Developmental regulation of DNA replication timing at the human β globin locus

Itamar Simon¹, Toyoaki Tenzen^{1,2},
Raul Mostoslavsky¹, Eitan Fibach³,
Laura Lande¹, Eric Milot⁴, Joost Gribnau⁴,
Frank Grosveld⁴, Peter Fraser^{4,5} and
Howard Cedar^{1,6}

¹Department of Cellular Biochemistry and ³Department of Hematology, Hebrew University Medical School, Jerusalem, Israel 91120, ²Department of Evolutionary Genetics, National Institute of Genetics, Mishima, Shizuoka-ken, Japan 411-8540 and ⁴MGC Department of Cell Biology and Genetics, Erasmus University, PO Box 1738, 3000 DR Rotterdam, The Netherlands

⁵Present address: Laboratory of Chromatin and Gene Expression, The Babraham Institute, Babraham, Cambridge CB2 4AT, UK

⁶Corresponding author
e-mail: cedar@md2.huji.ac.il

I.Simon and T.Tenzen contributed equally to this work

The human β globin locus replicates late in most cell types, but becomes early replicating in erythroid cells. Using FISH to map DNA replication timing around the endogenous β globin locus and by applying a genetic approach in transgenic mice, we have demonstrated that both the late and early replication states are controlled by regulatory elements within the locus control region. These results also show that the pattern of replication timing is set up by mechanisms that work independently of gene transcription.

Keywords: β globin/LCR/replication timing

Introduction

The mammalian genome is made up of defined time zones that undergo DNA replication in a programmed manner during S phase. By studying individual genes, it has been shown that there is a relatively straightforward correlation between replication timing and gene expression (Holmquist, 1987; Selig *et al.*, 1992). Thus, housekeeping genes replicate early in all cell types, while some tissue-specific gene regions are developmentally regulated, replicating late in most cell types while undergoing DNA synthesis early in the tissue of expression. This relationship can also be observed at the chromosomal level where it has been shown that early replicating bands co-map with the open DNase I-sensitive regions of the genome (Kerem *et al.*, 1984).

Studies in yeast indicate that replication timing is controlled by interactions between *cis*-acting sequences and *trans*-acting factors, which ultimately impact on the firing of local origins (Simon and Cedar, 1996). Very little is known, however, about how replication timing units are organized and regulated in the mammalian genome. The human β globin domain represents a good system in which to study this process. This entire locus, which

encompasses a number of different developmentally controlled globin genes, replicates relatively late in non-expressing cell types, but has been shown to be early replicating in erythroid cells in culture (Epner *et al.*, 1988; Dhar *et al.*, 1989). Studies using lymphoblasts from patients with Hispanic thalassemia fused to MEL cells provided genetic evidence that this switch is probably mediated by sequences located within a large 40 kb region upstream of the globin gene cluster (Forrester *et al.*, 1990), but little was done to map the control elements or determine how they regulate developmentally specific replication timing.

Here, we have used fluorescence *in situ* hybridization (FISH) analysis to map and characterize the replication time zones surrounding the endogenous human globin locus on chromosome 11. These studies allowed us to delineate a late replicating domain in non-erythroid cells as well as an expanded early replicating domain in erythroid cells, and suggest that this timing mechanism is controlled by nearby *cis*-acting sequences. By employing a series of transgenic mice, we demonstrate that sequences located within the locus control region (LCR) are sufficient for setting up the correct developmentally regulated replication timing pattern *in vivo*. These data also shed new light on the relationship between gene expression and replication timing.

Results

Mapping replication time zones

In order to understand better the organization of replication time zones around the β globin locus on chromosome 11 in human cells we utilized a series of PAC clones to map replication timing in this region by FISH analysis in interphase nuclei. In this method, the two alleles show two single dots before replication (called SS) and two double dots when both are replicated (called DD) (Selig *et al.*, 1992). A high percentage of single dots in S phase cells is indicative of late replication. In contrast to biochemical techniques, this method does not require cell cycle fractionation or synchronization and can even be used on small cell populations *in vivo* (Simon *et al.*, 1999). It is also particularly appropriate for detecting differences in replication timing between the two alleles. In general, replication at individual sites in the genome is synchronous, with almost all nuclei having either SS or DD signals and very few (~15–20%) showing one single and one double dot (called SD), indicating that one allele has replicated earlier than the other. However, regions containing imprinted genes or other monoallelically expressed arrays replicate asynchronously with a high single–double count (>25%) (Kitsberg *et al.*, 1993a; Chess *et al.*, 1994; Mostoslavsky *et al.*, 1999).

Analysis of a large region surrounding the globin gene locus in non-erythroid cells (lymphoblasts or fibroblasts) reveals that many individual probes from this portion of chromosome 11 replicate asynchronously with >30% of the nuclei containing an SD pattern (red bars in Figure 1). This finding was not completely unexpected, since this region is known to harbor both monoallelically expressed olfactory receptor genes (Bulger *et al.*, 1999) and large imprinted domains that include the IGF2, H19 and P57KIP2 genes (Feinberg, 1999), all known to replicate asynchronously in a variety of cell types. In contrast to this overall picture, the two alleles within an ~200 kb domain (denoted as HBB on the map) appear to replicate synchronously (blue bars), as indicated by an SD count of <20%. Since these probes show a relatively high percentage of cells with SS signals (gray bars), it appears that this domain undergoes DNA synthesis in middle to late S phase, and because of its synchrony is strikingly different from the surrounding sequences (Figure 1).

This pattern is dramatically altered in globin-expressing human erythroblasts, where a much larger region of ~1 Mb becomes synchronous and very early replicating (as indicated by the low SS count). It should be noted that for all of these cell types the demarcation between synchronous (blue) and asynchronous (red) replication profiles is quite clear cut, and statistically significant (Figure 1B). Thus, despite the regional dominance of asynchronous replication, both alleles of the globin locus are set to replicate in unison at a specific point in S phase with the domain itself being relatively small (200–300 kb) in non-erythroid cells, but much larger (~1 Mb) in the erythroid cell type. This suggests that sequence elements within this region may be involved in setting up both the early and late replication timing patterns, which appear to be separate from the surrounding time zones.

The LCR directs replication timing in transgenic mice

On the assumption that upstream sequences may be involved in the switch to early replication, we next asked whether the LCR itself plays a role in the regulation of replication timing *in vivo*. To this end, we analyzed the replication properties of a series of single or low copy mouse transgenic lines carrying the LCR attached either to the full complement of globin genes or the γ and β genes alone (Figure 2A). For each mouse strain, we isolated erythroid cells from fetal liver or embryonic blood, and non-erythroid cells (fibroblasts or lymphocytes), and measured replication timing of the human globin sequences using the FISH method. Since these studies were carried out on transgene heterozygotes, each nucleus shows only one hybridization locus composed of either a single or double signal. In the four LCR-wild-type lines

Fig. 1. Replication timing pattern in the HBB region. (A) FISH analysis was carried out on human EBV-transformed lymphoblasts, embryonic fibroblasts or peripheral blood-derived erythroblasts using a variety of different probes. In each experiment at least 100 S-phase cells (BrdU positive) were analyzed by counting SS, SD and DD patterns and then normalizing the results to the replication timing patterns of SNRPN and CD3D (see Materials and methods). For each probe, the percentage of cells with SS signals is presented as gray bars, and the percentage of cells with an SD signal as either red (asynchronous) or blue (synchronous) bars. Using this form of presentation, the amount of SS signals signifies the point in the cell cycle where the first allele replicates, while the percentage of SD signals indicates the additional movement through S phase before the second allele replicates. The results for lymphoblasts were averaged from analyses of two different cell lines. The positions of genetic markers in human chromosome region 11p15.5 and the probes (see Materials and methods for details) used in these experiments are shown on the map. The numbers indicate distances in Mb from the subtelomeric repeats. The smaller map includes details from the HBB region, with the LCR (red box), the various globin genes (blue lines) and local probes (orange). Using S-phase fractionation, we have confirmed that the two alleles of IGF2 replicate asynchronously (Simon *et al.*, 1999), while β globin replicates synchronously in lymphoblast cells (data not shown). (B) Data (% SD) for the asynchronously replicating (28–43%) probes (red) and synchronously replicating probes (blue) from all of the cell types are shown in graphic form. Although <20% SD is considered as synchronous replication, probe i in erythroblasts (26% SD) was included in this category since it is located at the border of a synchronous domain and is still less than any of the asynchronously replicating probes. These two populations are significantly different ($P < 0.001$) as determined by the Mann–Whitney *U*-test.

Fig. 2. Components of the LCR control replication timing. **(A)** Constructs used for generating the transgenic lines, showing the LCR (orange box), the five hypersensitive sites (HSs) (arrows) and the globin gene sequences (colored boxes). **(B)** The un-normalized times of replication (percent singlets) of the transgenic human (H) globin gene (probe 1329) and the endogenous mouse (M) globin region were determined by FISH by counting 200–300 nuclei per sample. $\Delta 2$ -b, $\Delta 2$ -c, $\Delta 4$ -a, $\Delta 4$ -b, $\Delta 4$ -c and μD -14 are specific mouse lines made from the constructs shown in (A). **(C)** A graphic representation of transgene replication timing. All of the data are normalized to the replication values (in the same cell population) from a combination of probes (Materials and methods). The adjusted replication times in embryonic liver (red) and fibroblasts (blue) are shown for each transgenic mouse. **(D)** The distribution of replication timing values in non-erythroid cells for the LCR mutants (LCR⁻) has a large standard deviation (± 13.7), which is significantly different ($P = 0.005$ using Levene's test for equality of variances) from that of the wild-type (LCR⁺) animals (± 1.1), whose values range from 58 to 61% singles. In addition to the transgenes shown in (B) and (C), this graph includes three additional mouse founders containing a construct carrying the globin genes without the LCR, which were also analyzed for replication timing in embryonic fibroblasts and these gave results of 40, 68 and 75% singles. Erythroid replication (28–30%) in the $\Delta 4$ series is similar ($P < 0.001$) to that of the wild-type animals (29–32%).

(LnS2, Ln72, Ln2 and Ln15) analyzed in this experiment, the transgene undergoes replication in the late half of S phase (~60% singles) in non-erythroid cells, but is clearly shifted to an earlier time in erythroid cell types derived from either fetal liver or embryonic blood (~30% singles) (Figure 2B).

Because each of these FISH determinations is obtained from individual cell populations that inherently may have slight variations in their cell cycle properties, it was

necessary to normalize these data by comparing with standard gene sequences. Figure 2B demonstrates how this can be done using the endogenous mouse globin gene as a single S-phase marker. The replication time of the LCR-wild-type constructs in fibroblasts, for example, varies in absolute values between 55 and 62% singles. In each individual sample, however, this exogenous DNA always replicates 10 (± 1) percentage points later than the endogenous mouse globin control, indicating that the

Fig. 3. Transgene replication utilizes the β globin origin. Replication direction analysis of the wild-type β -globin transgene construct in fibroblasts. Leading-strand BrdU-labeled DNA (1 μ g) was prepared as described previously (Kitsberg *et al.*, 1993b), placed on identical filters and hybridized to plus (+) or minus (-) strand riboprobes. Marker (M) DNA (1 μ g) is included in every experiment to correct for differences in hybridization efficiencies of the two complementary probes. Probes F and G are on opposite sides of the presumed origin from I and J (Kitsberg *et al.*, 1993b). In the case of I, for example, the plus probe hybridizes poorly to the BrdU DNA, but the minus probe gives a strong signal. Thus, in this region, the plus DNA represents the leading strand, and we can conclude that this fragment replicates to the right. Probe F is homologous to regions around both $G\gamma$ and $A\gamma$. To corroborate these results we also carried out double-label FISH analysis using two cosmid probes, one covering the LCR (HG4) and the other covering the β -like genes (HG-28TK). In spleen cells from Ln2, 12% (25/210) of the nuclei showed a double HG-28TK signal together with a single HG4 signal, and only 3% (7/210) showed the opposite pattern. Similar results were obtained for erythroid cells from fetal liver (10% 20/200, 2% 4/200). These results indicate that replication in this region proceeds leftward, and this is consistent with firing at the presumed origin.

transgene actually undergoes DNA synthesis at a fixed position in S phase with relatively little variation. A similar picture emerges from the data for replication timing in erythroid cells as well (Figure 2B). In order to attain even further accuracy we systematically normalized all of the replication timing data using a panel of three different endogenous gene sequences (see Materials and methods) and the results of this analysis are presented graphically in Figure 2C. When this is done, it can clearly be seen that for each of these mice, the human globin transgene replicates relatively late within a small window of time close to 60% singles in non-erythroid cells (Figure 2C, blue squares) and close to 30% singles (early) in erythroid cells (red squares). It should be noted that this regulation process takes place despite the fact that each construct is probably integrated at a different chromosomal site (Milot *et al.*, 1996), clearly suggesting that these transgenes must contain *cis*-acting sequences that can direct replication timing in a dominant manner.

In vivo, the initiation of replication at a specific time in S phase is apparently carried out through the action of local *cis*-acting elements that control the firing of nearby replication origins (Stillman, 1993). Previous studies have shown that in both erythroid and non-erythroid cells the endogenous human globin locus undergoes replication from a single specific origin located near the β gene (Kitsberg *et al.*, 1993b; Aladjem *et al.*, 1995). In order to test whether the transgene also utilizes this same *cis*-acting sequence we employed leading-strand replication analysis to map origin activity in fibroblast cells derived from one of the founder mouse lines (Ln2). In this method, newly synthesized (BrdU labeled) leading-strand DNA is hybridized to plus and minus strand probes covering the full length of the construct and the specificity of this hybridization serves to determine the direction of DNA synthesis. An origin is defined as that point on the DNA where replication fork movement changes direction (Handeli *et al.*, 1989). As shown in Figure 3, transgene replication indeed proceeds bidirectionally away from a

position near the β globin gene in a manner identical to that observed for the endogenous human globin locus itself (Kitsberg *et al.*, 1993b; Aladjem *et al.*, 1995). These results were also confirmed in both erythroid and non-erythroid cells using FISH methodology to determine the direction of fork movement between two adjacent probes in this region (see legend to Figure 3). These findings indicate that the transgene constructs used in these experiments are not only capable of directing differential replication timing, but do this by operating on the same origin as is normally used *in vivo*.

In light of the observation that *cis*-acting elements within the locus operate in a dominant manner to set up a developmentally controlled replication pattern, we next asked whether these effectors can also extend their influence to adjacent cellular DNA. To this end, we cloned the integration site of one transgenic line, Ln2 (Strouboulis *et al.*, 1992), and by means of this sequence, then isolated a representative BAC probe that could be used for FISH analysis. When tested in a normal mouse line (μ D) the intact endogenous sequence replicates in middle S phase (~50% singles) in both non-erythroid and erythroid cell types (Figure 4A). However, when this same sequence serves as the integration site for the globin transgene (i.e. in Ln2 itself), it becomes early replicating in erythroid cells (26% singles), and thus undergoes DNA synthesis at approximately the same time as the integrated human globin gene (29% singles). Indeed, using double-label FISH, it is actually possible to observe visually that the integration site and the human globin insert are almost always in the same replication state (Figure 4B). These results clearly indicate that elements within the transgene can dictate the replication timing properties of adjacent cellular sequences.

LCR mutation analysis

In order to prove that replication timing is indeed controlled by elements within the LCR itself, we next generated transgenic mice using a DNA fragment that

contains all of the hypersensitive sites (HS1–HS5), without any additional globin sequences. Once again, these mice exhibited normal replication timing control with the transgene replicating late in fibroblasts and early in erythroid cells (Figure 2B and C). We next asked which sequences within the LCR may be necessary for directing replication timing, and this was done by utilizing transgenic mice harboring mutant LCR sequences. Three independent mouse lines carrying a deletion (Δ 4) of HS4 (Milot *et al.*, 1996), for example, exhibited normal early replication in erythroid cells (~30% singles) but were unable to direct a fixed late replication timing pattern in fibroblasts (see Figure 2B and C).

In order to understand better how this mechanism works, we isolated integration site sequences from the Δ 4-a and Δ 4-c mice and then used FISH to determine their replication profiles at the normal as well as the disrupted locus (Figure 4). Both of these genomic sites undergo replication in the middle of S phase (~50% singles) in erythroid as well as non-erythroid cell types. Strikingly, however, when juxtaposed to the globin transgene these same sites become early replicating in erythroid cells (33% singles), clearly suggesting that the mutant LCR is still able to direct erythroid-specific replication timing in a dominant manner. In contrast, it appears that in non-erythroid cells the globin transgene has lost its ability to set the late replication time profile, and instead, falls under the control of the adjacent endogenous time zone at the integration site (Figure 4A), thus explaining why in both of these cases, the exogenous globin sequences replicate in the middle of S phase (50 and 48% singles, respectively).

As a further attempt to characterize the elements involved in replication timing, we analyzed mice carrying a globin transgene with a deletion of HS2 (Δ 2) (Milot *et al.*, 1996). In these lines, both erythroid- and non-erythroid-specific replication times are disrupted and the locus now replicates at about the same time point in both cell types. Similarly, a transgenic construct carrying a compact LCR (μ D) made up exclusively of the four hypersensitive sites themselves, without intervening sequences (see Figure 2A) (Ellis *et al.*, 1996), showed a completely disrupted timing pattern, with replication taking place even later in erythroid than in non-erythroid cells (Figure 2B and C). This suggests that the full complement of hypersensitive sites alone is not sufficient to direct replication timing control, and additional elements within the LCR must be required.

The results using these mutants clearly support the conclusion that the LCR itself can direct proper replication timing in a developmentally regulated manner. In order to further confirm that this is indeed the case, however, we carried out a statistical analysis on the data obtained from all of the transgenic mice used in this study, together with several other lines that carry the globin genes without the LCR. As can be seen in Figure 2C and summarized in D, all of the transgenes that contain an intact LCR replicate at a fixed time in non-erythroid cells regardless of their integration sites. In contrast, transgenes that are mutant, or completely lack the LCR, demonstrate a wide spectrum of replication times ($P = 0.005$), probably because they are

Fig. 4. Replication analysis of transgene integration sites. (A) Replication timing of the integration site on the allele carrying the transgene in mouse lines Ln2, Δ 4-a and Δ 4-c was scored by double labeling using probe CosHG-28TK to detect the human globin sequences. The normal replication timing of each integration site was determined on wild-type alleles using mice that do not have a transgene at this site (i.e. lines μ D, Δ 2-b or Ln15). The instances where the transgene alters replication timing of the integration site are highlighted in red. These data have been normalized for S-phase position as described in the legend to Figure 2. (B) Examples of BrdU-positive (blue AMCA-labeled) nuclei analyzed by FISH. The nucleus on the left is from Ln2 fetal liver. Globin is labeled with fluorescein and the integration site (probe 212a06) with rhodamine. Note that in this cell, the integration site on the transgenic allele replicates earlier (double dot) than the normal allele (single dot). The nucleus on the right is from Ln Δ 4-a fibroblasts. Photomicrographs were prepared as described (Selig *et al.*, 1992).

influenced by the replication timing properties of their sites of integration.

Discussion

The initiation of DNA replication in eukaryotes involves two types of *cis*-acting components (Stillman, 1993). DNA synthesis itself begins at defined origin sequences, but the timing of this process appears to be controlled by a separate set of *cis*-acting elements. It has been shown in yeast, for example, that a single origin can be made to fire

at different times during S phase simply by placing it near elements within the genome that regulate early or late replication timing (Ferguson and Fangman, 1992; Friedman *et al.*, 1996). In the human β globin locus, as well, a single origin upstream of the β gene is used for both early and late replication modes under normal circumstances (Kitsberg *et al.*, 1993b; Aladjem *et al.*, 1995, 1998).

Our studies have begun to shed light on the second component of the replication machinery, that which interacts with the origin to control replication timing. Previous studies had shown that the region surrounding the human globin genes replicates early in erythroid cells and late in non-erythroid cell types, but because of the limited scope of these analyses, it was not possible to decipher how this is controlled in the context of the chromosome. By examining replication over a large span of chromosome 11, we have been able to actually define the early and late replication domains. This mapping experiment not only serves to outline the boundaries of these replication time zones, but also demonstrates that both early and late replication patterns must result from local control elements that set up a fixed replication time on both alleles, and are not just the default state dictated by surrounding asynchronously replicating DNA.

Using transgenes, we have demonstrated that the LCR region (HS1–5) is sufficient for directing replication timing in a developmentally specific manner *in vivo*. This cluster of elements appears to work dominantly to set up both the early and late replication timing patterns independently of the integration site, and can even take over the replication of surrounding genomic sequences. In contrast, without the full LCR, transgenes passively adopt the replication time properties of the insertion locus itself (see Figure 4A). We have shown that the LCR can operate on its own natural β globin origin when it is present nearby, but studies using constructs that lack this origin (LCR 3, 4 and 8) clearly indicate that the LCR can also activate alternate nearby origins if needed. This is consistent with autoradiography and origin mapping studies that have clearly demonstrated that each replication time zone actually contains multiple origins under coordinate timing control (for review see Simon and Cedar, 1996).

Although the LCR (HS1–5) is clearly sufficient for directing replication timing in transgenic mice, a targeted endogenous deletion of this region does not appear to affect the ability of the locus to switch to early replication in erythroid cell hybrids (Cimbora *et al.*, 2000), indicating that it may not be necessary for this process *in vivo*. Early replication, however, can not be established in hybrid cells from patients with Hispanic thalassemia where the genotype is characterized by a much more extensive deletion, which includes an additional 27 kb 5' to HS5 (Forrester *et al.*, 1990). When taken together, these observations are consistent with the idea that the full genomic LCR must be larger than originally thought and probably includes additional redundant regulatory elements located further upstream.

The limited mutation analysis carried out in this work does not reveal a great deal about the precise sequences that control replication timing. However, it is quite evident that multiple elements both within and outside the

hypersensitive site fragments themselves are required for proper regulation. In addition, our data with $\Delta 4$ mutants suggest that separate elements may be involved in the setting up of early and late replication timing patterns. When taken together, these genetic studies suggest that replication timing must be regulated by multiple complex elements, as is also the case in yeast (Friedman *et al.*, 1996).

It is well established that the LCR (HS1–5) plays a role in setting up regional erythroid-specific chromatin structure in transgenic mice, and it is likely that this function is intertwined with the ability to direct early replication timing. This is generally borne out by analyses of the mutant LCRs used in this study. $\Delta 4$ transgenes, for example, are able to generate a DNase I-sensitive conformation in erythroid cells and also show an early replication pattern, while the $\Delta 2$ transgenes are DNase I insensitive and have lost the ability to replicate properly (Milot *et al.*, 1996) (Figure 2). It should be noted, however, that experiments with the μD transgene suggest that it is possible to separate some of the elements that control these two structural parameters, since this construct is known to adopt a DNase I-sensitive conformation in blood cells (Ellis *et al.*, 1996) even though its replication pattern is defective. It is interesting to note that replication timing decisions are evidently made during a small window of time following mitosis and that this is coincident with the spatial re-positioning of chromosomal domains within the nucleus (Dimitrova and Gilbert, 1999), adding further support to the idea that replication timing is intimately linked to chromatin structure.

Although early replication timing is generally correlated with gene expression, it has not been possible to decipher the cause and effect relationship between these two parameters (Simon and Cedar, 1996). Recent results using a targeted deletion of the LCR (HS1–5) showed that early replication timing and an open chromatin structure do not by themselves guarantee high levels of globin transcription in erythroid cells (Cimbora *et al.*, 2000). Conversely, we have demonstrated that the μD transgene, which correctly expresses the β globin gene at full levels (Ellis *et al.*, 1996), undergoes erythroid replication inappropriately in middle/late S phase, strongly suggesting that it is not transcription itself that causes early replication. Taken together, these findings indicate that the control of replication timing is mediated by a designated class of *cis*-acting elements, independently of transcription.

It should be noted that all previous studies have put emphasis on the relationship between early replication and globin transcription in erythroid cells. One of the important findings to come out of the experiments described here is that elements within the LCR also function in non-expressing cell types. It is thus possible that one of the major roles of replication timing control at the globin locus is to set up late replication and its accompanying inactive chromatin structure in non-erythroid cells, and in this way perhaps bring about the repression of background transcription. One way that this may be accomplished is by restricting the exposure of newly assembled nucleosomes to histone deacetylases specifically during replication in late S phase (Allshire and Bickmore, 2000), and recent studies showing that HDAC2 is preferentially associated

with late replication foci (Rountree *et al.*, 2000) strongly support this concept.

Materials and methods

Fluorescence in situ hybridization

FISH was performed as described previously (Lichter *et al.*, 1988, 1990). Briefly, denaturation was carried out by incubation in 70% deionized formamide, 2× SSC at 68°C for 2 min, and then slides were dehydrated by a series of ice-cold ethanol washes (70, 90 and 100% for 5 min each). RNA-free cosmid, BAC or PAC DNA was labeled by nick-translation, substituting dTTP with bio-16-dUTP (Boehringer Mannheim) or with dTTP and digoxigenin-11-dUTP (Boehringer Mannheim) in a ratio of 2:1. The critical size range of probe molecules (<500 bp and preferably 150–250 bp) was achieved by empirically varying the amount of DNase I in the nick-translation reaction. Unincorporated nucleotides were separated from the probe DNA by centrifugation through 1 ml Sephadex G-50 columns (Boehringer Mannheim). Probe DNA (10–50 ng) was mixed with cot-1 (Life Technologies) (2–3 µg) and sufficient salmon sperm DNA to obtain a total of 10 µg in a 10 µl hybridization cocktail. After denaturation of the probe mixture (80–90°C for 5 min), pre-annealing of repetitive DNA sequences was carried out for 10 min in 37°C before application to denatured nucleic acid specimens. Following incubation overnight and subsequent post-hybridization washes, the specimens were treated with blocking solution (3% bovine serum albumin (BSA), 4× SSC) for 10 min at 37°C. All detection reagents were incubated with the specimen for 10–15 min at 37°C in 1% BSA, 4× SSC and 0.1% Tween 20 and slides were then washed at room temperature three times for 3 min each in 4× SSC and 0.1% Tween 20. Biotin-labeled probes were detected with rhodamine-conjugated avidin DCS (1:500 dilution) (Vector Laboratories) and digoxigenin-labeled probes were detected with an anti-digoxigenin antibody conjugated to FITC (Boehringer Mannheim) (1:100 dilution). BrdU was detected by anti-BrdU antibody (NeoMarkers) (1:100) followed by either rhodamine- (1:50) or AMCA- (1:20) conjugated anti-mouse antibody (Jackson Immunoresearch Laboratories). Counterstaining, where needed, was done using diamidinophenylindole (DAPI) (200 ng/ml) in Vector antifade solution. Amplification of the digoxigenin-labeled probes was carried out with anti-sheep antibody conjugated to FITC (Vector) and of the biotin-labeled probes with biotinized anti-avidin (Vector).

Replication timing was determined ($\pm 3\%$ with a 95% confidence interval) by counting the number of single and double dots in 100–200 nuclei per slide. These data were then normalized to other replication time markers. Replication analysis in human cells (Figure 1), for example, was normalized by adjusting the number of singles and doubles with reference to two fixed endogenous genes, SNRPN and CD3D, in lymphoblasts (69% singles and 52% doubles, respectively), fibroblasts (65%, 50%) and erythroblasts (72%, 50%). To normalize the data on transgene or integration site replication (Figures 2 and 4), we employed three different mouse-specific probes as controls, the mouse endogenous globin locus (MG), a cosmid probe from chromosome 6 (Chr6) and a random BAC probe. We analyzed fibroblast and embryonic liver cells from 13 different founder mice, and determined the average (Av) value (% singles) for each control probe. The specific data for mouse Ln2 are shown in Table I. The actual measurement for human globin in fibroblasts, for example, was 56% singles. However, each of the control probes varied (Var) from their average values by an amount of $(-5 - 3 - 1)/3 = -3$. For this reason, the replication time of human globin was adjusted to $56 + 3 = 59\%$. A similar procedure was used to normalize the value (32%) in embryonic liver. In this case, the average variation (Var) of the control probes was $(+4 + 5 - 1)/3 = +2.7$, so the corrected value for human globin replication

came to $32 - 2.7 = 29\%$. Each and every cell population was analyzed in this same manner, and the final normalized results for the transgenes are shown in graphic form in Figure 2C, and for the integration sites in Figure 4A.

Cells and transgenic mice

EBV-transformed lymphoblast cell lines were derived from normal individuals, and embryonic fibroblasts from amniotic fluid. Normal human peripheral blood-derived erythroid progenitors undergoing maturation into hemoglobin-containing cells (erythroblasts) were isolated and grown in culture as described (Fibach *et al.*, 1989).

The following transgenic mouse lines were used in this study. Ln2 (one copy), Ln72 (one copy) and Ln15 (1.5 copies) are transgenic lines made from a wild-type vector constructed from two cosmids spanning the entire globin locus (Strouboulis *et al.*, 1992). Δ2 and Δ4 are derivatives of these constructs containing small deletions in specific HSs (Milot *et al.*, 1996). Mouse line LnS2 (one copy) was made from a construct containing the wild-type LCR attached to the γ and β genes (Berry *et al.*, 1992). Mouse line μD was made from a construct that contains HSI-4 themselves (without the intervening sequences) attached to the β globin gene (Ellis *et al.*, 1996), and LCR3 (two copies), LCR4 (one copy) and LCR8 (two copies) were made using the 22 kb *SalI*–*Clai* fragment from a plasmid derivative (pTR-150) of cosmid HSI-V (Ryan *et al.*, 1989). Most of these lines have already been characterized. Thus, lines containing the full LCR element (Ln2, Ln15, Ln72 and LnS2), as well as lines Δ4-a (one copy), Δ4-b (two copies), Δ4-c (three copies) and μD-14 (one copy), have non-centromeric integration sites and all show normal developmentally regulated globin expression and DNase I sensitivity in erythroid cells. However, in lines Δ2-b (three copies) and Δ2-c (one copy), integration is near centromeric sequences, globin levels are reduced (to ~6% of normal) and the transgenes themselves are insensitive to DNase I (Milot *et al.*, 1996). When present in multiple copies, transgene organization is tandem.

Isolation of erythroid and non-erythroid cells from transgenic mice was carried out as described (Stanworth *et al.*, 1995). Briefly, transgenic males (homozygous or heterozygous) were mated with (C57B6 × CBA)F1 females and embryos were taken from pregnant females at 12.5 d.p.c. Peripheral blood cells were collected by allowing the embryos to bleed out into RPMI + 10% FCS containing 10 U/ml of preservative-free heparin. Fetal livers were dissected out and mechanically disrupted in RPMI + 0% FCS by passing them through a syringe. The remainder of the embryos were disrupted with a 2 ml syringe into DMEM + 10% FCS, in order to obtain embryonic fibroblasts. These cells were grown for 1–2 h. Purity of the erythroid cells from embryonic liver (85–90%) was determined by FACS analysis using a specific monoclonal antibody (Ikuta *et al.*, 1990) (TER-119, Pharmingen). Spleen lymphocytes from adult heterozygous transgenic mice were grown in culture and prepared for FISH analysis as described (Webb *et al.*, 1989).

For FISH analysis, BrdU (3×10^{-5} M) was added to all cultures 1 h before harvesting. Cells were then treated with hypotonic KCl solution (0.5%), fixed in methanol:acetic acid (3:1) and dropped onto slides (Selig *et al.*, 1992).

Probes

The following DNA probes were used to analyze the human globin region on chromosome 11 (their positions are shown in Figure 1A): PDJ895k23 (a), pDJ1075f20 (b), pDJ74k15 (c), pDJ192k15 (d), pDJ1173a5 (e), pDJ443n7 (i) and pDJ1112m17 (j) are PAC clones from the RPCI-1 and RPCI-5 human library of the Roswell Park Cancer Institute. Their locations were determined by The Genome Science and Technology Center in the University of Texas Southwestern Medical Center, and a map showing the locations of most of these probes appears on their web site (<http://gestec.swmed.edu/chromoso/htm>). The locations of the remaining probes appeared in a previous version of this map and were confirmed using FISH. Cos88 (f) and cos15 (h) are cosmids from the human globin region as indicated in Figure 1A; 1359 (g) is a plasmid (Talbot *et al.*, 1989) containing HSI-4 and the β globin gene (see Figure 2A). CosHG-28TK is a 38 kb cosmid covering the region that begins 4 kb 5' of the human Gγ gene and extends to 3 kb 3' to the β globin gene. The human SNRPN probe was purchased from Oncor, and CD3D was obtained from G.Evans. The mouse globin region was detected by a combination of two plasmids, β major (containing the mouse β major gene) and pβ12g (a 16 kb fragment containing the murine LCR) or pBSKs Sma#22 (a 16 kb fragment *SmaI* fragment containing ey and βH1). Chr6 is a random cosmid clone from mouse chromosome 6.

The transgene integration sites from mouse lines Ln2, Δ4-a and Δ4-c were isolated by inverse PCR (Ochman *et al.*, 1993). Total genomic DNA

Table I.

	Fibroblasts				Liver			
	HG	MG	CHR6	BAC	HG	MG	CHR6	BAC
Av		50	61	39		33	65	41
Ln2	56	45	58	38	32	37	69	40
Var.		-5	-3	-1		+4	+5	-1
Final	59				29			

was digested with the restriction enzyme *MboI* or *MspI* and self-ligated with T4 DNA ligase. Nested PCR was carried out using primers complementary to DNA at the 3' end of the transgene in an attempt to amplify the flanking sequence (5'-ATGTTAAATTAATACCACTC-3' and 5'-ATGTATACCTTGTAAGATGA-3' for the first PCR and 5'-AAGCTAATTAACATACCCAT-3' and 5'-TGTGTAAGTAAGATA-GTGA-3' for the second PCR). PCR products were then cloned in a pGEM-T vector and sequenced. New sets of PCR primers were used for screening and isolating BAC clones 212a06 (5'-AGAGCTTCC-AGGCTCATGCCA-3' and 5'-ACCTTCCTCGACATTTTCA-3') for L_n2, 2s172 (5'-GTGCTGAGAGTGTCTATTGA-3' and 5'-GTGACA-GCACTCCACAGACC-3') for Δ4-c and 186n09 or 112g01 (5'-TAG-ATCAGCTGATCTTAACG-3' and 5'-AAAAGTGGACACTAATA-CCG-3') for Δ4-a from BAC ES mouse DNA libraries (releases I and 2 from Genome System Inc). FISH analysis confirmed that these clones indeed represent globin integration sites in the three mice (see Figure 4B).

Acknowledgements

We would like to thank G.Evans, D.Ward and M.Groudine, who kindly provided probes that were used for FISH analysis, T.M.Townes for the LCR plasmid pTR-150, T.Jakubowicz for help in preparing the manuscript and Pnina Ever Hadani for the statistical analysis. This work was supported by grants from the Israel Academy of Sciences (H.C.), the NIH (H.C.) and the Israel Cancer Research Fund (H.C.).

References

Aladjem,M.I., Groudine,M., Brody,L.L., Dieken,E.S., Fournier,R.E.K., Wahl,G.M. and Epner,E.M. (1995) Participation of human β-globin locus control region in initiation of DNA replication. *Science*, **270**, 815–819.

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

Allshire,R. and Bickmore,W. (2000) Pausing for thought on the boundaries of imprinting. *Cell*, **102**, 705–708.

Berry,M., Grosveld,F. and Dillon,N. (1992) A single point mutation is the cause of the Greek form of hereditary persistence of fetal haemoglobin. *Nature*, **358**, 499–502.

Bulger,M., von Doorninck,J.H., Saitoh,N., Telling,A., Farrell,C., Bender,M.A., Felsenfeld,G., Axel,R. and Groudine,M. (1999) Conservation of sequence and structure flanking the mouse and human β-globin loci: The β-globin genes are embedded within an array of odorant receptor genes. *Proc. Natl Acad. Sci. USA*, **96**, 5129–5134.

Chess,A., Simon,I., Cedar,H. and Axel,R. (1994) Allelic inactivation regulates olfactory receptor gene expression. *Cell*, **78**, 823–834.

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 β-globin locus are independent of the locus control region. *Mol. Cell. Biol.*, **20**, 5581–5591.

Dhar,V., Skoultschi,A.I. and Schildkraut,C.L. (1989) Activation and repression of a β-globin gene in cell hybrids is accompanied by a shift in its temporal regulation. *Mol. Cell. Biol.*, **9**, 3524–3532.

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

Ellis,J., Tan-Un,K.C., Harper,A., Michalovich,D., Yannoutsos,N., Philipsen,S. and Grosveld,F. (1996) A dominant chromatin-opening activity in 5' hypersensitive site 3 of the human β-globin locus control region. *EMBO J.*, **15**, 562–568.

Epner,E., Forrester,W.C. and Groudine,M. (1988) Asynchronous DNA replication within the human β-globin gene locus. *Proc. Natl Acad. Sci. USA*, **85**, 8081–8085.

Feinberg,A.P. (1999) Imprinting of a genomic domain of 11p15 and loss of imprinting in cancer: an introduction. *Cancer Res.*, **59**, Suppl. 7, 1743S–1746S.

Ferguson,B.M. and Fangman,W.L. (1992) A position effect on the time of replication origin activation in yeast. *Cell*, **68**, 333–339.

Fibach,E., Manor,D., Oppenheim,A. and Rachmilewitz,E.A. (1989) Proliferation and maturation of human erythroid progenitors in liquid culture. *Blood*, **73**, 100–103.

Forrester,W.C., Epner,E., Driscoll,M.C., Enver,T., Brice,M.,

Papayannopoulou,T. and Groudine,M. (1990) A deletion of the human β-globin locus activation region causes a major alteration in chromatin structure and replication across the entire β-globin locus. *Genes Dev.*, **4**, 1637–1649.

Friedman,K.L., Diller,J.D., Ferguson,B.M., Nyland,S.V.M., Brewer,B.J. and Fangman,W.L. (1996) Multiple determinants controlling activation of yeast replication origins late in S phase. *Genes Dev.*, **10**, 1595–1607.

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

Holmquist,G.P. (1987) Role of replication time in the control of tissue specific gene expression. *Am. J. Hum. Genet.*, **40**, 151–173.

Ikuta,K., Kina,T., MacNeil,L., Uchida,N., Peault,B., Chien,Y. and Weissman,I.L. (1990) A developmental switch in thymic lymphocyte maturation potential occurs at the level of hematopoietic stem cells. *Cell*, **62**, 863–874.

Kerem,B.S., Goitein,R., Diamond,G., Cedar,H. and Marcus,M. (1984) Mapping of DNase-I sensitive regions on mitotic chromosomes. *Cell*, **38**, 493–499.

Kitsberg,D., Selig,S., Brandeis,M., Simon,I., Keshet,I., Driscoll,D.J., Nicholls,R.D. and Cedar,H. (1993a) Allele-specific replication timing of imprinted gene regions. *Nature*, **364**, 459–463.

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

Lichter,P., Cremer,T., Borden,J., Manuelidis,L. and Ward,D.C. (1988) Delineation of individual human chromosomes in metaphase and interphase cells by *in situ* suppression hybridization using recombinant DNA libraries. *Hum. Genet.*, **80**, 224–234.

Lichter,P., Tang,C.C., Call,K., Hermanson,G., Evans,G.A., Housman,D. and Ward,D.C. (1990) High resolution mapping of human chromosome 11 by *in situ* hybridization with cosmid clones. *Science*, **247**, 64–69.

Milot,E. et al. (1996) Heterochromatin effects on the frequency and duration of LCR-mediated gene transcription. *Cell*, **87**, 105–114.

Mostoslavsky,R., Kirillov,A., Ji,Y.-H., Goldmit,M., Holzman,M., Wirth,T., Cedar,H. and Bergman,Y. (1999) Demethylation and the establishment of κ allelic exclusion. *Cold Spring Harb. Symp. Quant. Biol.*, **64**, 197–206.

Ochman,H., Ayala,F.J. and Hartl,D.L. (1993) Use of polymerase chain reaction to amplify segments outside boundaries of known sequences. *Methods Enzymol.*, **218**, 309–321.

Rountree,M.R., Bachman,K.E. and Baylin,S.B. (2000) DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. *Nature Genet.*, **25**, 269–277.

Ryan,T.M., Behringer,R.R., Martin,N.C., Townes,T.M., Palmiter,R.D. and Brinster,R.L. (1989) A single erythroid-specific DNase I super-hypersensitive site activates high levels of human β-globin gene expression in transgenic mice. *Genes Dev.*, **3**, 314–323.

Selig,S., Okumura,K., Ward,D.C. and Cedar,H. (1992) Delineation of DNA replication time zones by fluorescence *in situ* hybridization. *EMBO J.*, **11**, 1217–1225.

Simon,I. and Cedar,H. (1996) Temporal order of DNA replication. In DePamphilis,M.L. (ed.), *DNA Replication in Eukaryotic Cells*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 387–408.

Simon,I., Tenzen,T., Reubinoff,B.E., Hillman,D., McCarrey,J.R. and Cedar,H. (1999) Asynchronous replication of imprinted genes is established in the gametes and maintained during development. *Nature*, **401**, 929–932.

Stanworth,S.J., Robert,N.A., Sharpe,J.A., Sloan-Stanley,J.A. and Wood,W.G. (1995) Established epigenetic modifications determine the expression of developmentally regulated globin genes in somatic cell hybrids. *Mol. Cell. Biol.*, **15**, 3969–3978.

Stillman,B. (1993) Replicator renaissance. *Nature*, **366**, 506–507.

Strouboulis,J., Dillon,N. and Grosveld,F. (1992) Developmental regulation of a complete 70 kb human β-globin locus in transgenic mice. *Genes Dev.*, **6**, 1857–1864.

Talbot,D., Collis,P., Antoniou,M., Vidal,M., Grosveld,R. and Greaves,D.R. (1989) A dominant control region from the human β-globin locus conferring integration site-independent gene expression. *Nature*, **338**, 352–355.

Webb,C.F., Das,C., Coffman,R.L. and Tucker,P.W. (1989) Induction of immunoglobulin mRNA in a B cell transfectant stimulated with interleukin-5 and a T-dependent antigen. *J. Immunol.*, **143**, 3934–3939.

Received May 10, 2001; revised and accepted August 29, 2001