MOLECULAR AND CELLULAR BIOLOGY, May 1995, p. 2482–2489 Vol. 15, No. 5
0270-7306/95/$04.00+0

Mapping of Replication Initiation Sites in Human Ribosomal DNA by Nascent-Strand Abundance Analysis

YEUP YOON,† J. AQUILES SANCHEZ,‡ CHRISTINE BRUN, AND JOEL A. HUBERMAN*

Department of Molecular and Cellular Biology, Roswell Park Cancer Institute, Buffalo, New York 14263

Received 12 October 1994/Returned for modification 1 December 1994/Accepted 3 February 1995

New techniques for mapping mammalian DNA replication origins are needed. We have modified the existing nascent-strand size analysis technique (L. Vassilev and E. M. Johnson, Nucleic Acids Res. 17:7693–7705, 1989) to provide an independent means of studying replication initiation sites. We call the new method nascent-strand abundance analysis. We confirmed the validity of this method with replicating simian virus 40 DNA as a model. We then applied nascent-strand abundance and nascent-strand size analyses to mapping of initiation sites in human (HeLa) ribosomal DNA (rDNA), a region previously examined exclusively by two-dimensional gel electrophoresis methods (R. D. Little, T. H. K. Platt, and C. L. Schildkraut, Mol. Cell. Biol. 13:6600–6613, 1993). Our results partly confirm those obtained by two-dimensional gel electrophoresis techniques. Both studies suggest that replication initiates at relatively high frequency a few kilobase pairs upstream of the transcribed region and that many additional low-frequency initiation sites are distributed through most of the remainder of the ribosomal DNA repeat unit.

Although DNA replication origins in the budding yeast *Saccharomyces cerevisiae* have been moderately well characterized (reviewed in references 14 and 20), those in mammalian cells are not yet well understood (reviewed in reference 10). Understanding mammalian replication origins will require identification of both the "replicators" (*cis*-acting sequences required for origin function [31]) and the "initiation sites" (sequences which serve as templates for the 5′ ends of the first nascent strands) associated with each origin.

Although it would be desirable to map initiation sites with single-nucleotide resolution, all of the current techniques—in vitro runoff (3, 32), determination of replication timing (15, 22, 35), two-dimensional (2D) gel electrophoresis of replicating restriction fragments (37, 47), measurement of leading-strand polarity (12, 24), measurement of lagging-strand polarity (11), and measurement of nascent-strand size (45)—offer precisions of several hundred nucleotides at best. Furthermore, the current techniques sometimes produce apparently contradictory observations. For example, the results obtained by most techniques are consistent with the presence of two discrete initiation sites downstream of the dihydrofolate reductase (DHFR) gene in Chinese hamster ovary (CHO) cells (11, 24, 35, 46). In contrast, the results obtained by 2D gel electrophoresis methods argue in favor of a single broad (55-kbp) "initiation zone" downstream of the DHFR gene, within which replication initiates at so many locations that they cannot be resolved by the 2D gel electrophoresis techniques (18, 47). It should be noted that the 2D gel electrophoresis methods are capable of localizing the discrete initiation sites of *S. cerevisiae* with a precision of several hundred base pairs (7, 28).

It is possible that these apparently contradictory observations represent different aspects of a potentially complex process of initiation at mammalian replication origins. If so, the development of new approaches capable of providing new information about the process of initiation is likely to contribute to improved understanding of both initiation and origins. For this reason, we have modified the nascent-strand size analysis method (45) to provide information about nascent-strand abundance, and we have used the modified technique to study initiation of replication within the DNA encoding rRNA (rDNA) of human cells.

In all eukaryotic organisms, the genes encoding rRNA are naturally repeated. In human cells, there are ~400 tandem iterations per genome of a 44-kbp repeat unit (see Fig. 4). These are distributed in five clusters on chromosomes 13, 14, 15, 21, and 22. Within each repeat unit, a 13-kb primary transcript is processed to yield the mature 18S, 5.8S, and 28S rRNAs. Transcriptional control elements, the promoter and sequences required for transcriptional termination, are located in the remaining 31 kbp of nontranscribed spacer (NTS) near the 5′ and 3′ ends, respectively, of the transcribed region (reviewed in reference 33). In several eukaryotic organisms—*S. cerevisiae* (8, 36), *Tetrahymena thermophila* (44), *Physarum polycephalum* (49), pea (26), and sea urchin (5)—discrete replication origins have been detected within the NTS upstream of the transcribed region. However, results obtained with vertebrates are less clear. One group has reported a discrete replication origin in the NTS of rat rDNA (1). Other results suggest a more widely dispersed distribution of initiation sites in vertebrate rDNA. 2D gel electrophoresis results obtained with very early embryos of *Xenopus laevis* suggest that replication initiates at sites distributed throughout the entire repeat unit, including the transcribed region (29). In contrast, results obtained by both analysis of replication timing and electron microscopy suggest that initiation in more highly differentiated *Xenopus* cells is confined primarily to the NTS (6). Similarly, 2D gel electrophoresis results indicate that although replication in cultured human cells initiates in a broad zone, this zone is confined to most of the NTS; initiation is not detected in the promoter, the transcribed region, or the transcription termination region (37).

So far, replication of human rDNA has been studied only by 2D gel electrophoresis methods (37). However, as discussed

* Corresponding author. Mailing address: Department of Molecular and Cellular Biology, Roswell Park Cancer Institute, Elm and Carlton Sts., Buffalo, NY 14263-0001. Phone: (716) 845-3047. Fax: (716) 845-8126. Electronic mail address: huberman@acsu.buffalo.edu.

† Present address: 9-1102, Jangmi Apt, Shicheon dong, Songpa Gu, Seoul, Republic of Korea.

‡ Present address: Department of Biology, Brandeis University, Waltham, MA 02254.

above, studies of the DHFR origin region by other techniques have yielded results which appear to conflict with the 2D gel electrophoresis results. Consequently, we thought that it would be informative to apply our new method, nascent-strand abundance analysis, to human rDNA. Would this technique provide results confirming the 2D gel electrophoresis results, or would it produce apparently contradictory results, as with the DHFR origin? As described in this paper, the results obtained by nascent-strand abundance analysis partly confirmed the conclusions arrived at by 2D gel electrophoresis analysis. Both sets of results suggest that preferred initiation sites are located a few kilobase pairs upstream of the transcribed region and are consistent with the existence of multiple initiation sites throughout the NTS.

MATERIALS AND METHODS

Cell culture and SV40 infection. Monkey CV-1 cells were grown as monolayer cultures (15-cm plates) in Dulbecco's modified Eagle's medium supplemented with 2% fetal calf serum plus 8% calf serum (all from Life Technologies). At 70% confluence, the CV-1 cells were infected with simian virus 40 (SV40). HeLa cells were grown in 300-ml suspension cultures in Joklik-modified minimal essential medium (Life Technologies) supplemented with 10% horse serum.

Labeling and isolation of nascent DNA. When the CV-1 cultures were infected with SV40 or when the HeLa cell cultures reached a density of 2×10^5 cells per ml, their bulk DNA was labeled with [^{14}C]deoxythymidine (dThd; NEN Dupont; 53.8 mCi/mmol; 25 nCi/ml) for 24 h (HeLa cells) or 36 h (SV40-infected CV-1 cells). The cells were then pulse-labeled for 10 min (SV40-infected CV-1 cells) or 30 min (HeLa cells) by adding [^{3}H]deoxycytidine (dCyd; NEN Dupont; 26.5 Ci/mmol) and bromodeoxyuridine (BrdUrd) to final concentrations of 1.0 μCi/ml and 15 μg/ml, respectively.

Labeling of HeLa cells was terminated by adding 2 to 3 volumes of ice-cold PBS (137 mM NaCl, 2.7 mM KCl, 4.3 mM $Na_2HPO_4 \cdot 7\ H_2O$, 1.4 mM KH_2PO_4 [pH 7.3]). HeLa cells were harvested by centrifugation ($650 \times g$ for 10 min at 4°C). After an additional wash in ice-cold PBS, the cells from 300 ml of culture were resuspended in 10 ml of TET buffer (50 mM Tris [pH 7.5], 1 mM EDTA, 0.04% Triton X-100) and broken to release nuclei by rapid passage three times through a 22-gauge needle. After dilution with the same volume of TE buffer (50 mM Tris [pH 7.5], 1 mM EDTA), the nuclei were harvested by centrifugation at $500 \times g$ for 5 min at 4°C. This method may require adjustment for cell lines other than HeLa and CV-1. Any method that yields stable, easily resuspendable nuclei with intact DNA should be satisfactory. The purified intact nuclei were resuspended in 1 ml of TE buffer and then dissolved and denatured by addition of 1 ml of alkaline denaturation buffer (1 M NaCl, 0.5 M NaOH, 50 mM EDTA, 0.4% Sarkosyl NL-30) followed by incubation at 45°C for 30 min. Labeling of CV-1 cells was terminated by washing twice with ice-cold PBS (30 ml per wash per plate). The cells were subsequently incubated with 10 ml of TET buffer for 3 min and scraped from the plate. The cells harvested from six plates were handled in the same way as for HeLa cells. The resulting solution of denatured nuclear DNA was gently poured on top of a 32-ml alkaline sucrose gradient (linear, 5 to 20% sucrose in 1/2× alkaline denaturation buffer). The DNA strands were then separated according to size by centrifugation in a Beckman SW28 rotor (19,000 rpm for 14 h at 21°C).

Fractions (1.8 ml) were collected from the gradient, and the size distribution of DNA strands in each fraction was determined by alkaline agarose gel electrophoresis of an aliquot from each fraction, followed by neutralization and ethidium bromide staining. Fractions from the upper portion of the sucrose gradient containing strands ranging from 0 to ~25 kb were pooled, cooled on ice, and neutralized with 40 μl of 21% phosphoric acid per ml of collected fractions until a pH of 7 was reached. CsCl was then added directly to the pooled fractions (1.32 g/ml of initial volume), and additional CsCl or water was added as necessary to achieve a final refractive index of 1.4155. Isopycnic centrifugation was carried out in a Beckman VTi50 rotor (40,000 rpm for 20 h at 21°C). The fractions containing dense, BrdUrd-, [^{3}H]dCyd-labeled DNA were located on the basis of their ^{3}H radioactivity, pooled, redenatured by addition of NaOH to 0.2 M for 20 min (room temperature), and then reneutralized with phosphoric acid (as above). Their density was readjusted until the refractive index reached 1.4155, and a second round of isopycnic centrifugation was carried out under the same conditions as before. The fractions containing DNA strands fully substituted with BrdUrd were pooled as before and precipitated by addition of 3 volumes of 70% ethanol in the presence of glycogen (50 μg/ml) as carrier.

Abundance analysis of the nascent DNA. The purified fully substituted nascent DNA was dissolved in alkaline electrophoresis buffer (40 mM NaOH, 2 mM EDTA) and subjected to alkaline agarose gel electrophoresis (0.8% agarose) at 17 V for 14 h. Genomic DNA digested with appropriate restriction enzymes was run in a parallel lane as a control. After electrophoresis, the gel was incubated briefly (10 min) in 0.25 M HCl and then blotted to a nylon membrane (GeneScreen; NEN Dupont) with an alkaline transfer buffer (0.4 M NaOH, 1 M NaCl). The transferred DNA was cross-linked to the membrane by exposure to 254-nm UV light from a transilluminator for 10 s. The probes described in the next paragraph were labeled with ^{32}P-deoxynucleoside triphosphates by the random-oligonucleotide-primed labeling procedure (21). The membrane was hybridized as recommended by the manufacturer in a buffer containing 10% dextran sulfate and 50% formamide. After being washed and subjected to autoradiographic exposure (Kodak XAR film), the membranes were stripped with 0.2 M NaOH for 1 h at room temperature and then hybridized with the next probe. This process was continued until the same membrane had been hybridized with all probes of interest. It was then rehybridized with the first probe to check for possible nonuniform loss of DNA from the membrane. No such loss was detected in the experiments reported here. However, in other experiments in which the probe was removed with boiling 2 mM EDTA–1% sodium dodecyl sulfate, selective loss of BrdUrd-labeled DNA from the membrane was observed. Autoradiograms were scanned with a Molecular Dynamics 300A densitometer. For abundance measurements, the ratio of integrated signal for the entire nascent DNA lane to the integrated signal for the entire control DNA lane was determined. These ratios were then normalized to the ratio for the probe yielding the highest absolute ratio (probe 1 for SV40; probe 7 for HeLa cells).

Hybridization probes. The probes used for analysis of SV40 replication were as follows: probe 1, a 345-bp *Bgl*I-*Hha*I fragment (nucleotides 5242 to 343); probe 2, a 490-bp *Hha*I fragment (nucleotides 344 to 833); probe 3, a 259-bp *Hin*dIII-*Hpa*I fragment (nucleotides 3477 to 3735); and probe 4, a 675-bp *Bam*HI-*Pst*I fragment (nucleotides 2534 to 3208).

The probes used for analysis of human rDNA were all provided by James Sylvester, Hahnemann University, Philadelphia, Pa. Their positions are indicated on the maps in Fig. 4 and 5. They are as follows: probe 1, a 1.2-kbp *Eco*RI-*Sal*I fragment (also known as probe B_{ES} [37]); probe 2, a 2.2-kbp *Sal*I fragment; probe 3, a 0.96-kbp *Eco*RI-*Sal*I fragment (also known as probe D_{ES} [37]); probe 4, a 0.3-kbp *Xba*I fragment (also known as probe D_{XX} [37]); probe 5, a 1-kbp *Sal*I-*Hin*dIII fragment (also known as probe D_{SH} [41]); probe 6, a 0.4-kbp *Eco*RI-*Bam*HI fragment (also known as probe C_{EB} [37]); and probe 7, a 0.5-kbp *Hin*dIII-*Bam*HI fragment (also known as probe C_{HB} [37]).

RESULTS

Experimental rationale. The diagram in Fig. 1A illustrates the basis for the nascent-strand abundance technique. Unsynchronized growing cells contain nascent strands ranging from a few nucleotides to the size of the chromosome. The smaller nascent strands form pyramidal distributions centered on replication origins (Fig. 1A, top). When the cells are incubated with the deoxyribonucleoside BrdUrd, segments of BrdUrd-labeled DNA are added on to the ends of the preexisting nascent strands, and the nascent strands resulting from initiations during the incubation period become fully labeled with BrdUrd (Fig. 1A, middle). The fully labeled strands can be purified by selection of shorter nascent strands by using alkaline sucrose gradient sedimentation followed by selection of dense strands by using CsCl density gradient fractionation. The maximal size of the desired short strands can be inferred from the length of the BrdUrd pulse and the rate of replication fork movement. The purified fully labeled strands should form origin-centered pyramidal size distributions, and hybridization probes located close to origins should detect a greater abundance of these strands than should origin-distal probes (Fig. 1A, bottom).

The nascent-strand abundance technique is related to the widely used method of nascent-strand size analysis (2, 4, 27, 30, 38, 40, 42, 43, 45, 48, 50). Both techniques require purification of short, BrdUrd-labeled nascent strands. The goal of nascent-strand size analysis is to identify the probe from the region of interest which detects the shortest nascent strands, because the shortest nascent strands should come from origins (Fig. 1A). Although the total pyramid of nascent strands detectable by an origin-proximal probe may be relatively abundant, the shortest detectable nascent strands are rare. Thus, one potential advantage of the nascent-strand abundance technique is that it may not require detection methods as sensitive as those needed for nascent-strand size analysis.

The methods of nascent-strand size analysis and nascent-strand abundance analysis can be combined to provide addi-

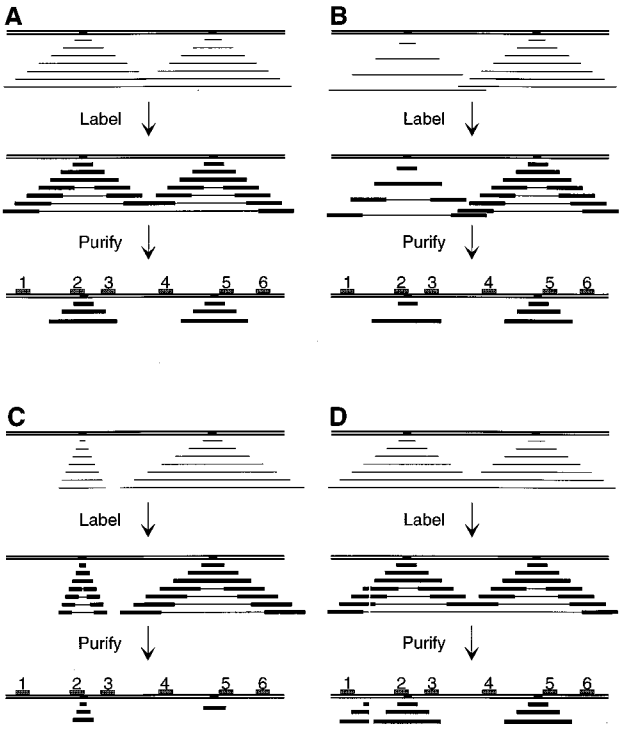


FIG. 1. Application of the nascent-strand abundance analysis technique to discrete origins and to initiation zones. The medium-thickness horizontal parallel lines represent the parental strands of a DNA molecule. The small solid boxes between these lines represent replication origins or initiation sites. The thin horizontal lines represent unlabeled (light) nascent strands, while the very thick horizontal lines represent BrdUrd-labeled (dense) nascent strands. The numbered boxes at the bottom represent hybridization probes. The pyramidal distributions of nascent strands represent the size range and abundance range of nascent strands which would be present in an unsynchronized cell population before (thin lines, top of each diagram) and after (thick and thin lines, middle of each diagram) pulse-labeling with BrdUrd. "Label" implies pulse-labeling with BrdUrd. "Purify" implies a two-step purification based on selection of short strands in an alkaline sucrose gradient followed by selection of fully BrdUrd-substituted strands by isopycnic centrifugation. (A) Two discrete origins which fire in 100% of S phases. Converging replication forks meet between the two origins, and nascent strands are ligated. (B) Two initiation sites which fire infrequently. The left-hand site fires at half the frequency of the right-hand site. Because of the low firing frequency, the two sites rarely fire simultaneously in the same DNA molecule. Therefore, in most cases, replication forks do not meet between the two origins. (C) Two initiation sites which fire at equal frequency generating replication forks which travel at different rates. Forks produced at the left-hand site move one-fourth as fast as those produced at the right-hand site. (D) Similar to panel A except for the presence of a site-specific nick in the nascent strands generated by forks moving leftward from the left-hand origin.

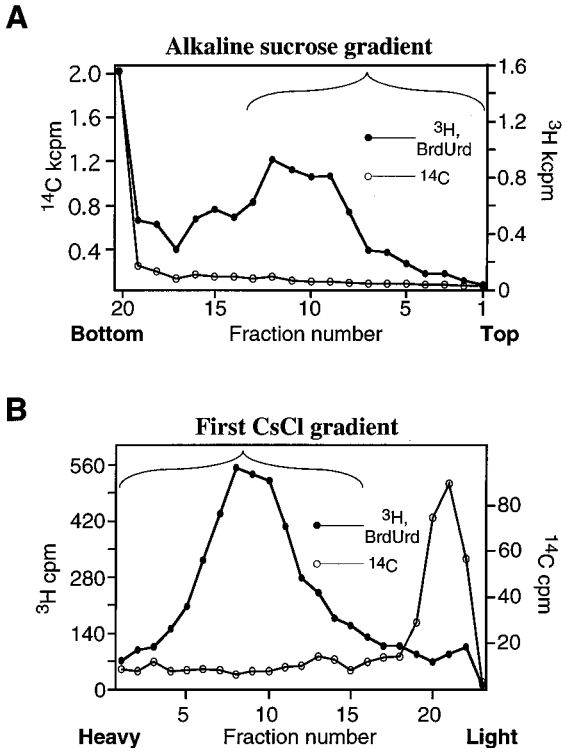


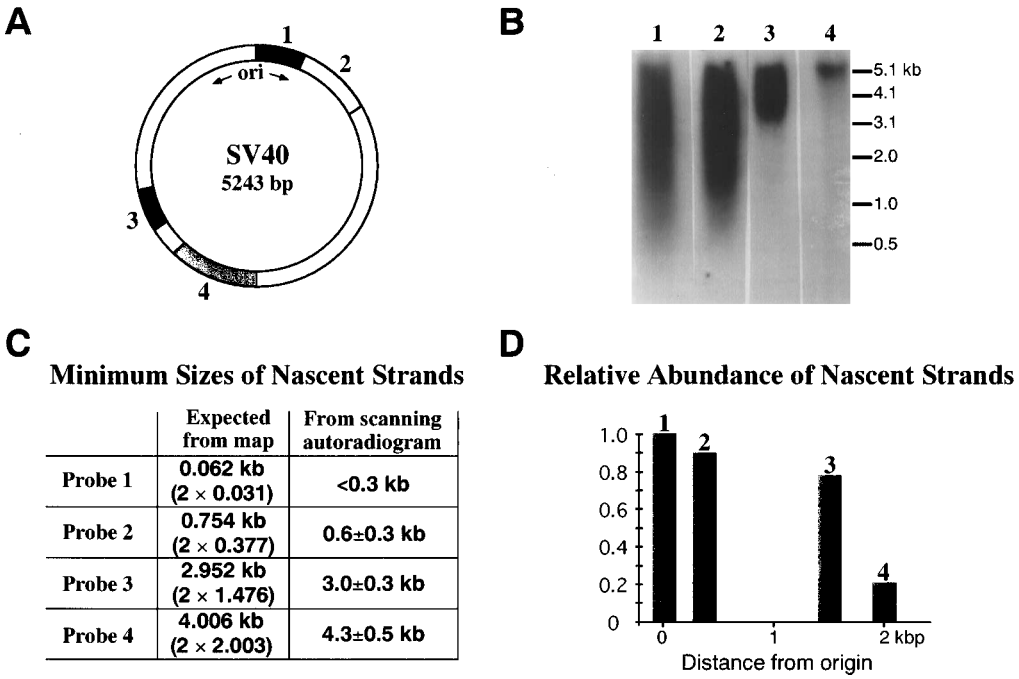
FIG. 2. Fractionation of nascent DNA strands according to size and density. The braces indicate the regions which were pooled for further analysis. (A) Alkaline sucrose gradient sedimentation of the denatured DNA from pulse-labeled HeLa cell nuclei. Fractions were collected from the top. (B) First isopycnic CsCl gradient separation of HeLa nascent strands from contaminating parental and longer progeny strands.

tional information. If the purified, fully labeled nascent strands (prepared as described above) are size fractionated by alkaline agarose gel electrophoresis and then blotted to a membrane, the hybridization probes that detect the most abundant nascent strands and the shortest nascent strands can be simultaneously identified.

Experimental conditions. In a typical experiment, asynchronously growing HeLa or SV40-infected CV-1 cells were incubated for 24 h (about one generation time) with [^{14}C]dThd to label the parental and older progeny DNA and then pulse-labeled for 30 min (HeLa cells) or 10 min (CV-1 cells) with [^{3}H]dCyd and BrdUrd to label the most recently synthesized (nascent) DNA. Following gentle alkaline lysis of intact nuclei, sedimentation through an alkaline sucrose gradient was used to separate the shorter, fully substituted nascent DNA from both parental DNA and the longer, partially substituted nascent DNA. Assuming a rate of replication fork movement of 0.6 to 3 kbp/min and a bidirectional replication process (19), fully labeled nascent DNA should range in size from a few nucleotides to 60–180 kb for a 30-min pulse. Parental and partially substituted strands should be still larger.

Figure 2A displays a typical distribution of [^{14}C]dThd-labeled parental and older progeny DNA (open circles) and [^{3}H]dCyd- and BrdUrd-labeled nascent DNA (solid circles) after alkaline sucrose gradient sedimentation. The ^{14}C profile indicates that a large proportion of the parental and older progeny DNA was very large and sedimented to the bottom of the gradient. As expected, a substantial portion of the pulse-labeled nascent DNA sedimented in the upper portions of the gradient. Next, we used alkaline agarose gel electrophoresis to measure the size distribution of the bulk DNA in each fraction of the gradient. On the basis of these measurements, we pooled fractions containing strands in the size range 0 to ~25 kb (HeLa cells) or 0 to ~10 kb (CV-1 cells). The choice of these size ranges ensured minimal contamination of the fully substituted strands with partially substituted, older progeny, or parental strands.

The fully substituted nascent DNA was then further purified by two consecutive rounds of isopycnic centrifugation in neutral CsCl gradients. A typical example of the separation achieved after just a single round is presented in Fig. 2B. The heavy peak (indicated by a brace) contained no detectable ^{14}C above background. Consequently, after the second round, there was no ^{14}C detectable anywhere in the gradient, even in the position expected for light DNA (51). The nascent DNA strands were recovered from the second CsCl gradient and



	Expected from map	From scanning autoradiogram
Probe 1	0.062 kb (2 × 0.031)	<0.3 kb
Probe 2	0.754 kb (2 × 0.377)	0.6±0.3 kb
Probe 3	2.952 kb (2 × 1.476)	3.0±0.3 kb
Probe 4	4.006 kb (2 × 2.003)	4.3±0.5 kb

FIG. 3. Nascent-strand size and abundance analysis of SV40 replication. (A) Diagram of the relative positions of the probes used to detect nascent strands. (B) Autoradiograms showing the size distributions of the nascent strands detected by each of the indicated probes. The same membrane was hybridized sequentially with each of the probes. (C) Comparison of the minimum nascent-strand sizes detected (Fig. 3B) with those predicted from the known origin location, 5210 to 5211 (25). The origin-proximal endpoint of each probe was used in making these calculations. (D) Relative abundance of nascent strands detected with each of the indicated probes, plotted as a function of distance from the known origin.

separated according to size by alkaline agarose gel electrophoresis before being blotted to a nylon membrane.

Analysis of nascent SV40 DNA. To test the validity of this experimental approach, we used SV40 as a model system. The SV40 origin of replication has been mapped with single-nucleotide resolution (25).

The nylon membrane containing the purified, size-fractionated, fully substituted nascent SV40 DNA strands was hybridized sequentially with probes derived from different locations in the SV40 DNA molecule (Fig. 3A). The resulting autoradiograms are shown in Fig. 3B. The probe closest to the replication origin (probe 1) detected nascent DNA from ~300 bases to 5.2 kb (the size of strands in mature SV40 DNA). The shortest strand sizes detected by the other probes increased in proportion to their distance from the origin (Fig. 3B; tabulated in Fig. 3C). If we had not previously known the location of the SV40 replication origin, it would have been possible to locate this origin from the results in Fig. 3B and C.

We then measured the relative abundances of the nascent strands detected by the various probes in Fig. 3. This was done by scanning two lanes in each of the autoradiograms which contributed to Fig. 3B—the lane shown in Fig. 3B and a control lane containing intact SV40 molecules. The normalized ratios of the experimental and control signals are displayed in Fig. 3D as a function of probe distance from the known origin. According to the rationale linking nascent-strand abundance with origin proximity (Fig. 1), it is possible to conclude from the data in Fig. 3D that the SV40 origin must be located closer to probe 1 than to any of the other tested probes. However, the abundance of strands detected by probe 3 is greater than expected based on the distance of probe 3 from the known origin. This departure from expected behavior could be a consequence of a nonuniform rate of replication fork movement, as explained in Discussion. Despite this deviation, analysis of nascent-strand abundance proved capable of correctly identifying the known SV40 replication origin. We concluded that this approach ought to yield valid results when applied to mammalian chromosomal DNA.

Analysis of nascent human rDNA. Since human ribosomal repeat units are 44 kbp, we chose to work with nascent strands smaller than ~25 kb. Such small nascent DNAs should arise only from repeat units containing active replication origins.

The results of one of our analyses are displayed in Fig. 4. The right lane in each panel (E, experimental) shows the size distribution of nascent DNA detected by the indicated probe. The left lane (C, control) displays the restriction fragments in total *Eco*RI-*Hin*dIII-digested genomic DNA detected by the same probe. This control lane was used to monitor the specificity and efficiency of hybridization with each probe. The diagram at the bottom of the figure illustrates the positions of the different probes within the human rDNA repeat. Probes 1, 2, and 7 detected two bands because of heterogeneity in the *Eco*RI site upstream of the transcription start site (17, 34) marked with a solid circle in the diagram. Probe 4 efficiently cross-hybridizes to two nearby sequences (39), which are indicated by open boxes.

To determine the abundances of nascent strands, we used a densitometer to scan both the nascent-strand signal and the control signal detected by each probe (Fig. 4). Then we determined the ratio of each nascent-strand signal to the corresponding control signal and normalized these ratios to the maximum ratio (the ratio of nascent-strand signal to control signal for probe 7). The results of the nascent-strand abundance measurements (Fig. 5) suggest that initiation occurs more frequently near probe 7 than near any of the other tested probes.

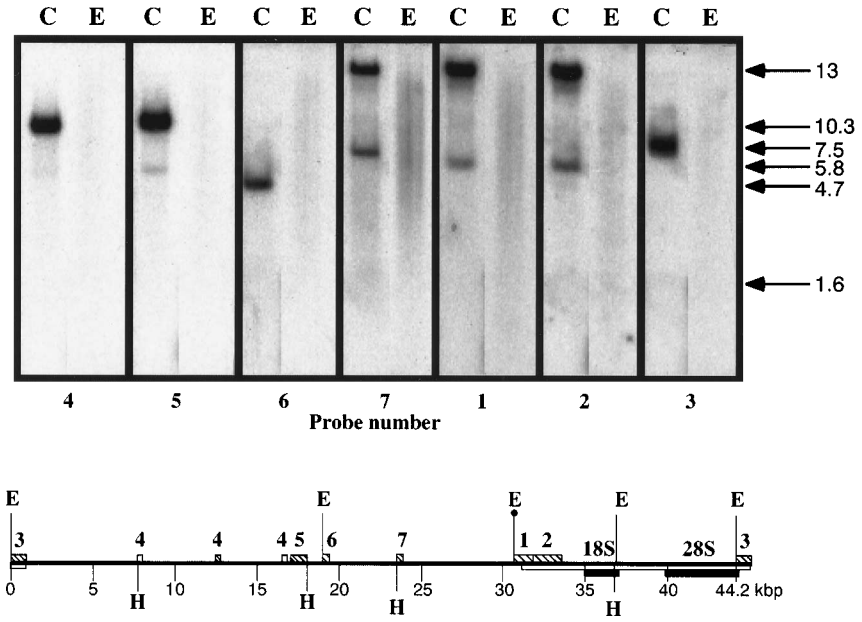


FIG. 4. Nascent-strand size and abundance analysis of human rDNA. In each autoradiogram frame, the left lane (C, control) contains total HeLa cell DNA digested with *Eco*RI and *Hin*dIII and the right lane (E, experimental) contains purified fully substituted nascent DNA. The same membrane was probed sequentially with each of the indicated probes. The sizes (in kilobases) indicated at the right are the sizes of the indicated genomic restriction fragments or of fragments of the "1-kb marker" (Life Technologies) which was included in the same gel (51). In the map below, the solid horizontal line indicates a stretch of human rDNA. Restriction sites: E, *Eco*RI; H, *Hin*dIII. The *Eco*RI site indicated by a circle is not present in all rDNA repeats. Distance in kilobase pairs is shown by the scale below the DNA line. Probes are indicated by hatched boxes. The two extra open boxes indicated for probe 4 represent regions that efficiently cross-hybridize with this probe (39). The transcribed region is indicated by a light-gray box, and the sequences encoding the mature 18S and 28S rRNAs are indicated by darker-gray boxes.

DISCUSSION

Interpretation of the HeLa rDNA results. The relatively high abundance of nascent strands detected by probe 7 (Fig. 5) could be explained in two ways. The fact that nascent-strand abundance declines with distance from probe 7 in both directions suggests that there may be a single major bidirectional initiation site near probe 7, analogous to the situation depicted in Fig. 1A. However, the simplified example in Fig. 1B shows that nascent-strand abundance can be affected by the frequency of initiation site usage as well as by initiation site proximity. Consequently, the abundance results in Fig. 5 are also consistent with the possibility that within the population of asynchronously growing cells, there are multiple initiation sites in the rDNA repeat, with those near probe 7 firing more frequently than those farther away.

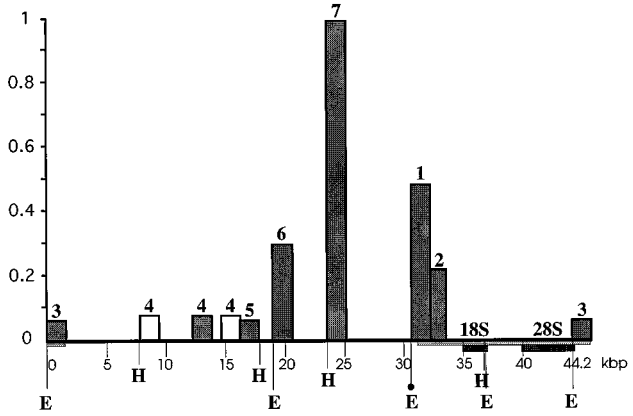


FIG. 5. Nascent-strand abundance analysis of human rDNA. The abundances were determined by densitometry of the signals in Fig. 4, according to the procedure described in Materials and Methods. The map of the human rDNA repeat is similar to that of Fig. 4.

To distinguish between these possibilities, we may consider the sizes of nascent strands displayed in Fig. 4. If there were a single major initiation site near probe 7, the shortest nascent strands detected by probes 1 to 6 ought to increase in size in proportion to distance from probe 7. This is not observed. Instead, all of the probes 1 to 7 detected a range of nascent-strand sizes, from <2 to >13 kb. Although the signals detected by probes 3, 4, and 5 are faint, longer exposures reveal that these signals extend over the same size range as the signals detected by probes 1, 2, 6, and 7.

The simplest interpretation of the combined results in Fig. 4 and 5 is that there are multiple initiation sites, distributed sufficiently frequently through the probed region that each probe is no more than ~1 kbp away from an initiation site, and the initiation sites near probe 7 fire more frequently than others, with the frequency declining in proportion to distance from probe 7.

Possible artifacts. Before one can have confidence in conclusions drawn from these data, it is necessary to consider several potential artifacts which might have contributed to the results.

Could degradation of nascent DNA partially or fully account for the fact that all of the rDNA probes detected similar nascent-strand size distributions? The distinct sedimentation patterns of parental and nascent DNA in the alkaline sucrose gradient (Fig. 2A) suggests that there was minimal strand breakdown prior to or during the alkaline sucrose gradient. However, there is reason to think that during the post-alkaline-sucrose-gradient steps, limited strand breakdown may have occurred in some cases. In two of our three repetitions of this experiment, the maximum strand size detected after the final alkaline agarose gel was only 12 to 13 kb, as in Fig. 4, despite

the fact that we had intended to pool strands up to ~25 kb from the alkaline sucrose gradient. Ethidium bromide staining of the agarose gel in these cases revealed that all of the purified fully BrdUrd-substituted nascent strands had a maximum size of 12 to 13 kb; thus, this size limit was not peculiar to nascent rDNA strands. Despite these differences in maximum strand sizes, in all three experiments probe 7 detected the most abundant nascent strands (as in Fig. 4 and 5) and strands less than 2 kb were detected by all the probes used (as in Fig. 4), even in the experiment in which bulk nascent strands up to 25 kb were recovered. Thus, whether the shorter size range (12 to 13 kb maximum) observed in two experiments was a consequence of limited DNA degradation after the alkaline sucrose gradient or was a consequence of pooling of strands smaller than intended from the alkaline sucrose gradient, the conclusion that there are multiple initiation sites appears valid. It should also be noted that because of the reduced concentration of partially BrdUrd-substituted strands in the population of short strands purified from the alkaline sucrose gradient, nonspecific strand breakdown after the alkaline sucrose gradient should have relatively minor effects on abundance measurements.

Two additional observations suggest that nascent-strand degradation did not significantly influence our results. First, the SV40 control experiment did not show evidence of nascent-strand degradation. Therefore, degradation which would produce nascent-strand fragments shorter than ~5 kb seems unlikely. Second, as expected on the basis of theory and as demonstrated in a control experiment (9), extensive nascent-strand degradation before the alkaline sucrose gradient should have permitted the fully substituted termini of partially substituted strands (strands initiated before the start of the pulse [Fig. 1]) to copurify with the fully substituted nascent strands. This would be expected to minimize differences in nascent-strand abundance, because all regions of the genome should have contained replication forks at the start of the pulse and thus should have been labeled to approximately the same extent with BrdUrd. Thus, the fact that significant differences in nascent-strand abundance were detected (Fig. 4 and 5) suggests that nascent-strand degradation before the alkaline sucrose gradient could not have been extensive.

Could variations in base composition within the rDNA contribute to the observed abundance differences? Probes 1 to 3 have low A+T contents (25, 22, and 29%, respectively). They flank the transcribed region, which is generally low in A+T (28%). The other probes that have been sequenced (probes 4, 6, and 7) have higher A+T contents (49, 53, and 44%, respectively). These variations in base composition could cause problems in either of two ways. First, sequences low in T might not incorporate sufficient BrdUrd to produce sufficient density shift to be efficiently recovered. Second, in the absence of BrdUrd, sequences low in A and T are heavier than AT-rich sequences. For this reason, nonreplicating AT-deficient strands might preferentially contaminate the final preparation of dense nascent strands.

To check whether there was selective loss or gain of strands corresponding to any of the probed regions, we carried out two control experiments. In the first experiment, we labeled asynchronously growing cells with BrdUrd for 24 h (~one generation time) to prepare DNA in which about half of the strands were uniformly BrdUrd substituted. Total genomic DNA was then isolated and sheared to generate a broad size distribution covering the range from ~1 to ~30 kb. The DNA was then denatured and centrifuged to equilibrium in a neutral CsCl gradient. DNA molecules from individual fractions of the gradient were purified, size fractionated by alkaline agarose gel electrophoresis, and blotted onto a nylon membrane. The membrane was then sequentially hybridized with the probes used in the experiment. For each probe, for several different strand sizes, we determined the relative position within the CsCl gradient of the BrdUrd-substituted and unsubstituted DNAs. In the size range 3 to 30 kb, we could not find any significant difference in the positions of the substituted and unsubstituted DNAs detected by any of the probes (51). In the second control experiment, we hybridized dot blots of aliquots of individual fractions from across the first CsCl gradient with each of the probes that we used. For each probe, most of the signal was detected in the light peak, emphasizing the importance of two sequential CsCl gradients to purify the dense nascent strands. Each of the probes detected maxima in the same fractions; no significant differences were seen between the distributions of signals detected by any of the probes. We conclude that the variations in abundance detected in Fig. 4 and 5 are probably not an artifact of variations in BrdUrd content.

Could annealing of substituted nascent DNA strands with contaminating unsubstituted strands during neutral CsCl gradient centrifugation have produced artifactual loss of more highly repeated sequences from the pool of heavy DNA strands? There are two reasons for thinking that this is unlikely to have accrued. First, the only highly repeated sequences known to be present in human rDNA are *Alu* repeats (17, 23, 37), and these appear to be concentrated in the region between probes 7 and 1, the region that yielded the highest strand abundance (Fig. 4 and 5). Significant nascent-strand loss due to annealing of repeated sequences should have reduced the abundance of strands from this region. Second, if reannealing during neutral CsCl gradient centrifugation were to occur, it should have been a more significant problem in the SV40 control experiment, because SV40 DNA molecules are present at very high copy number ($>10^5$ per cell) at 36 h postinfection (13), the time when we carried out our pulse-labeling. Nevertheless, the expected result was obtained (Fig. 3), and the SV40 neutral CsCl gradients did not show evidence of the peak of hybrid heavy-light DNA which would have been expected had significant annealing occurred (51). For both these reasons, we conclude that annealing during CsCl centrifugation could not have been a significant problem.

Is it possible that replication forks move more slowly when emanating from initiation sites near probe 7, leading to a relative abundance of nascent strands detected by probe 7 and flanking probes? The implications of altered fork movement rate are considered in Fig. 1C, which illustrates two initiation sites with equal firing frequencies but unequal fork rates. Forks emanating from the left-hand site are assumed to move one-fourth as fast as those from the right-hand site. If the largest strands pooled from the alkaline sucrose gradient are selected to be shorter than the strands that can be fully labeled at the left-hand site during incubation with BrdUrd, many of the strands generated at the right-hand site will be excluded from the pool and there will be a relative enrichment for strands from the left-hand site. Notice that despite this relative loss of signal from the right-hand site, a probe located near the right-hand site (probe 5 in Fig. 1C) will detect more signal than will probes that are further away (probes 4 and 6). Thus, variations in fork movement rate need not interfere with detection of initiation sites. In addition, the nascent strands detected by probe 7 are >10-fold more abundant than those detected by probes 3 to 5, but variations in fork rate in mammalian cells rarely exceed 3-fold and the greatest detected differences are ~6-fold (19). These differences in fork rate are insufficient to account for the observed differences in nascent-strand abundance (Fig. 5). Finally, using 2D gel electrophoresis, Little et

al. (37) obtained evidence for several sites where replication forks pause downstream of the transcribed region between probes 3 and 4. These sites are likely to be the portions of the rDNA repeat in which replication forks move most slowly. They are too far away to explain the abundant signal detected by probe 7.

Could site-specific nicking (Fig. 1D) partially or fully explain the relatively high abundance of nascent strands detected by probe 7 and neighboring probes? Because neutral-alkaline 2D gel electrophoresis characterization of the nascent strands detected by probes 6 and 7 and a probe adjacent to probe 1 did not reveal extensive site-specific nicking of these strands (Fig. 5 in reference 37), this appears to be an untenable explanation.

Biological implications. Previous studies of mammalian replication origins have yielded apparently contradictory results (see Introduction). A possible explanation of the apparent conflicts between these results is that each of the techniques previously used to study mammalian origins provides limited information about one aspect of origin function. If so, development of new techniques which provide information about different aspects of origin function should help in obtaining a more complete picture.

For this reason, we modified an existing method (nascent-strand size analysis [45]) to develop a new, independent method for mapping origins (nascent-strand abundance analysis). For nascent-strand size analysis, one measures the sizes of nascent strands detectable by probes scattered through the genomic region of interest. Since nascent strands become longer as replication forks move outward from origins, the probes which detect the shortest nascent strands should be located closest to origins. For nascent-strand abundance analysis, one measures the abundance of those nascent strands which become uniformly labeled during a brief incubation with BrdUrd. The probes which detect the most abundant nascent strands should be closest to origins (Fig. 1). Nascent-strand abundance analysis offers the advantages of simplicity, sensitivity, lack of necessity for single-strand-specific probes, applicability to unsynchronized cell populations, and relative insensitivity to possible nascent-strand breakdown during the later stages of nascent-strand purification.

When we applied nascent-strand abundance analysis to replicating SV40 DNA, we obtained the correct origin location (Fig. 3D). The same origin location was also obtained by nascent-strand size analysis (Fig. 3B). However, when nascent-strand abundance and size analyses were applied to a mammalian chromosomal region, human rDNA, the results obtained by the abundance method suggested different but not necessarily contradictory conclusions from those obtained by the size method. We used probes scattered through the human rDNA repeat. Each of these probes detected similar nascent-strand size distributions. In each case, short nascent strands ($<$2 kb) could be detected (Fig. 4 and longer exposures). However, one probe (probe 7) detected more nascent strands than did any other, and the abundances of nascent strands detected by the other probes declined with distance from probe 7 (Fig. 5). Both the size and abundance results are consistent with the possibility that there are multiple initiation sites scattered through human rDNA, and the abundance results suggest that one or more sites near probe 7 fire more frequently than do sites further away from probe 7.

In some respects, our measurements are consistent with previous conclusions from 2D gel electrophoresis studies of human rDNA (37). Both analyses suggest the existence of multiple initiation sites scattered through the rDNA repeat, and both suggest that some initiation sites are used more frequently than others. Our abundance measurements are consistent with the 2D gel observation that the highest initiation frequency is detected in the region between probes 7 and 1 (37). This finding is also consistent with the observation that cloned restriction fragments from the region between probes 7 and 1 are relatively efficient substrates for an in vitro replication system involving proteins from human cells (16).

However, there also appear to be several disagreements between the conclusions suggested by our results and those of the 2D gel electrophoresis studies. The 2D gel electrophoresis results suggest that initiation does not occur in the transcribed region. Three of our probes (probes 1 to 3) are partly or entirely from the transcribed region, and each of these probes detects strands $<$2 kb in length, suggesting nearby initiation. In addition, our abundance results do not confirm the second, weaker preferred initiation region in the vicinity of probes 4 and 5 which was detected by 2D gel electrophoresis measurement (37). We do not know the reasons for these disagreements. Additional methods will probably have to be applied before a full understanding of initiation in the human rDNA repeat can be achieved.

ACKNOWLEDGMENTS

We are grateful to George Russev and Maarten Linskens for initial application of nascent-strand size analysis to replication of human rDNA and for helpful discussions during the course of these studies. We thank James Sylvester for cloned probes and unpublished sequence information from the human rDNA repeat. We also thank Jim Borowiec, Milam Brantley, Bill Burhans, Dharani Dubey, David Kowalski, John Yates, and anonymous referees for useful comments on the manuscript.

This work was supported by Public Health Service grants PO1 GM44119 and RO1 GM49294 from the National Institute of General Medical Sciences. J.A.S. was a Research Fellow of the Leukemia Society of America. C.B. was a fellow of the International Human Frontier Science Program.

REFERENCES

1. **Anachkova, B. B., V. Nosikov, and G. C. Russev.** 1986. Localization of a DNA replication origin in the nontranscribed spacer of rat ribosomal RNA genes. C. R. Acad. Bulg. Sci. **39:**113–116.
2. **Ariizumi, K., Z. Wang, and P. W. Tucker.** 1993. Immunoglobulin heavy chain enhancer is located near or in an initiation zone of chromosomal DNA replication. Proc. Natl. Acad. Sci. USA **90:**3695–3699.
3. **Berberich, S., and M. Leffak.** 1993. DNase-sensitive chromatin structure near a chromosomal origin of bidirectional replication of the avian α-globin locus. DNA Cell Biol. **12:**703–714.
4. **Biamonti, G., G. Perini, F. Weighardt, S. Riva, M. Giacca, P. Norio, L. Zentilin, S. Diviacco, D. Dimitrova, and A. Falaschi.** 1992. A human DNA replication origin: localization and transcriptional characterization. Chromosoma **102:**S24–S31.
5. **Botchan, R. M., and A. I. Dayton.** 1982. A specific replication origin in the chromosomal rDNA of *Lytechinus variegatus*. Nature (London) **299:**453–456.
6. **Bozzoni, I., C. T. Baldari, F. Amaldi, and M. Buongiorno-Nardelli.** 1981. Replication of ribosomal DNA in *Xenopus laevis*. Eur. J. Biochem. **118:**585–590.
7. **Brewer, B. J., and W. L. Fangman.** 1987. The localization of replication origins on ARS plasmids in S. cerevisiae. Cell **51:**463–471.
8. **Brewer, B. J., and W. L. Fangman.** 1988. A replication fork barrier at the 3′ end of yeast ribosomal RNA genes. Cell **55:**637–643.
9. **Brun, C., Y. Yoon, and J. A. Huberman.** Unpublished data.
10. **Burhans, W. C., and J. A. Huberman.** 1994. DNA replication origins in animal cells: a question of context? Science **263:**639–640.
11. **Burhans, W. C., L. T. Vassilev, M. S. Caddle, N. H. Heintz, and M. L. DePamphilis.** 1990. Identification of an origin of bidirectional DNA replication in mammalian chromosomes. Cell **62:**955–965.
12. **Burhans, W. C., L. T. Vassilev, J. Wu, J. M. Sogo, F. S. Nallaseth, and M. L. DePamphilis.** 1991. Emetine allows identification of origins of mammalian DNA replication by imbalanced DNA synthesis, not through conservative nucleosome segregation. EMBO J. **10:**4351–4360.
13. **Calothy, G., K. Hirai, and V. Defendi.** 1973. 5-Bromodeoxyuridine incorporation into simian virus 40 deoxyribonucleic acids: effects on simian virus 40 replication in monkey cells. Virology **55:**329–338.
14. **Campbell, J. L., and C. S. Newlon.** 1991. Chromosomal DNA replication, p. 41–146. *In* J. R. Broach, J. R. Pringle, and E. W. Jones (ed.), The molecular

biology and cellular biology of the yeast *Saccharomyces*: genome dynamics, protein synthesis, and energetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
15. **Carroll, S. M., M. L. DeRose, J. L. Kolman, G. H. Nonet, R. E. Kelly, and G. M. Wahl.** 1993. Localization of a bidirectional DNA replication origin in the native locus and in episomally amplified murine adenosine deaminase loci. Mol. Cell. Biol. **13:**2971–2981.
16. **Coffman, F. D., I. Georgoff, K. L. Fresa, J. Sylvester, I. Gonzalez, and S. Cohen.** 1993. *In vitro* replication of plasmids containing human ribosomal gene sequences: origin localization and dependence on an aprotinin-binding cytosolic protein. Exp. Cell Res. **209:**123–132.
17. **Dickson, K. R., D. C. Braaten, and D. Schlessinger.** 1989. Human ribosomal DNA: conserved sequence elements in a 4.3-kb region downstream from the transcription unit. Gene **84:**197–200.
18. **Dijkwel, P. A., and J. L. Hamlin.** 1992. Initiation of DNA replication in the dihydrofolate reductase locus is confined to the early S period in CHO cells synchronized with the plant amino acid mimosine. Mol. Cell. Biol. **12:**3715–3722.
19. **Edenberg, H. J., and J. A. Huberman.** 1975. Eukaryotic chromosome replication. Annu. Rev. Genet. **9:**245–284.
20. **Fangman, W. L., and B. J. Brewer.** 1991. Activation of replication origins within yeast chromosomes. Annu. Rev. Cell Biol. **7:**375–402.
21. **Feinberg, A. P., and B. Vogelstein.** 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132:**6–13.
22. **Gale, J. M., R. A. Tobey, and J. A. D'Anna.** 1992. Localization and DNA sequence of a replication origin in the rhodopsin gene locus of Chinese hamster cells. J. Mol. Biol. **224:**343–358.
23. **Gonzalez, I. L., S. Wu, W.-M. Li, B. A. Kuo, and J. E. Sylvester.** 1992. Human ribosomal RNA intergenic spacer sequence. Nucleic Acids Res. **20:**5846.
24. **Handeli, S., A. Klar, M. Meuth, and H. Cedar.** 1989. Mapping replication units in animal cells. Cell **57:**909–920.
25. **Hay, R. T., and M. L. DePamphilis.** 1982. Initiation of SV40 DNA replication in vivo: location and structure of 5′ ends of DNA synthesized in the *ori* region. Cell **28:**767–779.
26. **Hernandez, P. C., B. J. Bjerknes, S. S. Lamm, and J. Van't Hof.** 1988. Proximity of an ARS consensus sequence to a replication origin of pea (*Pisum sativum*). Plant Mol. Biol. **10:**413–422.
27. **Huberman, J. A., D. D. Dubey, K. A. Nawotka, G. Russev, J. A. Sanchez, Y. Yoon, and M. H. K. Linskens.** 1992. Directions of DNA replication in yeast and mammalian cells, p. 83–95. *In* P. Hughes, E. Fanning, and M. Kohiyama (ed.), DNA replication: the regulatory mechanisms. Springer-Verlag KG, Berlin.
28. **Huberman, J. A., L. D. Spotila, K. A. Nawotka, S. M. El-Assouli, and L. R. Davis.** 1987. The in vivo replication origin of the yeast 2μm plasmid. Cell **51:**473–481.
29. **Hyrien, O., and M. Méchali.** 1993. Chromosomal replication initiates and terminates at random sequences but at regular intervals in the ribosomal DNA of *Xenopus* early embryos. EMBO J. **12:**4511–4520.
30. **Iguchi-Ariga, S. M. M., N. Ogawa, and H. Ariga.** 1993. Identification of the initiation region of DNA replication in the murine immunoglobulin heavy chain gene and possible function of the octamer motif as a putative DNA replication origin in mammalian cells. Biochim. Biophys. Acta **1172:**73–81.
31. **Jacob, F., S. Brenner, and F. Cuzin.** 1963. On the regulation of DNA replication in bacteria. Cold Spring Harbor Symp. Quant. Biol. **28:**329–348.
32. **James, C. D., and M. Leffak.** 1986. Polarity of DNA replication through the avian alpha-globin locus. Mol. Cell. Biol. **6:**976–984.
33. **Khadzhiolov, A. A.** 1985. The nucleolus and ribosome biogenesis. Springer-Verlag, New York.
34. **La Volpe, A., A. Simeone, M. D'Esposito, L. Scotto, V. Fidanza, A. de Falco, and E. Boncinelli.** 1985. Molecular analysis of the heterogeneity region of the human ribosomal spacer. J. Mol. Biol. **183:**213–223.
35. **Leu, T.-H., and J. L. Hamlin.** 1989. High-resolution mapping of replication fork movement through the amplified dihydrofolate reductase domain in CHO cells by in-gel renaturation analysis. Mol. Cell. Biol. **9:**523–531.
36. **Linskens, M. H. K., and J. A. Huberman.** 1988. Organization of replication of ribosomal DNA in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **8:**4927–4935.
37. **Little, R. D., T. H. K. Platt, and C. L. Schildkraut.** 1993. Initiation and termination of DNA replication in human rRNA genes. Mol. Cell. Biol. **13:**6600–6613.
38. **Muller, F., and F. Grummt.** 1991. Mapping eukaryotic replication origins in vivo by size analysis of purified nascent DNA strands. DNA Cell Biol. **10:**149–157.
39. **Safrany, G., and E. J. Hidvegi.** 1989. New tandem repeat region in the nontranscribed spacer of human ribosomal RNA gene. Nucleic Acids Res. **17:**3013–3022.
40. **Shinomiya, T., and S. Ina.** 1993. DNA replication of histone gene repeats in *Drosophila melanogaster* tissue culture cells: multiple initiation sites and replication pause sites. Mol. Cell. Biol. **13:**4098–4106.
41. **Sylvester, J. E., D. A. Whiteman, R. Podolsky, J. M. Pozsgay, J. Respess, and R. D. Schmickel.** 1986. The human ribosomal genes: structure and organization of the complete repeating unit. Hum. Genet. **73:**193–198.
42. **Taira, T., S. M. M. Iguchi-Ariga, and H. Ariga.** 1994. A novel DNA replication origin identified in the human heat shock protein 70 gene promoter. Mol. Cell. Biol. **14:**6386–6397.
43. **Tasheva, E. S., and D. J. Roufa.** 1994. A mammalian origin of bidirectional DNA replication within the Chinese hamster *RPS14* locus. Mol. Cell. Biol. **14:**5628–5635.
44. **Truett, M. A., and J. G. Gall.** 1977. The replication of ribosomal DNA in the macronucleus of *Tetrahymena*. Chromosoma **64:**295–303.
45. **Vassilev, L., and E. M. Johnson.** 1989. Mapping initiation sites of DNA replication *in vivo* using polymerase chain reaction amplification of nascent strand segments. Nucleic Acids Res. **17:**7693–7705.
46. **Vassilev, L. T., W. C. Burhans, and M. L. DePamphilis.** 1990. Mapping an origin of DNA replication at a single-copy locus in exponentially proliferating mammalian cells. Mol. Cell. Biol. **10:**4685–4689.
47. **Vaughn, J. P., P. A. Dijkwel, and J. L. Hamlin.** 1990. Replication initiates in a broad zone in the amplified CHO dihydrofolate reductase domain. Cell **61:**1075–1087.
48. **Virta-Pearlman, V. J., P. H. Gunaratne, and A. C. Chinault.** 1993. Analysis of a replication initiation sequence from the adenosine deaminase region of the mouse genome. Mol. Cell. Biol. **13:**5931–5942.
49. **Vogt, V. M., and R. Braun.** 1977. The replication of ribosomal DNA in *Physarum polycephalum*. Eur. J. Biochem. **80:**557–566.
50. **Wu, C., M. Zannis-Hadjopoulos, and G. B. Price.** 1993. In vivo activity for initiation of DNA replication resides in a transcribed region of the human genome. Biochim. Biophys. Acta **1174:**258–266.
51. **Yoon, Y., and J. A. Huberman.** Unpublished data.