gene organization driven by the coordination of replication and transcription. *Genome Res* **17**: 1278-1285.

Ito, T., Bulger, M., Pazin, M.J., Kobayashi, R., and Kadonaga, J.T. 1997. ACF, an ISWI-containing and ATP-utilizing chromatin assembly and remodeling factor. *Cell* **90**: 145-155.

Izban, M.G. and Luse, D.S. 1991. Transcription on nucleosomal templates by RNA polymerase II in vitro: inhibition of elongation with enhancement of sequence-specific pausing. *Genes Dev* **5**: 683-696.

Jacob, F. and Brenner, S. 1963. [On the regulation of DNA synthesis in bacteria: the hypothesis of the replicon.]. *C R Hebd Seances Acad Sci* **256**: 298-300.

Jenuwein, T. and Allis, C.D. 2001. Translating the histone code. *Science* **293**: 1074-1080.

Jeon, Y., Bekiranov, S., Karnani, N., Kapranov, P., Ghosh, S., MacAlpine, D., Lee, C., Hwang, D.S., Gingeras, T.R., and Dutta, A. 2005. Temporal profile of replication of human chromosomes. *Proc Natl Acad Sci U S A* **102**: 6419-6424.

Johnson, W.E., Li, W., Meyer, C.A., Gottardo, R., Carroll, J.S., Brown, M., and Liu, X.S. 2006. Model-based analysis of tiling-arrays for ChIP-chip. *Proc Natl Acad Sci U S A* **103**: 12457-12462.

Joshi, A.A. and Struhl, K. 2005. Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to Pol II elongation. *Mol Cell* **20**: 971-978.

Kim, S.M., Dubey, D.D., and Huberman, J.A. 2003. Early-replicating heterochromatin. *Genes Dev* **17**: 330-335.

Kind, J., Vaquerizas, J.M., Gebhardt, P., Gentzel, M., Luscombe, N.M., Bertone, P., and Akhtar, A. 2008. Genome-wide analysis reveals MOF as a key regulator of dosage compensation and gene expression in Drosophila. *Cell* **133**: 813-828.

Kloc, A., Zaratiegui, M., Nora, E., and Martienssen, R. 2008. RNA Interference Guides Histone Modification during the S Phase of Chromosomal Replication. *Curr Biol* **18**: 490-495.

Klymenko, T. and Muller, J. 2004. The histone methyltransferases Trithorax and Ash1 prevent transcriptional silencing by Polycomb group proteins. *EMBO Rep* **5**: 373-377.

Kornberg, R.D. 1974. Chromatin structure: a repeating unit of histones and DNA. *Science* **184**: 868-871.

-. 2007. The molecular basis of eucaryotic transcription. *Cell Death Differ* **14**: 1989-1997.

Koryakov, D.E., Reuter, G., Dimitri, P., and Zhimulev, I.F. 2006. The SuUR gene influences the distribution of heterochromatic proteins HP1 and SU(VAR)3-9 on nurse cell polytene chromosomes of Drosophila melanogaster. *Chromosoma* **115**: 296-310.

Kouzarides, T. 2007. Chromatin modifications and their function. *Cell* **128**: 693-705.

Krogan, N.J., Keogh, M.C., Datta, N., Sawa, C., Ryan, O.W., Ding, H., Haw, R.A., Pootoolal, J., Tong, A., Canadien, V. et al. 2003. A Snf2 family ATPase complex required for recruitment of the histone H2A variant Htz1. *Mol Cell* **12**: 1565-1576.

Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. 2001. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410**: 116-120.

Laroche, T., Martin, S.G., Gotta, M., Gorham, H.C., Pryde, F.E., Louis, E.J., and Gasser, S.M. 1998. Mutation of yeast Ku genes disrupts the subnuclear organization of telomeres. *Curr Biol* **8**: 653-656.

Leatherwood, J. and Vas, A. 2003. Connecting ORC and heterochromatin: why? *Cell Cycle* **2**: 573-575.

Lee, C.K., Shibata, Y., Rao, B., Strahl, B.D., and Lieb, J.D. 2004. Evidence for nucleosome depletion at active regulatory regions genome-wide. *Nat Genet* **36**: 900-905.

Lewis, P.W., Beall, E.L., Fleischer, T.C., Georlette, D., Link, A.J., and Botchan, M.R. 2004. Identification of a Drosophila Myb-E2F2/RBF transcriptional repressor complex. *Genes Dev* **18**: 2929-2940.

Li, B., Carey, M., and Workman, J.L. 2007. The role of chromatin during transcription. *Cell* **128**: 707-719.

Li, F., Chen, J., Izumi, M., Butler, M.C., Keezer, S.M., and Gilbert, D.M. 2001. The replication timing program of the Chinese hamster beta-globin locus is established coincident with its repositioning near peripheral heterochromatin in early G1 phase. *J Cell Biol* **154**: 283-292.

Lin, C.M., Fu, H., Martinovsky, M., Bouhassira, E., and Aladjem, M.I. 2003. Dynamic alterations of replication timing in mammalian cells. *Curr Biol* **13**: 1019-1028.

Lis, J. and Wu, C. 1993. Protein traffic on the heat shock promoter: parking, stalling, and trucking along. *Cell* **74**: 1-4.

Liu, C.L., Kaplan, T., Kim, M., Buratowski, S., Schreiber, S.L., Friedman, N., and Rando, O.J. 2005a. Single-nucleosome mapping of histone modifications in S. cerevisiae. *PLoS Biol* **3**: e328.

Liu, G., Malott, M., and Leffak, M. 2003. Multiple functional elements comprise a Mammalian chromosomal replicator. *Mol Cell Biol* **23**: 1832-1842.

Liu, L.P., Ni, J.Q., Shi, Y.D., Oakeley, E.J., and Sun, F.L. 2005b. Sex-specific role of Drosophila melanogaster HP1 in regulating chromatin structure and gene transcription. *Nat Genet* **37**: 1361-1366.

Lorch, Y., LaPointe, J.W., and Kornberg, R.D. 1987. Nucleosomes inhibit the initiation of transcription but allow chain elongation with the displacement of histones. *Cell* **49**: 203-210.

Loyola, A. and Almouzni, G. 2004. Histone chaperones, a supporting role in the limelight. *Biochim Biophys Acta* **1677**: 3-11.

Lu, B.Y., Emtage, P.C., Duyf, B.J., Hilliker, A.J., and Eissenberg, J.C. 2000. Heterochromatin protein 1 is required for the normal expression of two heterochromatin genes in Drosophila. *Genetics* **155**: 699-708.

Lucas, I., Palakodeti, A., Jiang, Y., Young, D.J., Jiang, N., Fernald, A.A., and Le Beau, M.M. 2007. High-throughput mapping of origins of replication in human cells. *EMBO Rep* **8**: 770-777.

Lucchesi, J.C., Kelly, W.G., and Panning, B. 2005. Chromatin remodeling in dosage compensation. *Annu Rev Genet* **39**: 615-651.

Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. 1997. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389**: 251-260.

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

MacAlpine, D.M. and Bell, S.P. 2005. A genomic view of eukaryotic DNA replication. *Chromosome Res* **13**: 309-326.

MacAlpine, D.M., Rodriguez, H.K., and Bell, S.P. 2004. Coordination of replication and transcription along a Drosophila chromosome. *Genes Dev* **18**: 3094-3105.

Machida, Y.J., Hamlin, J.L., and Dutta, A. 2005. Right place, right time, and only once: replication initiation in metazoans. *Cell* **123**: 13-24.

Makunin, I.V., Volkova, E.I., Belyaeva, E.S., Nabirochkina, E.N., Pirrotta, V., and Zhimulev, I.F. 2002. The Drosophila suppressor of underreplication protein binds to late-replicating regions of polytene chromosomes. *Genetics* **160**: 1023-1034.

Masuda, T., Mimura, S., and Takisawa, H. 2003. CDK- and Cdc45-dependent priming of the MCM complex on chromatin during S-phase in Xenopus egg extracts: possible activation of MCM helicase by association with Cdc45. *Genes Cells* **8**: 145-161.

Mavrich, T.N., Jiang, C., Ioshikhes, I.P., Li, X., Venters, B.J., Zanton, S.J., Tomsho, L.P., Qi, J., Glaser, R.L., Schuster, S.C. et al. 2008. Nucleosome organization in the Drosophila genome. *Nature*.

Mendjan, S. and Akhtar, A. 2007. The right dose for every sex. *Chromosoma* **116**: 95-106.

Mendjan, S., Taipale, M., Kind, J., Holz, H., Gebhardt, P., Schelder, M., Vermeulen, M., Buscaino, A., Duncan, K., Mueller, J. et al. 2006. Nuclear pore components are involved in the transcriptional regulation of dosage compensation in Drosophila. *Mol Cell* **21**: 811-823.

Mesner, L.D., Crawford, E.L., and Hamlin, J.L. 2006. Isolating apparently pure libraries of replication origins from complex genomes. *Mol Cell* **21**: 719-726.

Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.K., Koche, R.P. et al. 2007. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**: 553-560.

Mis, J., Ner, S.S., and Grigliatti, T.A. 2006. Identification of three histone methyltransferases in Drosophila: dG9a is a suppressor of PEV and is required for gene silencing. *Mol Genet Genomics* **275**: 513-526.

Mito, Y., Henikoff, J.G., and Henikoff, S. 2005. Genome-scale profiling of histone H3.3 replacement patterns. *Nat Genet* **37**: 1090-1097.

Mizuguchi, G., Tsukiyama, T., Wisniewski, J., and Wu, C. 1997. Role of nucleosome remodeling factor NURF in transcriptional activation of chromatin. *Mol Cell* **1**: 141-150.

Moir, R.D., Yoon, M., Khuon, S., and Goldman, R.D. 2000. Nuclear lamins A and B1: different pathways of assembly during nuclear envelope formation in living cells. *J Cell Biol* **151**: 1155-1168.

Moldovan, G.L., Pfander, B., and Jentsch, S. 2007. PCNA, the maestro of the replication fork. *Cell* **129**: 665-679.

Moshkin, Y.M., Armstrong, J.A., Maeda, R.K., Tamkun, J.W., Verrijzer, P., Kennison, J.A., and Karch, F. 2002. Histone chaperone ASF1 cooperates with the Brahma chromatin-remodelling machinery. *Genes Dev* **16**: 2621-2626.

Muller, H.J. and Altenburg, E. 1930. The Frequency of Translocations Produced by X-Rays in Drosophila. *Genetics* **15**: 283-311.

Muse, G.W., Gilchrist, D.A., Nechaev, S., Shah, R., Parker, J.S., Grissom, S.F., Zeitlinger, J., and Adelman, K. 2007. RNA polymerase is poised for activation across the genome. *Nat Genet* **39**: 1507-1511.

Ng, R.K. and Gurdon, J.B. 2008. Epigenetic memory of an active gene state depends on histone H3.3 incorporation into chromatin in the absence of transcription. *Nat Cell Biol* **10**: 102-109.

Nieduszynski, C.A., Knox, Y., and Donaldson, A.D. 2006. Genome-wide identification of replication origins in yeast by comparative genomics. *Genes Dev* **20**: 1874-1879.

Norio, P. 2006. DNA replication: the unbearable lightness of origins. *EMBO Rep* **7**: 779-781.

Norio, P., Kosiyatrakul, S., Yang, Q., Guan, Z., Brown, N.M., Thomas, S., Riblet, R., and Schildkraut, C.L. 2005. Progressive activation of DNA replication initiation in large domains of the immunoglobulin heavy chain locus during B cell development. *Mol Cell* **20**: 575-587.

Olins, A.L. and Olins, D.E. 1974. Spheroid chromatin units (v bodies). *Science* **183**: 330-332.

Osborne, C.S., Chakalova, L., Brown, K.E., Carter, D., Horton, A., Debrand, E., Goyenechea, B., Mitchell, J.A., Lopes, S., Reik, W. et al. 2004. Active genes dynamically colocalize to shared sites of ongoing transcription. *Nat Genet* **36**: 1065-1071.

Pak, D.T., Pflumm, M., Chesnokov, I., Huang, D.W., Kellum, R., Marr, J., Romanowski, P., and Botchan, M.R. 1997. Association of the origin recognition complex with heterochromatin and HP1 in higher eukaryotes. *Cell* **91**: 311-323.

Pal-Bhadra, M., Leibovitch, B.A., Gandhi, S.G., Rao, M., Bhadra, U., Birchler, J.A., and Elgin, S.C. 2004. Heterochromatic silencing and HP1 localization in Drosophila are dependent on the RNAi machinery. *Science* **303**: 669-672.

Peel, D.J., Johnson, S.A., and Milner, M.J. 1990. The ultrastructure of imaginal disc cells in primary cultures and during cell aggregation in continuous cell lines. *Tissue Cell* **22**: 749-758.

Perry, P., Sauer, S., Billon, N., Richardson, W.D., Spivakov, M., Warnes, G., Livesey, F.J., Merkenschlager, M., Fisher, A.G., and Azuara, V. 2004. A dynamic switch in the replication timing of key regulator genes in embryonic stem cells upon neural induction. *Cell Cycle* **3**: 1645-1650.

Peters, A.H. and Schubeler, D. 2005. Methylation of histones: playing memory with DNA. *Curr Opin Cell Biol* **17**: 230-238.

Pickersgill, H., Kalverda, B., de Wit, E., Talhout, W., Fornerod, M., and van Steensel, B. 2006. Characterization of the Drosophila melanogaster genome at the nuclear lamina. *Nat Genet* **38**: 1005-1014.

Pindyurin, A.V., Boldyreva, L.V., Shloma, V.V., Kolesnikova, T.D., Pokholkova, G.V., Andreyeva, E.N., Kozhevnikova, E.N., Ivanoschuk, I.G., Zarutskaya, E.A., Demakov, S.A. et al. 2008. Interaction between the Drosophila heterochromatin proteins SUUR and HP1. *J Cell Sci*.

Pindyurin, A.V., Moorman, C., de Wit, E., Belyakin, S.N., Belyaeva, E.S., Christophides, G.K., Kafatos, F.C., van Steensel, B., and Zhimulev, I.F. 2007. SUUR joins separate subsets of PcG, HP1 and B-type lamin targets in Drosophila. *J Cell Sci* **120**: 2344-2351.

Pokholok, D.K., Harbison, C.T., Levine, S., Cole, M., Hannett, N.M., Lee, T.I., Bell, G.W., Walker, K., Rolfe, P.A., Herbolsheimer, E. et al. 2005. Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell* **122**: 517-527.

Prasanth, S.G., Prasanth, K.V., Siddiqui, K., Spector, D.L., and Stillman, B. 2004. Human Orc2 localizes to centrosomes, centromeres and heterochromatin during chromosome inheritance. *Embo J* **23**: 2651-2663.

Prioleau, M.N., Gendron, M.C., and 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.

Quivy, J.P., Gerard, A., Cook, A.J., Roche, D., and Almouzni, G. 2008. The HP1-p150/CAF-1 interaction is required for pericentric heterochromatin replication and S-phase progression in mouse cells. *Nat Struct Mol Biol*.

R_Development_Core_Team. 2006. R: A language and environment for statistical computing. In (ed. R.F.f.S. Computing), Vienna, Austria.

Raghuraman, M.K., Winzeler, E.A., Collingwood, D., Hunt, S., Wodicka, L., Conway, A., Lockhart, D.J., Davis, R.W., Brewer, B.J., and Fangman, W.L. 2001. Replication dynamics of the yeast genome. *Science* **294**: 115-121.

Rando, O.J. 2007. Global patterns of histone modifications. *Curr Opin Genet Dev* **17**: 94-99.

Reinke, H. and Horz, W. 2003. Histones are first hyperacetylated and then lose contact with the activated PHO5 promoter. *Mol Cell* **11**: 1599-1607.

Remus, D., Beall, E.L., and Botchan, M.R. 2004. DNA topology, not DNA sequence, is a critical determinant for Drosophila ORC-DNA binding. *Embo J* **23**: 897-907.

Riddle, N.C. and Elgin, S.C. 2006. The dot chromosome of Drosophila: insights into chromatin states and their change over evolutionary time. *Chromosome Res* **14**: 405-416.

Roh, T.Y., Cuddapah, S., Cui, K., and Zhao, K. 2006. The genomic landscape of histone modifications in human T cells. *Proc Natl Acad Sci U S A* **103**: 15782-15787.

Rothstein, R., Michel, B., and Gangloff, S. 2000. Replication fork pausing and recombination or "gimme a break". *Genes Dev* **14**: 1-10.

Sadoni, N., Cardoso, M.C., Stelzer, E.H., Leonhardt, H., and Zink, D. 2004. Stable chromosomal units determine the spatial and temporal organization of DNA replication. *J Cell Sci* **117**: 5353-5365.

Sasaki, T. and Gilbert, D.M. 2007. The many faces of the origin recognition complex. *Curr Opin Cell Biol* **19**: 337-343.

Sasaki, T., Ramanathan, S., Okuno, Y., Kumagai, C., Shaikh, S.S., and Gilbert, D.M. 2006. The Chinese hamster dihydrofolate reductase replication origin decision point follows activation of transcription and suppresses initiation of replication within transcription units. *Mol Cell Biol* **26**: 1051-1062.

Sasaki, T., Sawado, T., Yamaguchi, M., and Shinomiya, T. 1999. Specification of regions of DNA replication initiation during embryogenesis in the 65-kilobase DNApolalpha-dE2F locus of Drosophila melanogaster. *Mol Cell Biol* **19**: 547-555.

Schepers, A., Ritzi, M., Bousset, K., Kremmer, E., Yates, J.L., Harwood, J., Diffley, J.F., and Hammerschmidt, W. 2001. Human origin recognition complex binds to the region of the latent origin of DNA replication of Epstein-Barr virus. *Embo J* **20**: 4588-4602.

Schliep, A., Georgi, B., Rungsarityotin, W., Costa, I.G., and Schönhuth, A. 2004. The General Hidden Markov Model Library: Analyzing Systems with Unobservable States. *Proceedings of the Heinz-Billing-Price*: 121-136.

Schotta, G., Ebert, A., Dorn, R., and Reuter, G. 2003. Position-effect variegation and the genetic dissection of chromatin regulation in Drosophila. *Semin Cell Dev Biol* **14**: 67-75.

Schotta, G., Ebert, A., Krauss, V., Fischer, A., Hoffmann, J., Rea, S., Jenuwein, T., Dorn, R., and Reuter, G. 2002. Central role of Drosophila SU(VAR)3-9 in histone H3-K9 methylation and heterochromatic gene silencing. *Embo J* **21**: 1121-1131.

Schreiber, S.L. and Bernstein, B.E. 2002. Signaling network model of chromatin. *Cell* **111**: 771-778.

Schubeler, D., Francastel, C., Cimbora, D.M., Reik, A., Martin, D.I., and Groudine, M. 2000. Nuclear localization and histone acetylation: a pathway for chromatin opening and transcriptional activation of the human beta-globin locus. *Genes Dev* **14**: 940-950.

Schubeler, D., MacAlpine, D.M., Scalzo, D., Wirbelauer, C., Kooperberg, C., van Leeuwen, F., Gottschling, D.E., O'Neill, L.P., Turner, B.M., Delrow, J. et al. 2004. The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. *Genes Dev* **18**: 1263-1271.

Schubeler, D., Scalzo, D., Kooperberg, C., van Steensel, B., Delrow, J., and Groudine, M. 2002. Genome-wide DNA replication profile for Drosophila melanogaster: a link between transcription and replication timing. *Nat Genet* **32**: 438-442.

Schulze, S.R., Sinclair, D.A., Fitzpatrick, K.A., and Honda, B.M. 2005. A genetic and molecular characterization of two proximal heterochromatic genes on chromosome 3 of Drosophila melanogaster. *Genetics* **169**: 2165-2177.

Schwaiger, M. and Schubeler, D. 2006. A question of timing: emerging links between transcription and replication. *Curr Opin Genet Dev* **16**: 177-183.

Schwartz, Y.B., Kahn, T.G., Nix, D.A., Li, X.Y., Bourgon, R., Biggin, M., and Pirrotta, V. 2006. Genome-wide analysis of Polycomb targets in Drosophila melanogaster. *Nat Genet* **38**: 700-705.

Schwartz, Y.B. and Pirrotta, V. 2007. Polycomb silencing mechanisms and the management of genomic programmes. *Nat Rev Genet* **8**: 9-22.

Sclafani, R.A. and Holzen, T.M. 2007. Cell cycle regulation of DNA replication. *Annu Rev Genet* **41**: 237-280.

Segal, E., Fondufe-Mittendorf, Y., Chen, L., Thastrom, A., Field, Y., Moore, I.K., Wang, J.P., and Widom, J. 2006. A genomic code for nucleosome positioning. *Nature* **442**: 772-778.

Seum, C., Reo, E., Peng, H., Rauscher, F.J., 3rd, Spierer, P., and Bontron, S. 2007. Drosophila SETDB1 is required for chromosome 4 silencing. *PLoS Genet* **3**: e76.

Shaffer, C.D., Stephens, G.E., Thompson, B.A., Funches, L., Bernat, J.A., Craig, C.A., and Elgin, S.C. 2002. Heterochromatin protein 2 (HP2), a partner of HP1 in Drosophila heterochromatin. *Proc Natl Acad Sci U S A* **99**: 14332-14337.

Shareef, M.M., King, C., Damaj, M., Badagu, R., Huang, D.W., and Kellum, R. 2001. Drosophila heterochromatin protein 1 (HP1)/origin recognition complex (ORC) protein is associated with HP1 and ORC and functions in heterochromatin-induced silencing. *Mol Biol Cell* **12**: 1671-1685.

Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J.R., Cole, P.A., Casero, R.A., and Shi, Y. 2004. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* **119**: 941-953.

Shimada, K. and Gasser, S.M. 2007. The origin recognition complex functions in sister-chromatid cohesion in Saccharomyces cerevisiae. *Cell* **128**: 85-99.

Shogren-Knaak, M., Ishii, H., Sun, J.M., Pazin, M.J., Davie, J.R., and Peterson, C.L. 2006. Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* **311**: 844-847.

Shumaker, D.K., Dechat, T., Kohlmaier, A., Adam, S.A., Bozovsky, M.R., Erdos, M.R., Eriksson, M., Goldman, A.E., Khuon, S., Collins, F.S. et al. 2006. Mutant nuclear lamin A leads to progressive alterations of epigenetic control in premature aging. *Proc Natl Acad Sci U S A* **103**: 8703-8708.

Sims, R.J., 3rd, Belotserkovskaya, R., and Reinberg, D. 2004. Elongation by RNA polymerase II: the short and long of it. *Genes Dev* **18**: 2437-2468.

Smith, E.R., Allis, C.D., and Lucchesi, J.C. 2001a. Linking global histone acetylation to the transcription enhancement of X-chromosomal genes in Drosophila males. *J Biol Chem* **276**: 31483-31486.

Smith, L., Plug, A., and Thayer, M. 2001b. Delayed replication timing leads to delayed mitotic chromosome condensation and chromosomal instability of chromosome translocations. *Proc Natl Acad Sci U S A* **98**: 13300-13305.

Spann, T.P., Moir, R.D., Goldman, A.E., Stick, R., and Goldman, R.D. 1997. Disruption of nuclear lamin organization alters the distribution of replication factors and inhibits DNA synthesis. *J Cell Biol* **136**: 1201-1212.

Spellman, P.T. and Rubin, G.M. 2002. Evidence for large domains of similarly expressed genes in the Drosophila genome. *J Biol* **1**: 5.

Srinivasan, S., Armstrong, J.A., Deuring, R., Dahlsveen, I.K., McNeill, H., and Tamkun, J.W. 2005. The Drosophila trithorax group protein Kismet facilitates an early step in transcriptional elongation by RNA Polymerase II. *Development* **132**: 1623-1635.

Stevenson, J.B. and Gottschling, D.E. 1999. Telomeric chromatin modulates replication timing near chromosome ends. *Genes Dev* **13**: 146-151.

Straub, T. and Becker, P.B. 2007. Dosage compensation: the beginning and end of generalization. *Nat Rev Genet* **8**: 47-57.

Straub, T., Gilfillan, G.D., Maier, V.K., and Becker, P.B. 2005. The Drosophila MSL complex activates the transcription of target genes. *Genes Dev* **19**: 2284-2288.

Suganuma, T., Gutierrez, J.L., Li, B., Florens, L., Swanson, S.K., Washburn, M.P., Abmayr, S.M., and Workman, J.L. 2008. ATAC is a double histone acetyltransferase complex that stimulates nucleosome sliding. *Nat Struct Mol Biol*.

Sullivan, S.A. and Landsman, D. 2003. Characterization of sequence variability in nucleosome core histone folds. *Proteins* **52**: 454-465.

Taddei, A., Van Houwe, G., Hediger, F., Kalck, V., Cubizolles, F., Schober, H., and Gasser, S.M. 2006. Nuclear pore association confers optimal expression levels for an inducible yeast gene. *Nature* **441**: 774-778.

Tagami, H., Ray-Gallet, D., Almouzni, G., and Nakatani, Y. 2004. Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. *Cell* **116**: 51-61.

Takei, Y., Assenberg, M., Tsujimoto, G., and Laskey, R. 2002. The MCM3 acetylase MCM3AP inhibits initiation, but not elongation, of DNA replication via interaction with MCM3. *J Biol Chem* **277**: 43121-43125.

Thomae, A.W., Pich, D., Brocher, J., Spindler, M.P., Berens, C., Hock, R., Hammerschmidt, W., and Schepers, A. 2008. Interaction between HMGA1a and the origin recognition complex creates site-specific replication origins. *Proc Natl Acad Sci U S A* **105**: 1692-1697.

Thomas, M.C. and Chiang, C.M. 2006. The general transcription machinery and general cofactors. *Crit Rev Biochem Mol Biol* **41**: 105-178.

Tolhuis, B., de Wit, E., Muijrers, I., Teunissen, H., Talhout, W., van Steensel, B., and van Lohuizen, M. 2006. Genome-wide profiling of PRC1 and PRC2 Polycomb chromatin binding in Drosophila melanogaster. *Nat Genet* **38**: 694-699.

Torres-Rosell, J., De Piccoli, G., Cordon-Preciado, V., Farmer, S., Jarmuz, A., Machin, F., Pasero, P., Lisby, M., Haber, J.E., and Aragon, L. 2007. Anaphase onset before complete DNA replication with intact checkpoint responses. *Science* **315**: 1411-1415.

Tremethick, D.J. 2007. Higher-order structures of chromatin: the elusive 30 nm fiber. *Cell* **128**: 651-654.

Turner, B.M., Birley, A.J., and Lavender, J. 1992. Histone H4 isoforms acetylated at specific lysine residues define individual chromosomes and chromatin domains in Drosophila polytene nuclei. *Cell* **69**: 375-384.

Tzeng, T.Y., Lee, C.H., Chan, L.W., and Shen, C.K. 2007. Epigenetic regulation of the Drosophila chromosome 4 by the histone H3K9 methyltransferase dSETDB1. *Proc Natl Acad Sci U S A* **104**: 12691-12696.

Vakoc, C.R., Mandat, S.A., Olenchock, B.A., and Blobel, G.A. 2005. Histone H3 lysine 9 methylation and HP1gamma are associated with transcription elongation through mammalian chromatin. *Mol Cell* **19**: 381-391.

Varga-Weisz, P.D., Wilm, M., Bonte, E., Dumas, K., Mann, M., and Becker, P.B. 1997. Chromatin-remodelling factor CHRAC contains the ATPases ISWI and topoisomerase II. *Nature* **388**: 598-602.

Verdel, A., Jia, S., Gerber, S., Sugiyama, T., Gygi, S., Grewal, S.I., and Moazed, D. 2004. RNAi-mediated targeting of heterochromatin by the RITS complex. *Science* **303**: 672-676.

Vogelauer, M., Rubbi, L., Lucas, I., Brewer, B.J., and Grunstein, M. 2002. Histone acetylation regulates the time of replication origin firing. *Mol Cell* **10**: 1223-1233.

Volpe, T.A., Kidner, C., Hall, I.M., Teng, G., Grewal, S.I., and Martienssen, R.A. 2002. Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* **297**: 1833-1837.

Wakimoto, B.T. and Hearn, M.G. 1990. The effects of chromosome rearrangements on the expression of heterochromatic genes in chromosome 2L of Drosophila melanogaster. *Genetics* **125**: 141-154.

White, E.J., Emanuelsson, O., Scalzo, D., Royce, T., Kosak, S., Oakeley, E.J., Weissman, S., Gerstein, M., Groudine, M., Snyder, M. et al. 2004. DNA replication-timing

analysis of human chromosome 22 at high resolution and different developmental states. *Proc Natl Acad Sci U S A* **101**: 17771-17776.

Wirbelauer, C., Bell, O., and Schubeler, D. 2005. Variant histone H3.3 is deposited at sites of nucleosomal displacement throughout transcribed genes while active histone modifications show a promoter-proximal bias. *Genes Dev* **19**: 1761-1766.

Wolfe, K.H., Sharp, P.M., and Li, W.H. 1989. Mutation rates differ among regions of the mammalian genome. *Nature* **337**: 283-285.

Wolffe, A.P. and Hayes, J.J. 1999. Chromatin disruption and modification. *Nucleic Acids Res* **27**: 711-720.

Woodfine, K., Beare, D.M., Ichimura, K., Debernardi, S., Mungall, A.J., Fiegler, H., Collins, V.P., Carter, N.P., and Dunham, I. 2005. Replication timing of human chromosome 6. *Cell Cycle* **4**: 172-176.

Woodfine, K., Fiegler, H., Beare, D.M., Collins, J.E., McCann, O.T., Young, B.D., Debernardi, S., Mott, R., Dunham, I., and Carter, N.P. 2004. Replication timing of the human genome. *Hum Mol Genet* **13**: 191-202.

Woodward, A.M., Gohler, T., Luciani, M.G., Oehlmann, M., Ge, X., Gartner, A., Jackson, D.A., and Blow, J.J. 2006. Excess Mcm2-7 license dormant origins of replication that can be used under conditions of replicative stress. *J Cell Biol* **173**: 673-683.

Wu, R., Singh, P.B., and Gilbert, D.M. 2006. Uncoupling global and fine-tuning replication timing determinants for mouse pericentric heterochromatin. *J Cell Biol* **174**: 185-194.

Wustmann, G., Szidonya, J., Taubert, H., and Reuter, G. 1989. The genetics of position-effect variegation modifying loci in Drosophila melanogaster. *Mol Gen Genet* **217**: 520-527.

Wysocka, J., Swigut, T., Xiao, H., Milne, T.A., Kwon, S.Y., Landry, J., Kauer, M., Tackett, A.J., Chait, B.T., Badenhorst, P. et al. 2006. A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* **442**: 86-90.

Yabuki, N., Terashima, H., and Kitada, K. 2002. Mapping of early firing origins on a replication profile of budding yeast. *Genes Cells* **7**: 781-789.

Yasuhara, J.C., DeCrease, C.H., and Wakimoto, B.T. 2005. Evolution of heterochromatic genes of Drosophila. *Proc Natl Acad Sci U S A* **102**: 10958-10963.

Yasuhara, J.C. and Wakimoto, B.T. 2006. Oxymoron no more: the expanding world of heterochromatic genes. *Trends Genet* **22**: 330-338.

-. 2008. Molecular landscape of modified histones in Drosophila heterochromatic genes and euchromatin-heterochromatin transition zones. *PLoS Genet* **4**: e16.

Yuan, G.C., Liu, Y.J., Dion, M.F., Slack, M.D., Wu, L.F., Altschuler, S.J., and Rando, O.J. 2005. Genome-scale identification of nucleosome positions in S. cerevisiae. *Science* **309**: 626-630.

Zappulla, D.C., Sternglanz, R., and Leatherwood, J. 2002. Control of replication timing by a transcriptional silencer. *Curr Biol* **12**: 869-875.

Zeitlinger, J., Stark, A., Kellis, M., Hong, J.W., Nechaev, S., Adelman, K., Levine, M., and Young, R.A. 2007. RNA polymerase stalling at developmental control genes in the Drosophila melanogaster embryo. *Nat Genet* **39**: 1512-1516.

Zhang, J., Xu, F., Hashimshony, T., Keshet, I., and Cedar, H. 2002. Establishment of transcriptional competence in early and late S phase. *Nature* **420**: 198-202.

Zhimulev, I.F. and Belyaeva, E.S. 2003. Intercalary heterochromatin and genetic silencing. *Bioessays* **25**: 1040-1051.

Zhou, J., Ashouian, N., Delepine, M., Matsuda, F., Chevillard, C., Riblet, R., Schildkraut, C.L., and Birshtein, B.K. 2002a. The origin of a developmentally regulated Igh replicon is located near the border of regulatory domains for Igh replication and expression. *Proc Natl Acad Sci U S A* **99**: 13693-13698.

Zhou, J., Ermakova, O.V., Riblet, R., Birshtein, B.K., and Schildkraut, C.L. 2002b. Replication and subnuclear location dynamics of the immunoglobulin heavy-chain locus in B-lineage cells. *Mol Cell Biol* **22**: 4876-4889.

Zhou, J., Saleque, S., Ermakova, O., Sepulveda, M.A., Yang, Q., Eckhardt, L.A., Schildkraut, C.L., and Birshtein, B.K. 2005. Changes in replication, nuclear location, and expression of the Igh locus after fusion of a pre-B cell line with a T cell line. *J Immunol* **175**: 2317-2320.

Zink, D. 2006. The temporal program of DNA replication: new insights into old questions. *Chromosoma* **115**: 273-287.

# 5. Curriculum Vitae

**Michaela Schwaiger**
Kirchgasse 7
CH-4058 Basel
Switzerland
Email: michaela.schwaiger@fmi.ch

Date of birth: August 18th, 1980
Nationality: Austrian

## *Education*

| | | |
|---|---|---|
| Graduate: | **Friedrich Miescher Institute for Biomedical Research (FMI)** Basel, Switzerland<br>Project title: "Dynamics and control of replication timing in *Drosophila* cells"<br>Adivisor: **Dr. Dirk Schübeler** | since 2004/03 |
| Diploma thesis: | **Institute of Molecular Pathology (IMP)**<br>Vienna, Austria<br>Project title: "The role of SMC (structural maintenance of chromosomes) proteins in *Xist* localisation"<br>Adivisor: **Dr. Anton Wutz** | 2002/04-2003/07 |
| Undergraduate: | Biology/Genetics<br>University of Vienna<br>Vienna, Austria | 1998/10-2003/07 |

## *Awards and Fellowships*

| | |
|---|---|
| FASEB Poster Award | 2008/06/27 |
| Epigenome NoE travel award | 2007/05/02 |
| Boehringer Ingelheim Fonds: PhD scholarship | 2004/10-2007/03 |
| Erasmus program: undergraduate studies at University of Leicester, U.K. | 2001/01-2001/06 |

## *Selected Work Experience*

**University of British Columbia** 2003/11–2004/01
Vancouver, Canada; Department: Medical Genetics
Project title: RNAi in *C.elegans*
Supervisor: **Prof. Ann Rose**

**Boehringer-Ingelheim Austria** 2000/08-09
Vienna, Austria; Department: NCE-Discovery 2001/08-09
Project title: The effect of Bombesin on the development of *C.elegans*
Supervisor: Dr. Jacques Van Meel/ Dr. Marlene Dengg

### *Publications*

**Schwaiger M**, Stadler MB, Kohler H, Oakekely EJ, Schubeler, D: Chromatin structure marks cell-type and gender-specific replication of the *Drosophila* genome. *In review at Genes Dev*

Hiratani I, Ryba T, Itoh M, Yokochi T, **Schwaiger M**, Chang C, Lyou Y, Townes T, Schubeler D, Gilbert DM: Global Re-organization of Replication Domains During Embryonic Stem Cell Differentiation. *PloS Biology, in press*

Bell O, Wirbelauer C, Hild M, Scharf AN, **Schwaiger M**, Macalpine DM, Zilbermann F, van Leeuwen F, Bell SP, Imhof A, Garza D, Peters AH, Schubeler D: Localized H3K36 methylation states define histone H4K16 acetylation during transcriptional elongation in *Drosophila*. *EMBO J,* 2007:26(24): 4974-4984

**Schwaiger M**, Schubeler, D: A question of timing: Emerging links between transcription and replication. *Curr Opin Genet Dev*, 2006:16(2):177-83

### *Posters*

Michaela Schwaiger, Michael B. Stadler, Hubertus Kohler, Edward J. Oakeley, and Dirk Schübeler, "Chromatin structure marks cell-type and gender specific replication of the *Drosophila* genome." **FASEB Summer Research Conference on Transcriptional Regulation During Cell Growth, Differentiation, and Development. Snowmass Village, CO, USA. June 22nd to 27th, 2008**

Michaela Schwaiger and Dirk Schübeler, "Defining dynamic replication timing and its link with chromatin and transcription in *Drosophila* cells." **Gordon Conference on Epigenetics. Holderness School, Holderness, NH, USA. August 5 to 10, 2007**

Michaela Schwaiger and Dirk Schübeler, "Dynamics and control of replication timing in *Drosophila* cells." **Meeting on DNA replication and genome integrity. Salk Institute, La Jolla, CA, USA. August 9 to 13, 2006**

Michaela Schwaiger and Dirk Schübeler, "Dynamics and control of replication timing in *Drosophila* cells**." Meeting on Eukaryotic DNA replication. Cold Spring Harbor Laboratory, NY, USA. September 7 to 11, 2005**

### *Presentations*

Michaela Schwaiger, Edward J. Oakeley, Michael B. Stadler and Dirk Schübeler, "The genomic landscape of DNA replication timing and its crosstalk with chromatin and transcription in Drosophila cells**." FMI Annual Meeting. Grindelwald, Switzerland. September 19 to 21, 2007**

Michaela Schwaiger and Dirk Schübeler, "Defining dynamic replication timing and its link with chromatin and transcription in *Drosophila* cells." **Meeting on Eukaryotic DNA Replication and Genome Maintenance. Cold Spring Harbor Laboratory, NY, USA. September 5 to 9, 2007**

Michaela Schwaiger and Dirk Schübeler, "Defining dynamic replication timing and its link with chromatin and transcription in *Drosophila* cells." **Gregor Mendel Institute of Molecular Plant Biology, Vienna, Austria. April 5th, 2007**

Michaela Schwaiger and Dirk Schübeler, "Dynamics and control of the timing of DNA replication in *Drosophila melanogaster*." **22nd Summer Seminar of the Boehringer Ingelheim Fonds in Hirschegg, Austria. August 20 to 26, 2006**

Michaela Schwaiger and Dirk Schübeler, "Dynamics and control of the timing of DNA replication in *Drosophila melanogaster*." **21st Summer Seminar of the Boehringer Ingelheim Fonds in Hirschegg, Austria. August 21 to 27, 2005**