From: DNA SYNTHESIS: Present and Future
Edited by Ian Molineux and Masamichi Kohiyama
[Plenum Publishing Corporation, 1978]

THE ORGANIZATION OF DNA REPLICATION IN A MAMMALIAN CELL LINE

Barbara R. Jasny*, Joel E. Cohen and Igor Tamm

The Rockefeller University

New York, N.Y. 10021, U.S.A.

INTRODUCTION

The first evidence that DNA replication proceeds in an organized, nonrandom manner at the level of the chromosome was obtained by Taylor (1960) using autoradiographic techniques to examine labeled metaphase chromosomes. His work, and that of Hsu (1964), demonstrated that DNA replication is initiated at multiple sites along the chromosome and that each chromosome shows a reproducible, characteristic pattern of replication in different stages of S phase. It has also been shown that hetero-chromatin tends to be late replicating (see Lima-de-Faria and Jaworska, 1968). Even better resolution is now possible with the fluoresence labeling technique developed by Latt (1973). Both Latt (1975) and Stubblefield (1975) have suggested that chromosome bands which have been visualized as structural units by classical staining procedures are actually units of replication as observed by fluoresence labeling.

It has not yet been possible to correlate this organized pattern of replication with replication at the level of the double-stranded DNA itself. The technique of DNA fiber autoradiography, originated by Cairns (1963) and further developed by Huberman and Riggs (1968), has been used to determine the rate of fork progression, the distance between initiation sites, and the direction of replication. Hand and Tamm (1974) have suggested that there is synchrony in time of initiation of clusters of replication units, based on the occurrence of predominantly prepulse or postpulse figures in a microscopic field. Hand (1975) has obtained additional evidence for such synchrony from studies of individual fibers. The

*A.S.I. *Participant*

autoradiographic technique has the advantages that observations can be made on individual fibers and that many replicating fibers can be readily examined. The disadvantages are that only a small fraction of the DNA is spread well enough to be analyzed (Edenberg and Huberman, 1975), that there is subjectivity in visual observations, and that there can be wide variation in measurement of such parameters as interinitiation site distance.

To examine the organization of replication units within the genome and to determine whether initiation events occur at random along the DNA fiber, it is necessary to evaluate the processes giving rise to the autoradiograms and the replicating strands themselves in a more analytical manner than has previously been used. It is the aim of this work to provide some new insights into the organization of the initiation process.

MATERIALS AND METHODS

Cells

Monolayer cultures of uncloned L929 cells (a continuous line of mouse fibroblasts) were grown in Eagle's medium supplemented with 5% fetal calf serum. Before each experiment, the L929 cells were passaged once in 75 cm^2 Falcon flasks under conditions of exponential growth, then passaged under similar conditions into 28 cm^2 plastic dishes at approximately 1×10^5 cells per plate and allowed to double twice before use. Two dishes were used per experiment.

Radioisotopes and Chemicals

(^{3}H)thymidine (50-52 Ci/mmole) was obtained from New England Nuclear, Boston, Mass., and thymidine from Sigma Chemical Company, St. Louis, Mo. 5-Fluoro-2'-deoxyuridine (FUdR) was a gift from Hoffman-La Roche, Inc., Nutley, N.J.

Cell Labeling and DNA Spreading Procedures

The medium in culture dishes was replaced with fresh warm medium containing 2×10^{-6} M FUdR to deplete the thymidine pool and the cells were incubated at 37°C for 30-40 minutes. They were labeled for 10 minutes at 37°C with high specific activity (^{3}H)-thymidine (250 μCi/ml, 50-52 Ci/mmol, 5×10^{-6}M) in fresh medium containing 2×10^{-6} M FUdR (the "hot" pulse). Unlabeled thymidine was then added (as a 200 μM solution) so as to give a 10-fold reduction in specific activity and incubation continued for 3 hours at 37°C (the "warm" pulse). The cells were washed 2-4 times in ice-

cold phosphate buffered saline deficient in Ca^{2+} and Mg^{2+} (PBS def.), removed from the plates with trypsin, and washed and resuspended in PBS def.

One drop containing 1000-3000 cells was placed on a subbed slide adjacent to a drop of a solution of 1% SDS and 0.01 M EDTA in PBS def. The drops were mixed by touching both simultaneously with a glass rod, and were then spread down the slide. By this procedure the cells were lysed and the DNA spread out as long strands. In some experiments the labeled cells were diluted with unlabeled cells before spreading to reduce the amount of labeled DNA per slide. The slides were allowed to dry, followed by precipitation in 5% TCA and dehydration in ethanol. Slides were dipped in melted Kodak NTB2 emulsion using the red safety light, dried for 2 hours in total darkness and stored for approximately 5 months at -20°C. They were developed in Kodak D19 developer at 17°C for 4 minutes.

Analysis of Autoradiograms

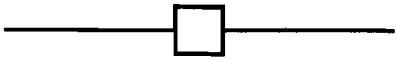
Autoradiograms of (^{3}H)thymidine-labeled DNA were examined by light microscopy at magnifications of 260 x and 640 x. All measurements were made at 640 x with an ocular micrometer. Statistical analysis of the grain density was performed on DNA strands labeled lightly by (^{3}H)thymidine of low specific activity in order to distinguish breaks in strands from apparent gaps representing stretches with a low probability of grain formation. The grains over a ten division span of the ocular micrometer (17.1 μm) on 100 randomly chosen pieces of lightly labeled DNA were counted. The mean number of grains per division (which varied from 0.822 to 1.07 in different experiments) was used to determine the probability of finding a given number of strains in a given interval according to the Poisson distribution formula: $p_n = m^n e^{-n}/n!$, where m equals the mean number of grains in the interval, and n equals the number of grains actually found. If the probability of finding no grains on a stretch of lightly labeled DNA was less than 0.01, then the interval was considered a break. Typically, unlabeled stretches greater than 7 to 11 μm (with the range indicating variation between experiments) were considered significant breaks in the warm pulse-labeled DNA.

The end of a warm-pulse labeled strand is considered to be "free" if it does not terminate in a mass of nuclear material. Thus strands with two free ends have a naturally defined total length which results from breaks or termination of label at each end.

The protocol used gave rise to two types of figures: (1) Pre-pulse figures, in which initiation occurred before the hot pulse. These figures showed a central unlabeled stretch bordered by high

grain density regions (shown as rectangles in the following diagram), followed by lightly labeled regions (shown as lines).

Prepulse initiation is assumed to have occurred at the midpoint of the central unlabeled stretch, followed by bidirectional replication. Equilibration of the thymidine pool required time, so that up to 2.6 μm of lightly labeled DNA on either border of the gap region was allowed; (2) Postpulse figures, in which initiation occurred after the start of the pulse. These showed a central heavily labeled region, bounded on both sides by lightly labeled DNA. Postpulse initiation is assumed to have occurred at the midpoint of the central heavily labeled region.



Determinations of the occurrence of hot pulse-labeled regions were based on clear visibility of high grain density at both the lower and higher magnifications. Intervals of $\leqslant$ 0.87 μm ($\leqslant$ 0.5 division at the high magnification) appearing either totally clear or lightly labeled within a hot region were not considered significant.

Huberman and Riggs (1968) used the term replication section to refer to the stretch of DNA replicated by a single growing point. They proposed the term replication unit to mean the basic unit of control of the initiation of replication - presumably an adjacent pair of diverging replication sections.

Measurements of the distances between adjacent initiation sites provide estimates of the size of replication units. It should be noted, however, that the stretch of DNA between two initiation sites is not equivalent to the replication unit as defined by Huberman and Riggs (1968), but rather comprises the neighboring halves of two adjacent replication units. The interinitiation distance was measured as a center to center distance for all internal figures. A prepulse initiation figure (see diagram) was considered internal regardless of the presence or absence of a warm pulse-labeled stretch of DNA at the strand terminus. A postpulse initiation figure (see diagram) was considered internal if there was a stretch of $\geqslant$ 3.42 μm of warm pulse-labeled DNA distal to the hot pulse-labeled region.

RESULTS

Rate of Fork Progression and Temporal Synchrony

The rate of fork progression was calculated by measuring the lengths of the internal hot pulse-labeled halves of prepulse figures

(i.e., not at the ends of strands) and dividing such lengths by the duration of the hot pulse, i.e., 10 minutes. The frequency distribution of rate measurements is shown in Fig. 1. The mean rate of fork progression in L cells is 0.55 μm/minute, which is within the range of previous estimates from 0.4 to 1.2 μm/minute (Hand and Tamm, 1972).

To evaluate the temporal synchrony with which initiation events observed with our techniques are taking place, strands containing prepulse figures were examined, since prepulse figures potentially span a longer period of replication than postpulse figures. Very large prepulse figures may not be recognized as part of one strand because the unlabeled stretch may have been interpreted as a break rather than a prepulse gap. Only strands with lengths greater than or equal to 40 μm were used in this analysis of synchrony. For each prepulse figure on the strand, the mean distance from the assumed initiation point at the center of the figure to the ends of the hot-labeled segments was calculated (i.e., the average of the left and right halves of each prepulse figure). Then, for each possible pair of prepulse figures on a single strand, the smaller mean distance was subtracted from the greater mean distance to give a *difference* in mean distance. Fig. 2 shows the distribution of the *differences* between these mean lengths of DNA replicated from the time of initiation to the end of the 10-minute hot pulse for all of the strands. Based on the mean replication rate of 0.55 μm/minute, it is concluded that most of the observed initiations occurred within 10 minutes of each other.

The preceding results define the "window" in time through which observations are made. They indicate that there is considerable

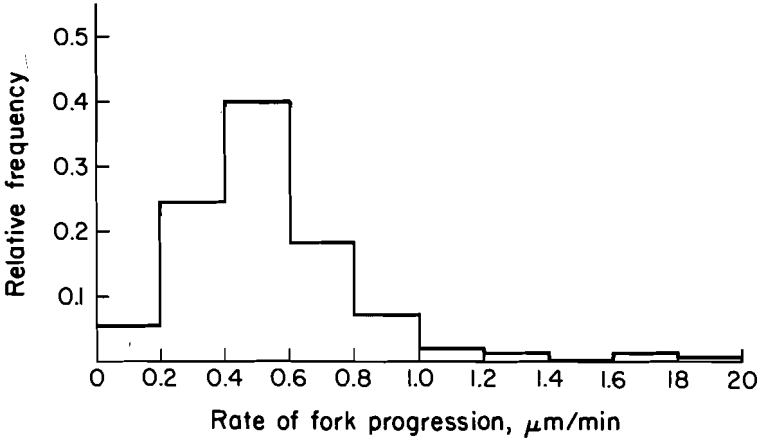


Figure 1. Frequency distribution of rates of fork progression. Mean distribution, based on the mean of three experiments (357 strands total), is shown. Strands examined did not necessarily have two free ends.

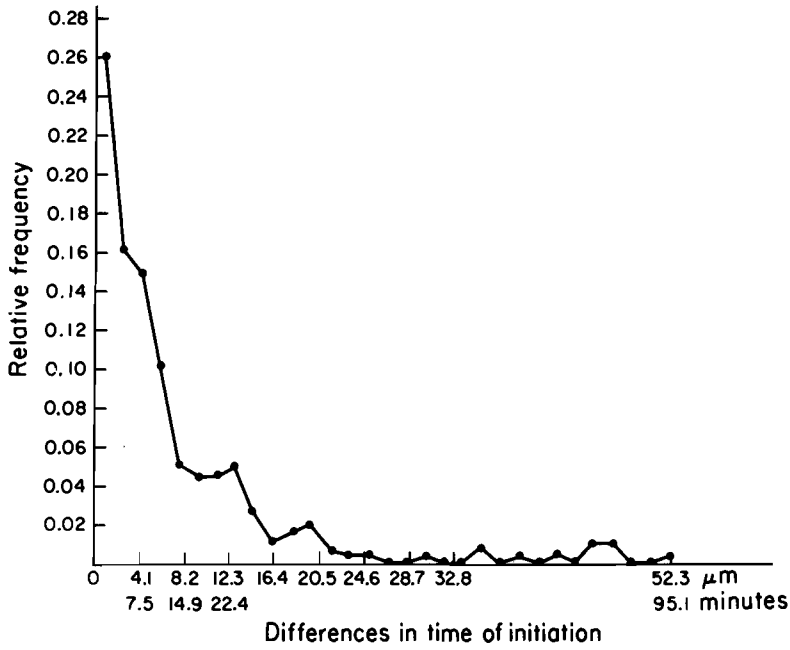


Figure 2. Frequency distribution of differences in time of initiation on individual strands. The mean distance from the assumed initiation point at the center of the figure to the ends of the two hot pulse-labeled segments was calculated for prepulse figures on each strand and for each pair of prepulse figures, the difference between these mean distances was determined. These differences were converted into time differences by multiplying the lengths by the mean rate of fork progression (0.55 μm/min = 0.33 ocular divisions/min). There were 268 differences plotted, representing data from 92 strands, not necessarily with 2 free ends, taken from three experiments.

temporal synchrony, and that a reasonable picture of the initiation pattern can be developed from the events observed on the strand using a 10-minute hot pulse followed by a 3-hour period of low specific activity labeling.

Spatial Organization

In order to determine whether initiation events occur at locations distributed at random or according to some demonstrable organization, we developed quantitatively precise predictions, based on a number of idealized assumptions, concerning what would be expected assuming randomness. First, the DNA fiber was treated as if the length of an unbroken DNA fiber greatly exceeded the average length of a strand with 2 free ends, which appears to be

true. Second, within every replication unit, initiation was assumed to occur at a single point each located by a single prepulse or postpulse figure. Here, the length of hot pulse-labeled regions of prepulse or postpulse figures is irrelevant. The model which we have compared with our data is that there are two independent and strictly random (Poisson) processes occurring simultaneously on a DNA fiber. One process distributes nicks at random, with a negative exponential distribution of distances between two adjacent nicks and with an average number λ of nicks per unit length of DNA fiber. The nicks considered here are only those nicks which will generate strands with 2 free ends. Thus, we are testing the assumption that strands with prepulse and post-pulse figures are broken from the DNA fiber in a random manner. A second hypothetical process distributes initiation events, seen as prepulse or postpulse figures, at random, again with a negative exponential distribution of distances between adjacent events and an average number ν of events per unit length of DNA fiber.

For every strand with 2 free ends on two slides per experiment, we recorded the total length of the strand, the number of replication figures per strand, and the lengths of the external segments. An external segment is defined as the region of warm-labeled DNA between the end of the strand and the nearest hot-labeled region. Only strands containing one or more hot pulse-labeled regions were selected for observations, and all such strands on a slide were recorded.

Nonrandom Breakage of Strands

From the random model, it may be calculated that the probability density function of the length y of an external segment is:

$$g(y) = (\nu + \lambda)e^{-(\nu + \lambda)y}$$

λ = average number of nicks per unit length

ν = average number of initiation events per unit length

$y \geq 0$

This is a negative exponential distribution. If the external segments y are ranked in order of increasing size and plotted as the cumulative frequency against external segment size, the dotted line in Fig. 3 is generated. A theoretical curve (solid line) can be constructed by taking the integral of the probability density function in the limits from zero to y. The data deviate

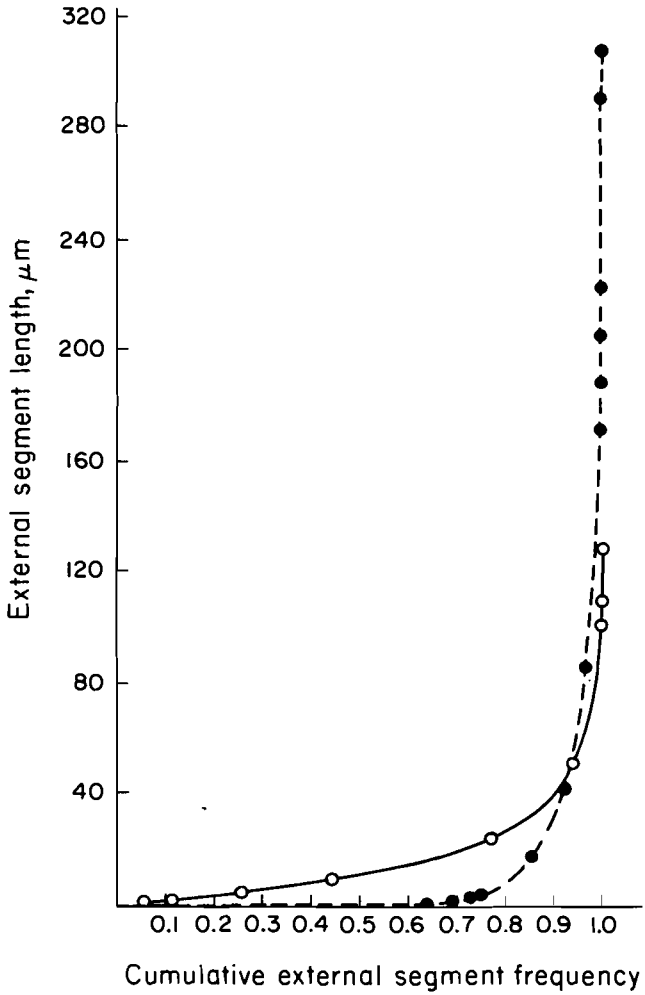


Figure 3. Frequency distribution of external segment lengths. ------ experimental data, _____ theoretical curve derived as described in the text. All strands had 2 free ends. The data represent 3476 external segments pooled from three experiments.

greatly from the model. Far too many short external segments are observed, and there appear to be longer segments than would have been predicted by a random model of strand breakage.

A simple explanation for the increased probability of breakage of a strand adjacent to a hot pulse-labeled region would be that the high specific activity label causes excess breaks. Although this

may be a contributing factor, we do not believe that it adequately explains the nonrandom breakage. If breaks occurred only adjacent to hot regions, then the external segment lengths should have a distribution similar to that of the inter-initiation distances (the distances between adjacent initiation sites). If nonrandom breakage were occurring between adjacent hot pulse figures at single stranded regions generated during replication, then the external segment distribution should show a higher frequency of shorter lengths than the inter-initiation distance distribution. Our evidence does not support either of the above possibilities. In Fig. 4 are plotted only external segments of lengths greater than or equal to 3.42 μm which is the limit of resolution for the inter-initiation distances. All such external segments were taken from every strand with 2 free ends observed on the slides. The mean external segment length was 44.98 μm, which was 35% higher than the mean inter-initiation distance of 33.26 μm. In this comparison, each experiment was weighted according to the sample size. The difference was significant at the 0.005 probability level by t test.

To investigate breakage of the DNA fiber into strands further, we tested whether or not breakage at one end of the strand is independent of breakage at the other end. This was done by examining the number of strands having 0, 1, or both external segments longer than some given length y. If F is the fraction of all external segments which are longer than y, then, assuming that the two external segments are independent, the proportion of all strands with both external segments longer than y should be F^2, the proportion with exactly one external segment longer than y should be $2F(1 - F)$, and the proportion with zero external segments longer than y should be $(1 - F)^2$. This model is formally identical to the Hardy-Weinberg equilibrium of population genetics. There was no significant deviation from the model for any value of y in any of the three experiments as determined by χ^2 analysis (data not shown). Thus, it can be concluded that there is no interaction between the length of the external segment at one end of the strand, and the length of external segment at the other end. Possible implications of this result are discussed below.

Inter-Initiation Distance Distribution

The distances between initiation sites reveal one aspect of the organization of replication units. Although the situation is complicated by the non-random breakage of strands and by our finding that the inter-initiation distances vary with strand lengths (data not shown), it is possible to test our data for randomness in the distribution of inter-initiation distances. If the activated initiation sites are distributed randomly (i.e., with negative exponential distances between them) on the unbroken DNA fiber, then

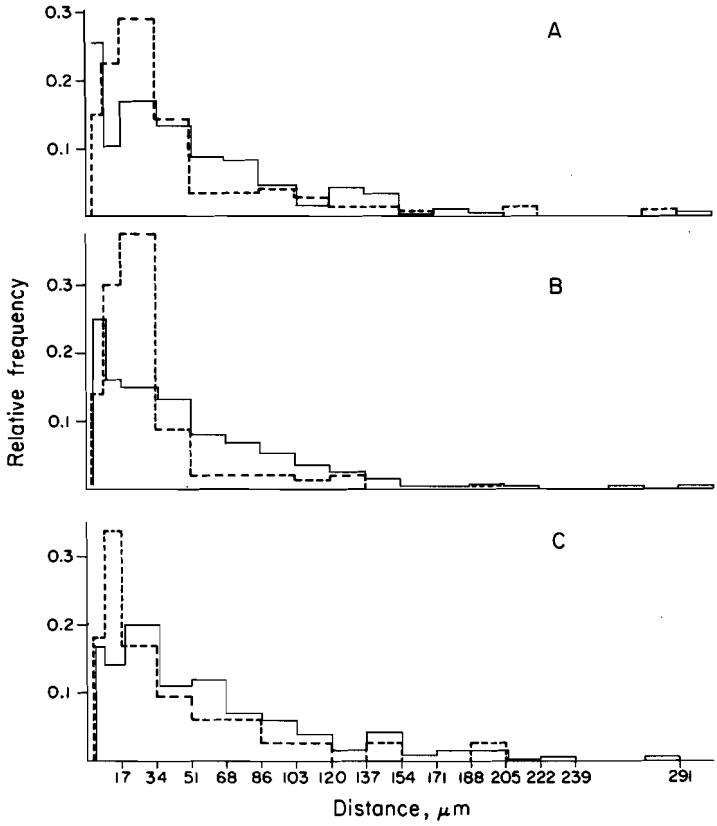


Figure 4. Frequency distribution of external segments and inter-initiation distances ≥ 3.42 μm (2 ocular divisions) in length. Panels A-C show results from 3 separate experiments. ----- inter-initiation site distance, ——— external segment length. The data represent all strands with 2 free ends on 2 slides per experiment, plus a population of strands longer than 20-30 divisions picked from additional slides. The longer strands also had 2 free ends. For the inter-initiation distance distributions, in experiment A, n = 147; in B, n = 150; in C, n = 83. For the external segment distributions in experiment A, n = 288; in B, n = 486; in C, n = 403.

it is possible to derive a function B such that:

$$B = 2n(\bar{x})/s,$$

n = number of observations,

$\bar{x}$ = mean inter-initiation distance,

s = standard deviation of sample mean.

As n becomes larger, B has approximately the distribution of χ^2 with 2n degrees of freedom. Then (Wilson-Hilferty, 1931) if $([B/(2n)]^{1/3} + 1/(9n) - 1)(9n)^{1/2} = z$, z has normal distribution with zero mean and unit variance. If $|z|$ is less than 2.56, then by this test the distances do not deviate significantly from exponential at the 1% level. Table 1A shows inter-initiation distances from three experiments in which all strands with 2 free ends were recorded. $|z|$ was less than 2.56 in all cases. When the mean inter-initiation distance was calculated from the 3 experiments (203 distances), z = -2.48, which is suggestive of a random distribution. Since strands are not equally represented by the above calculation (i.e., strands with more than two replication units can generate many inter-initiation distances), a distance distribution was calculated for hot pulse-labeled autoradiograms in strands containing only 2 activated initiation sites. For all such strands the total external segment length was subtracted from the length of the entire strand, generating a distance analogous to the inter-initiation distance (315 strands; Table 1B). A z equalling -2.13 was determined, which is also within the limits of ±2.56 set for a random distribution.

In each of the three experiments in Table 1, the mean distance between hot pulse-labeled regions on strands with two figures (Table 1B) exceeds the corresponding mean inter-initiation distances (by 6 μm in experiment 6; 11 μm in experiment 14; and 17 μm in experiment 5). The excess corresponds roughly to the mean length of a single hot pulse-labeled region. The difference arises because the two figure distances include the total lengths of the hot labeled regions in each of the two replication figures, whereas the inter-initiation distances extend from midpoint to midpoint, approximately, of adjacent replication figures.

DISCUSSION

The finding of nonrandom breakage has important implications for DNA replication. The excess breakage which may be caused by the presence of high specific activity label is not sufficient to explain the apparently nonrandom breakage at distances far removed from the hot pulse-labeled regions. If breakage were due to the presence of the single-stranded region at the replication fork, or was occurring at or near the hot pulse region , then we would not expect to observe, as we did, that the mean external segment was significantly larger than the mean inter-initiation distance.

We find that there is no interaction between the length of the external segment at one end of a strand and the length of the external segment at the other end. This finding appears to exclude the possibility that in some large stretches of the DNA fiber, containing many of the hypothesized clusters of replication units, there is a higher probability of breakage at any given distance

Table I

A. Interinitiation Distances

		Expt 5	Expt 6	Expt 14	Pooled data*
Number	n	97	39	67	203
Mean	$\bar{\chi}$ (μm)	30.44	34.7	44.5	35.76
Standard deviation	sd (μm)	31.42	48.0	56.0	42.95
	$z^{\dagger}$	-0.66	-1.82	-1.87	-2.48

B. Two Figure Distances

		Expt 5	Expt 6	Expt 14	Pooled data
Number	n	121	103	91	315
Mean	$\bar{\chi}$ (μm)	47.23	40.62	55.1	47.33
Standard deviation	sd (μm)	56.06	42.60	63.40	53.77
	$z^{\dagger}$	-1.80	-0.44	-1.27	-2.13

*In pooling the data, each distance was weighted equally, so that each experiment was weighted by n.

†Calculated from eqs. (2) and (3). If $|z| > 2.56$, we interpret the distribution of distances to be nonrandom.

from a hot pulse-labeled region than there is in other large stretches. If the DNA fiber were inhomogeneous in this way, we would have expected to observe an excess of strands with two short external segments and an excess with two long external segments, in comparison with the numbers expected using the probabilities based on strands pooled from all parts of the fiber. Thus, there do not appear to be regions, detectable by our procedures, which contain "weak spots" that cause the nonrandom breakage. It is unlikely that weak spots account for the presence of external segments which are longer than the distances between initiation sites.

Although we are not yet prepared to offer a definitive explanation of these findings, one reasonable possibility is that, at specific times during S phase, physical clusters of replication units are activated. These clusters are separated by distances longer than the measured, mean inter-initiation distance. The high degree of temporal synchrony observed (Fig. 2) indicates that the firing of initiation sites within a cluster occurs within a fairly short period of time. However, the physical site at which initiation occurs within such a cluster may be randomly determined, as suggested by the exponential distribution of inter-initiation distances and interunit distances on strands containing only 2 replication units.

Support for the theory of physical clustering of replication units can be found from studies using very different methodologies. Stubblefield (1975) and Latt (1975) have observed reproducible banding patterns when the chromosomes of cells which have been treated with BUdR during the entire S phase, except for a short period of thymidine incorporation, are analyzed by fluorescence labeling with the Hoechst stain 33258. These bands appear to correspond with banding patterns obtained when chromosomes are stained with quinacrine or by G banding procedures. They have theorized that entire regions of the genome, comprising lengths equivalent to hundreds of replication units, are made available for initiation at specified times within the S phase. Individual units in the cluster could initiate replication randomly. The finding of clusters of units in replicating drosophila DNA by means of electron microscopic techniques (Zakian, 1976) is also in accord with this hypothesis.

The data described in this paper do not rule out the possibility that *potential* initiation sites are located at non-random intervals. Such structures may be located at short intervals along the strand with only a small percentage of such sites being activated during any one S phase. The observed differences in inter-initiation site distances among species (Hand and Tamm, 1974) would agree with this possibility and could have evolutionary significance. There is also evidence that there are different subfractions of replicating DNA, which show different patterns of replication (Zakian, 1976; Hori and Lark, 1976). A sufficiently small percentage of the replicating DNA could have a nonrandom distribution of inter- initiation distances without being detected in our analysis. In Table I, in each experiment, the mean distance is less than the standard deviation, and z is negative. This pattern suggests the possibility of a quantitatively small deviation from randomness.

Further research in this area will entail development of more detailed structural models. We are also attempting to correlate the replication patterns observed at the fiber level with banding patterns observed on isolated chromosomes.

ACKNOWLEDGEMENTS

This investigation was supported in part by Research Grant Number CA-18608 and by Program Project Grant Number CA-18213, awarded by the National Cancer Institute, and by U.S. National Science Foundation grant BMS74-13276. B.R. Jasny is a predoctoral trainee under the Institutional National Research Service Award Number CA-09256, from the National Cancer Institute.

REFERENCES

1. Edenberg, H.J. and Huberman, J.A. (1975) Ann. Rev. Genet. 9:245-284

2. Hand, R. (1975) J. Cell Biol. 64:89-97

3. Hand, R. and Tamm, I. (1974) J. Mol. Biol. 82:175-183

4. Hori, T.A. and Lark, K.G. (1976) Nature 259:504-505

5. Hsu, T.C. (1964) J. Cell Biol. 23:53-62

6. Huberman, J.A. and Riggs, A.D. (1968) J. Mol. Biol. 32: 327-341

7. Latt, S.A. (1973) Proc. Nat. Acad. Sci. USA 70:3395-3399

8. Latt, S.A. (1975) Somat. Cell Genet. 1:293-321

9. Lima-de-Faria, A. and Jaworska, H. (1968) 217:138-142

10. Stubblefield, E. (1975) 53:209-221

11. Taylor, J.H. (1960) J. Biophys. Biochem. Cytol. 7:455-464

12. Wilson, E.B. and Hilferty, M.M. (1931) Proc. Nat. Acad. Sci. USA 17:684-686

13. Zakian, V.A. (1976) J. Mol. Biol. (1976) 108:305-331