The Units of DNA Replication in *Drosophila melanogaster* Chromosomes

Alan B. Blumenthal, Henry J. Kriegstein and David S. Hogness

Access the most recent version at doi:10.1101/SQB.1974.038.01.024

References

This article cites 26 articles, 7 of which can be accessed free at:
http://symposium.cshlp.org/content/38/205.refs.html

Article cited in:
http://symposium.cshlp.org/content/38/205#related-urls

Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here

To subscribe to *Cold Spring Harbor Symposia on Quantitative Biology* go to:
http://symposium.cshlp.org/subscriptions

Copyright © 1974 Cold Spring Harbor Laboratory Press
The Units of DNA Replication in Drosophila melanogaster Chromosomes

ALAN B. BLUMENTHAL,* HENRY J. KRIEGSTEIN, AND DAVID S. HOGNESS
Department of Biochemistry, Stanford University School of Medicine, Stanford, California 94305

The DNA which must be replicated in a chromosome of Drosophila melanogaster appears to exist as a single molecule of double-stranded DNA, which, for the largest chromosomes, has a length of about 2.1 cm, or 62,000 kb\(^1\) (Kavenoff and Zimm, 1973). We have studied both the topography of the units of replication in this chromosomal DNA and the rate of replication per unit in two different classes of Drosophila nuclei which exhibit very different S phases. Our purpose is to define the factors which determine the overall replication rate for these giant DNA molecules.

The two classes are the rapidly replicating cleavage nuclei and the slowly replicating nuclei in cell cultures. At 24°C, the cleavage nuclei divide every 9.6 min in the syncytium of the egg for a period of about two hours after fertilization (Rabinowitz, 1941). Since interphase occupies 3.4 min of this doubling time, we presume that each chromosomal DNA molecule is replicated within this short period. By contrast, the nuclei in cell cultures at 25°C exhibit an S phase of about 600 min (Dolfini et al., 1970), more than two orders of magnitude greater than that for cleavage nuclei.

The molecular rate of replication for the DNA in the largest chromosomes in the cleavage nuclei must then be equal to or greater than 18,000 kb · min\(^{-1}\) · molecule\(^{-1}\). This extremely rapid molecular replication is about four orders of magnitude greater than the upper estimates for the rate of movement of a DNA replication fork in animal chromosomes (Huberman and Riggs, 1968; Callan, this volume) and would therefore require the cooperative action of thousands of these forks per molecule. This condition has allowed an examination of both the structure and the topography of replication forks in cleavage nuclei by electron microscopy. Forks in the nuclei from cell cultures would, however, be extremely tedious to study in the electron microscope because of the much lower fork frequency, and we have therefore used the technique of radioautography to examine this state.

A situation in which multiple forks cooperate in the replication of a single, duplex molecule is depicted by configuration A in Figure 1, where two topologically related ways of replicating a large amount of DNA in a short time are indicated. Configuration A is similar to that suggested by Huberman and Riggs (1968) to account for the radioautographic patterns produced by replicating DNA from mammalian chromosomes. Here each of many origins \((o_i)\) in a single DNA molecule generates two replication forks, which move in opposite directions to create a serial array of "eye forms." Configuration B consists of a set of circular DNA molecules, each of which contains a single origin with this same capacity to generate bidirectional replication. The resulting "θ forms" are typical of the replicating configurations of many prokaryotic chromosomes and have been described in considerable detail for bacteriophage θ (Schnöss and Inman, 1970; Inman and Schnöss, 1971). Since configuration A can be formed from B by cutting each circular DNA at a position other than \(o_i\) and joining the cut ends of adjacent circles, the topological aspects of replication are much the same in the two configurations.

However, the two configurations exhibit an important difference for our present consideration. Whereas each origin in B must be activated to replicate all of the DNA in the set, the activation of only one origin is sufficient for that purpose in A. The time required to replicate all DNA in A can therefore be controlled by regulating the distribution of those origins which are activated, without

\[ A \]
\[ \cdots \quad o_1 \quad o_1 + 1 \quad o_1 + 2 \quad o_1 + 3 \quad \cdots \]

\[ B \]
\[ \cdots + \quad + \quad + \quad + \quad + \quad + \cdots \]

Figure 1. Two schemes for replicating a large amount of DNA in a short time.
recourse to any change in the rate of fork movement. Thus, if a unit of DNA replication is defined as that segment which is replicated by the forks which emanate from a single origin, then the length of a unit, though fixed in B, can be varied in A.

The replicating DNA in both classes of Drosophila nuclei examined here exhibits the configuration given in A. In spite of the great difference in S phase between these classes, no significant difference in the rate of fork movement was observed between them. Distinct differences in the distributions of active origins were, however, observed. These observations have been used to construct a general model for the topography of active origins during S phase which accounts for its duration in the different nuclei.

Methods

Electron Microscopic Measurements on DNA from Cleavage Nuclei

Fertilized eggs of D. melanogaster (Oregon R) laid over a 40-min interval at 25°C were collected and incubated for an additional 15 min before preparation of cleavage nuclei. Egg collection, preparation of cleavage nuclei, and isolation of the chromosomal DNA by equilibrium sedimentation in CsCl gradients are described in another article (Kriegstein and Hogness, 1973), which also contains a description of the methods used to prepare the DNA for electron microscopy. Contour length measurements were made with a Hewlett-Packard 9864A Digitizer and 9810A Calculator with a fully smoothed, length-calculation program giving an accuracy of ±0.5%. Lengths in kb units were obtained by comparison to the lengths of reference phage DNAs. The single- and double-stranded references were M13 (8.6 kb; Marvin and Hohn, 1969) and PM2 (9.9 kb—determined from the PM2/λ length ratio of 0.213 and the 46.5 kb length of λ DNA; Davidson and Szybalski, 1971).

Radioautographic Measurements on DNA from Cell Cultures

Cell line and culture. D. melanogaster cells of Schneider’s line #2 were grown at 25°C in Schneider’s medium (Grand Island Biological Co. #172) supplemented with bactopeptone and 15% heat-inactivated fetal calf serum, as described in Schneider (1972), except that spinner as well as surface cultures were used. The cells were diploid, contained two X chromosomes, and exhibited a doubling time of about 24 hr at 25°C.

[3H]thymidine labeling of the DNA. Asynchronous cultures in exponential growth were pulse labeled by the addition of [methyl-3H]thymidine (45 to 56 Ci/m mole; New England Nuclear) to a concentration of 200 μCi per ml of medium. Incorporation of the [3H]thymidine into DNA proceeded linearly for a period of at least 2 hr with no detectable lag. In most cases, the [3H]thymidine was added directly to the above culture medium, and the pulse terminated by addition of 2 volumes of ice-cold medium supplemented with unlabeled thymidine at 0.1 mM. In a few early experiments, the culture medium was replaced by a modified medium (the culture medium minus the yeast hydrolysate, bactopeptone, and serum) containing the [3H]thymidine. No significant differences in the radioautographic patterns given by the two methods of labeling were observed, and the data from both have been combined here. A [3H]thymidine pulse can be effectively chased (i.e., all tracks have sharp ends and do not exhibit the “tailing” described in Huberman and Riggs, 1968, for pulse-chase experiments with Chinese hamster cells) by replacing the labeled culture medium with fresh culture medium supplemented with thymidine at 0.1 mM.

Radioautography. Labeled cells from spinner cultures were rinsed twice with ice-cold Schneider’s medium (without the bactopeptone and serum) by centrifugation in a Sorvall HB4 rotor at 1000 rpm for 5 min and resuspending to 2–6 × 10⁶ cells per ml. Surface cultures were directly suspended in the ice-cold medium. Cells (3 ± 2 × 10⁴ per slide) were lysed, and their DNA spread on subbed slides according to the method of Lark et al. (1971), except that the lysis medium consisted of 2% sodium dodecyl sulfate in 2 mM Na EDTA. Dry slides were fixed for 30 min in 5% cold TCA and rinsed in 95% ethanol, after which they were coated with AR-10 (AR-10) in Kodak D-19 at 20°C Grain tracks were developed for 2 min (NTB-2) or for 5 or 10 min (AR-10) in Kodak D-19 at 20°C. Grain tracks were observed under oil without a cover glass with a Zeiss Planapo (40 ×) using either bright or dark field illumination, and fields of approximately 300 μ were photographed with Panatomic X film for analysis.

Grain track measurements. Photographic negatives were projected onto the Hewlett-Packard Digitizer, and the lengths of grain tracks and gaps between tracks in the linear arrays measured as described for the electron micrographs. Distances in microns were obtained by normalization to a micron grating photographed under the same conditions and then converted to kb by dividing by 0.34. The terminal grain tracks in a linear array were never included in the analysis, as they may
contain broken ends. Grain densities were determined to set criteria for the following three classes of tracks: (a) one daughter duplex containing one labeled strand (density = \(D\)); (b) a pair of such daughter duplexes created by a single fork and drawn together when the DNA is stretched during spreading (density = \(2D\)); and (c) aggregated DNAs (density > \(2D\)), which were eliminated from the sample. To be scored, gap lengths had to be greater than 1 \(\mu\) and 2 \(\mu\) for \(2D\) and \(1D\) tracks, respectively. No significant differences in the distribution of grain-track lengths and of distances between the centers of adjacent tracks were detected after 2, 3, or 4 months of exposure.

**Results and Discussion**

**Electron Microscope Observation of Replicating DNA from Cleavage Nuclei**

*The structure of “eye forms” and the mechanism of replication.* When the DNA isolated from cleavage nuclei is examined in the electron microscope, molecules containing multiple eye forms, such as that shown in Figure 2, are observed. Since the structure of these eye forms and the analysis of the single-stranded regions in their forks are reported in detail elsewhere (Kriegstein and Hogness, 1973), only a brief summary of these subjects is given in this section. That the eye forms result from DNA replication is indicated by three kinds of observations. First, the two segments that form an eye are invariably of the same length and are double-stranded, except for a small region at the fork which is described in the next paragraph. Second, the two segments in each eye exhibit partial denaturation maps that indicate they contain related sequences of base pairs. These two characteristics are those expected for replicated daughter segments. Finally, the frequency of eye forms exhibited by DNA isolated from the slowly replicating nuclei of cell cultures is less than that observed for cleavage nuclei by about two orders of magnitude.

Most forks (63%) have single-stranded regions associated with them (Fig. 3A, B). These single-stranded forks can be divided into the same two classes (SSG\(_1\) and SSG\(_2\)) observed for the replication forks in \(\lambda\) and T7 DNAs (Inman and Schnö, 1971; Wolfson and Dressler, 1972). For this reason, and because they have the configuration and orientation

---

**Figure 2.** Replicating chromosomal DNA from cleavage nuclei. The portion of the chromosomal molecule shown here is 110 kb in length and contains 23 eye forms. This electron micrograph is reproduced from Kriegstein and Hogness (1973), which should be consulted for the methods and conditions.
**Figure 3.** The structure of replication forks in the chromosomes of *D. melanogaster* (Kriegstein and Hogness, 1973). The percentages in (A), (B), and (C) derive from a population of 360 forks that were examined in the electron microscope after spreading DNA from cleavage nuclei in 40% formamide to allow visualization of both single- and double-stranded DNA. In (A), whiskers are shown to arise from forks with one single-stranded gap (SSG1) by branch migration that is assumed to occur while preparing the DNA. The lengths of the single-stranded regions in SSG1 and whisker forks yield equivalent distributions with means of 0.22 kb and 0.21 kb, respectively. The arrows in the forks indicate the 3'-OH termini in the newly replicated strands. The fact that the 3'→5'-specific *E. coli* exonuclease I destroys the whiskers confirms the orientation given in the figure. This orientation is also compatible with the property of DNA polymerases to add nucleotide subunits only to 3'-OH termini, which in turn forces synthesis to be discontinuous in the daughter, where chain elongation proceeds away from the fork. The SSG1 forks are a prediction of such a discontinuous synthesis, as are forks with two single-stranded gaps (SSG2). The ratio of SSG2 to SSG1 forks will depend on the rates of fork movement and of chain extension and will be small, as observed here, if the fork movement is much slower than chain extension. The percentage in (D) refers to the fact that all of the 37 eye forms in which both forks were SSG1 exhibited the *trans* configuration shown. These 37 are part of a total population of 128 eye forms in which each daughter segment could be unambiguously traced from one fork to another. The percentages for the different classes of eye forms given in the text derive from this population.

Eye lengths and eye-to-eye distances. The histograms given in Figure 4 indicate the distributions of eye lengths and of eye-to-eye distances. The mean length of 439 eye forms is 4.1 kb and the mean of 316 eye-to-eye distances is 9.7 kb. The eye lengths were taken from the same population of molecules used for the eye-to-eye distances and were therefore limited to those molecules containing multiple eye forms. When all scorable eye forms in each field were measured, the length distribution of 279 eye forms was very similar to that shown in Figure 4 but exhibited a slightly higher mean of 5.6 kb.

**Mean origin-to-origin distance.** The distance between origins in any given segment may be different from the observed eye-to-eye distance; (see legend, Fig. 3) predicted from models which are based on the properties of DNA polymerases, DNA ligases, and Okazaki fragments, they are presumed to represent replication forks.

An analysis of the distribution of forks in the eye forms indicates that the eyes are created by bidirectional replication. All eye forms in which both forks are SSG1 exhibit the *trans* configuration shown in Figure 3D. The observed frequencies of eye forms in which (a) both forks contain single-stranded regions (42%), (b) one fork is single-stranded and one is all duplex (41%), and (c) both forks are all duplex (16%) are not significantly different from the frequencies expected for independently occurring pairs of forks; i.e., the expected frequencies are (a) (0.63)^2, or 40%, (b) 2(0.63)(0.37), or 47%, and (c) (0.37)^2, or 14%, respectively. These are the characteristics expected if both forks in an eye are replication forks.

**Figure 4.** Distributions of eye lengths and eye-to-eye distances in the replicating DNA from cleavage nuclei. The eye length is the measured contour length of either daughter segment in an eye. The eye-to-eye distance is measured from the center of one eye to the center of the adjacent eye.
Figure 5. The distribution of a segment of chromosomal DNA in a random population of cleavage nuclei representing all stages in the 10-min nuclear cycle. The single line at the top of the figure (zero time) represents the segment in nuclei which have just divided. Replication of the segment begins in nuclei at $t_1$ and is just completed in nuclei at $t_2$, so that nuclei at later times in the cycle contain two copies of the segment.

either because there are origins in the segment that have not been activated by the time of sampling but would be activated later or because two or more eye forms have merged before sampling. The effect of these factors is depicted in Figure 5 between times $t_1$, when the first origins in a segment are activated, and $t_2$, when the replication of the segment is complete. It is seen that the eye-to-eye distance at first decreases, as more origins are activated and no mergers take place. At later times, the effect of mergers becomes dominant, and the eye-to-eye distance should increase from a minimum value. It is this minimum value which best approximates the origin-to-origin distance in the segment.

We have obtained this minimum value by analyzing the eye-to-eye distances in the following manner. The data for the eye-to-eye distances was grouped into ten classes according to the fraction of the segment containing the eyes that had been replicated. That fraction equals the sum of the eye lengths in the segment divided by the total length of the segment. The range between 0 and 100% replication was thereby divided into 10 equal class intervals and the mean value for the eye-to-eye distance determined for each. These mean values have been plotted against the midpoint of their respective class intervals to give the upper curve shown in Figure 6. The mean values drop precipitously to a plateau, which is reached when about 20-30% of the segment is replicated. This plateau value is maintained until about 70 to 80% replication, when the mean values rise again. The simplest interpretation of this curve is that initiation events are highly synchronized and occur during the first 30% of replication. The eyes formed by these events then grow bidirectionally, with little merging until ca. 70% replication, when mergers become a dominant factor. The plateau value of 7.9 kb would then represent the mean origin-to-origin distance.

Rate of fork movement. If we adopt the above interpretation, that the plateau represents a period when no new eyes are initiated and no existing eyes merge, then the linear increase in mean eye length observed during this period (Fig. 6) can be used to determine the average rate of fork movement. It is necessary to make a transformation to units of time on the abscissa of Figure 6 for this determination. The diagram in Figure 5 illustrates this transformation. The line at the top of the diagram represents a segment in the chromosomal DNA just after nuclear division. Activation of origins first occurs at time $t_1$, and replication is completed by time $t_2$, yielding two segments per nucleus in place of one until the nucleus divides at 10 min (Rabinowitz, 1941). We are interested in calculating the time it takes to replicate the segment, or $(t_2 - t_1)$, since this will represent the interval 0-100% replicated on the abscissa of Figure 6.

The population of segments examined in the electron microscope derive from a random population of cleavage nuclei in which nuclei from any given time interval in the nuclear cycle are equally represented. The fraction of segments which are observed to contain one or more eye forms, $f_e$, is therefore given by

$$f_e = (t_2 - t_1)/(t_2 + 2(10 - t_2)),$$

from which we derive

$$(t_2 - t_1) = f_e(20 - t_2). \quad (1)$$

The value of $f_e$ for a random population of segments having the same mean length as that for the population used to obtain the data in Figure 6 (see legend Fig. 6) is 0.07. Substituting this value in equation (1) yields

$$(t_2 - t_1) = 1.4 - 0.07t_2. \quad (2)$$
Figure 6. Mean eye-to-eye distances and mean eye lengths as a function of the fraction of the segment that is replicated. The data are the same as that given in Figure 4, but classified according to the extent of replication in each segment that was examined. The fraction of a segment that is replicated (upper scale on the abscissa) is defined as the sum of the eye lengths divided by length of the segment, where the boundaries of the segment are the midpoints of the two terminal eyes (i.e., one-half of the length of each of the two terminal eyes and the entire length of all other eyes are included in the sum of eye lengths). The segments were divided into ten equal class intervals according to the percent replicated, and the mean value for each class was plotted at the midpoint of the interval. The number of eye lengths contributing to the mean for each class varied from 21 to 66 and averaged 44; the number of eye-to-eye distances contributing to each class mean varied from 13 to 51 and averaged 32. The mean length of all segments was 25 kb. This is less than the mean length of 90 kb observed for the molecules in the preparation, because the boundaries of the field in the electron microscope were allowed to restrict the size of the segments that were analyzed here.

Since \( t_2 \) is assumed to be less than or equal to the interphase interval, or 3.4 min (Rabinowitz, 1941), then the minimum value for \( t_2 - t_1 \) is 1.2 min. The maximum value for \( t_2 - t_1 \) is 1.3 min and is calculated from equation (2) by substituting \( t_2 = 1.2 \), since \( 1.2 \leq t_2 - t_1 \). This small range of values for \( t_2 - t_1 \) allows us to use its midpoint, or 1.25 min, as the time interval covered by the abscissa in Figure 5 without introducing errors greater than those involved in determining \( f_e \) or the nuclear doubling time (Rabinowitz, 1941).

The rate at which the length of the eye forms increases during the plateau period can then be determined from the slope of the eye length-versus-time curve (Fig. 6), which is 5.3 kb \cdot \text{min}^{-1} (determined by method of least squares over the plateau interval). Since the eyes expand bidirectionally, the rate of fork movement is one-half this value, or 2.6 kb \cdot \text{min}^{-1}.

Given the fork rate and the time required to replicate a segment, one can then arrive at an independent estimate of the mean origin-to-origin distance, as this should equal twice the fork rate multiplied by the replication time. The resulting value of 6.6 kb is in reasonable agreement with the plateau value of 7.9 kb exhibited in Figure 6.

**Distribution of active origins in cleavage nuclei.** According to the arguments given in the preceding sections, the distribution of origin-to-origin distances will be closely approximated by the distribution of eye-to-eye distances in segments that fall into the plateau region observed in Figure 6. The eye-to-eye distributions for segments which have experienced 30(±10), 40(±10), 50(±10), and 60(±10)% replication are given in Figure 7A-D, and Figure 7E gives the distribution for a range of segments that includes most of the plateau region (30 to 70% replicated). The striking characteristic of these distributions is the periodic nature of the modes that are observed. Five significant modes appear at 3.5, 7.5, 10.5, 13.5, and 16.5 kb in Figure 7E and are usually seen in the four distributions with smaller samples. These modal values exhibit a good fit to the integral multiples of 3.4 kb: i.e., 6.8, 10.2, 13.6, and 17.0 kb. There appears to be a sixth mode about 3 kb from the fifth, and there is a suggestion of a seventh mode about 3 kb from the sixth.

These curves indicate that the origins are not randomly distributed but tend to be spaced at distances equal to 3.4a kb, where \( a \) is an integer. One explanation for this spacing is that specific sequences determine the positions of origins, that these origin-specific sequences tend to be spaced at a distance of 3.4 kb, and that the probability, \( P \), that a given sequence is activated to form an origin is less than one. The probability of activation can be estimated from the mean eye-to-eye distance for those distributions in which each eye is derived from a single origin (i.e., before mergers begin) according to

\[
P = \frac{(3.4)}{\text{mean eye-to-eye distance}}.
\]

The mean for the plateau region is essentially constant at 7.9 kb (Fig. 6), from which we estimate \( P = 0.43 \). However, the mean steadily decreases with increasing replication in the preplateau region (0 to 30% replication), indicating the progressive increase in \( P \) that is given in Table 1.

In this model, the frequency of origin-to-origin distances in the first peak (\( a = 1 \)) will be \( P \), that in
Figure 7. Distribution of eye-to-eye distances in the plateau period. The eye-to-eye distances in segments that were classified with respect to the extent of replication as described in Figure 6 are plotted in (A) through (D) for different parts of the plateau period, and in (E) for the entire period, using 1-kb class intervals. Although the data for adjacent distributions overlap, those used in (A) are independent from those in (C), and those in (B) are independent from those in (D). Eye-to-eye distances greater than 25 kb, which are not shown, are: (A) 30–32 kb, 2.0%; (B) 30–32 kb, 1.1%; (C) 32–34 kb, 1.2% and 38–40 kb, 1.2%; (D) 32–34 kb, 1.3% and 38–40 kb, 1.3%; (E) 30–32 kb, 32–34 kb, and 38–40 kb, 0.6%. Eye lengths in segments from the postplateau period are plotted in (F), where the distribution for the total population is also shown (taken from Fig. 4) for comparative purposes. Eye lengths at 26–28 kb (2.1%), 28–30 kb (6.3%), 36–38 kb (2.1%), and 50–54 kb (2.1%) in the 90 (+10)% replicated segments are not shown.
Table 1. Frequency Distributions of Origin-to-Origin Distances During Replication

<table>
<thead>
<tr>
<th>Percent Replication</th>
<th>Probability P</th>
<th>1st Peak</th>
<th>2nd Peak</th>
<th>All Other Peaks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>obs. a</td>
<td>calc.</td>
<td>obs. a</td>
</tr>
<tr>
<td>Preplateau</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 (±10)</td>
<td>0.220</td>
<td>0.220</td>
<td>0.220</td>
<td>0.20</td>
</tr>
<tr>
<td>15 (±15)</td>
<td>0.280</td>
<td>0.280</td>
<td>0.280</td>
<td>0.23</td>
</tr>
<tr>
<td>20 (±10)</td>
<td>0.343</td>
<td>0.343</td>
<td>0.343</td>
<td>0.20</td>
</tr>
<tr>
<td>25 (±5)</td>
<td>0.382</td>
<td>0.39</td>
<td>0.382</td>
<td>0.20</td>
</tr>
<tr>
<td>Plateau</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30, 40, 50, &amp; 60</td>
<td>0.430</td>
<td>0.43</td>
<td>0.430</td>
<td>0.30</td>
</tr>
</tbody>
</table>

* The observed values are taken from the eye-to-eye distances in the indicated distributions, which for the plateau region, are shown in Figure 7A–D.

b These values give the range of frequencies observed for the four distributions in the plateau region, Figure 7A–D.

The second peak, \( P(1 - P) \), and generally, that in the \( n \)th peak, \( P(1 - P)^{n-1} \). We have calculated the expected frequencies for the first and second peaks and the sum of the frequencies for all other peaks \( [1 - P - P(1 - P) = 1 - 2P + P^2] \) in the preplateau and plateau distributions and compared these values with the observed values in Table 1. There is reasonable agreement between the observed and calculated values, although in the plateau region, the second peak is consistently larger than expected, and the third peak (Fig. 7) smaller. Part of this discrepancy may be due to a low frequency of mergers taking place during the plateau period.

It is evident from the relative constancy of the eye-to-eye distributions in the plateau period (Fig. 7 and Table 1) that mergers, like activations, are infrequent between 30 and 60% replication. However, the distributions of eye-to-eye distances for segments exhibiting greater extents of replication indicate that mergers rapidly become a dominant factor. There is a shift toward greater distances, which increases the mean (Fig. 6), and a loss or "blurring" of the peaks present in the plateau distribution; indeed, it is this blurring effect of mergers which makes it difficult to recognize the 3.4-kb spacing in the distribution of eye-to-eye distances shown in Figure 4.

These considerations lead to the expectation that the lengths of many or most eyes in postplateau segments will be determined by mergers. Thus, (a) two eyes from origins that are 3.4 kb apart will yield an eye that is 6.8 kb in length at merger; (b) three eyes from adjacent origins separated by two 3.4-kb spacings will yield an eye 10.2 kb in length; (c) four eyes from origins separated by three 3.4-kb spacings will yield an eye 13.6 kb in length, as will two eyes from adjacent origins that are 6.8 kb apart (this should occur only late in replication of the segment, since it will take 1.3 min at a fork rate of \( 2.6 \text{ kb} \cdot \text{min}^{-1} \cdot \text{fork}^{-1} \); and (d) five eyes from origins separated by four 3.4-kb spacings will yield an eye 17.0 kb in length, as will three eyes from adjacent origins spaced 3.4 and 6.8 kb apart. These four length classes, which are also multiples of 3.4 kb, are closely approximated by the positions of peaks in the distribution of eye lengths for segments that have been 90±10% replicated (Fig. 7F).

The largest peak, at 4.5 kb in Figure 7F, is the expected result of eye growth from single origins, with no mergers. The synchrony of activation, which limits the formation of most if not all origins to the range 15 ± 15% replication, and the bidirectional growth of eyes with a fork rate of \( 2.6 \text{ kb} \cdot \text{min}^{-1} \cdot \text{fork}^{-1} \) should yield a peak of eye lengths at 4.9 kb; i.e., \((0.9 - 0.15)(1.25 \text{ min})(2 \text{ forks})(2.6 \text{ kb} \cdot \text{min}^{-1} \cdot \text{fork}^{-1}) = 4.9 \text{ kb} \). Were there no mergers, this would be the only peak expected; mergers decrease both the frequency of this peak and its modal value. We can estimate that this peak contains only 18% of all origins in this distribution and that about 81% of the origins are in eyes that result from mergers (the remaining 3% are in the eyes of the small peak seen at 1.5 kb in Figure 7F and may represent a low frequency of activation that occurs during the plateau and postplateau periods).

The distribution of eye lengths in segments that exhibit progressively lesser extents of replication are consistent with these interpretations. The frequency of origins in the peak attributable to eyes that have not merged increases, and its modal value decreases with decreasing extent of replication; whereas the frequency of the length classes that can be attributed to mergers progressively decreases. For instance, at 40(±10)% replication, the peak of nonmerged eyes contains 89% of the origins and has
DNA REPLICATION IN CHROMOSOMES

a modal value between 1.5 and 2.5 kb, whereas the 7.5- and 13.5-kb length classes that result from mergers contain only 7 and 4% of the origins, respectively.

Comparison of the number of activable sequences in cleavage nuclei with the number of chromomeres. The data presented in the preceding section are qualitatively consistent with the model that the activable, origin-specific sequences are repeated along the chromosomal DNA with a spacing that is rather tightly centered about a mean value of 3.4 kb, and that, in cleavage nuclei, these sequences are randomly activated to produce origins with a probability that approaches one-half over a short time interval (ca. 0.4 min). The haploid genome of 165,000 kb (Rudkin, 1964, 1972; Rasch et al., 1971) would then contain about 50,000 of these sequences. This is an upper limit, since our data are equally compatible with the possibility that the distribution of activable sequences is directly represented by the eye-to-eye distributions in Figure 7A–E (a model for this possibility is presented further on in this paper). In this case, the activation probability would approach one, and the mean distance between these sequences would then be 7.9 kb. This leads to a minimum number of 20,000 activable sequences per haploid genome in cleavage nuclei.

The slightly more than 5000 chromomeres, counted by Bridges as bands in polytene chromosomes, account for 78% of the haploid genome (Rudkin, 1972). The number of activable sequences in this fraction of the genome is therefore between 16,000 and 40,000, or an average of three to six sequences per chromomere. If there is validity to the notion that individual chromomeres correspond to individual units of replication (see the review by Rudkin, 1972), then it would appear that the chromomeric structure somehow restricts the number of these sequences that are available for activation and that this aspect of the chromomeric structure is absent from cleavage nuclei.

In regard to the possibility that the class of origin-specific sequences is divisible into activable and nonactivable subclasses according to the chromosomal state, we note that among the more than one-thousand eye forms that we have examined, no instance of an eye within an eye—i.e., the reactivation of an origin within a daughter segment—has ever been observed. This observation indicates that once replicated, a DNA sequence cannot be activated to an origin until the next cycle of replication. Such a restriction could result from changes during the cycle in the DNA itself (e.g., methylation), or, more generally, in the chromatin structure. One can also imagine that this restriction is due to the absence of one or more activation factors, except during a narrow time interval at the beginning of replication. However, we have observed small eyes (e.g., 0.5 kb) next to larger eyes (e.g., 3.3 kb) where the eye-to-eye distance was about equal to the unit spacing of 3.4 kb. This topography is most simply explained by the formation of the origin creating the small eye after the origin-specific sequence in the larger eye has been replicated.

We reserve a more detailed consideration of the concept of activable and nonactivable sequences until after the following data on DNA replication in cell culture nuclei has been presented.

Radioautographic Observations of Replicating DNA from Nuclei in Cell Cultures

Experimental plan. Although the average DNA segment is replicated in 1.2 to 1.3 min in cleavage nuclei, the S phase appears to include most, if not all, of the estimated 3.4 min of interphase (Rabinowitz, 1941). This conclusion is inferred from the observation that the largest eye-to-eye distance that is consistently present in distributions where it is unlikely to have arisen from mergers or to be erased by the formation of new origins is 19.5 kb (Fig. 7B, C); and the time required to replicate this interval between two origins at the average fork rate is 3.7 min. The 600 min devoted to S phase in cell cultures (Dolfini et al., 1970) at the same temperature (25°C) is therefore about 2 x 10^2-fold larger than in cleavage nuclei.

This increase in S phase could result from (a) a decrease in the fork rate, (b) an increase in the distances separating adjacent origins, and/or (c) an increase in the time period during which the origins are activated. In this initial study of the factors controlling the overall rate of chromosomal DNA replication in Drosophila, we shall compare the fork rates and the distribution of origins in the two kinds of nuclei. A cursory examination of the DNA from cell culture nuclei in the electron microscope revealed a very much lower frequency of replicating forks than in DNA from cleavage nuclei. This observation indicated that a decreased fork rate cannot be the sole cause of the increased S phase and that electron microscopy would be too time-consuming a method for the analysis of this DNA. We therefore turned to radioautography, whereby the replicating DNA can be selectively examined (Fig. 8).

The problems associated with the analysis of the linear series of grain tracks produced in the radioautographs are summarized in Figure 9. Proceeding from left to right in this figure, we see that individual grain tracks formed from eyes which exist at
Figure 8. Linear arrays of grain tracks obtained from pulse-labeled cells. Cells of *D. melanogaster* in spinner culture at 25°C were pulse labeled for 5 min with [3H]thymidine and radioautographs obtained as described under Methods. The linear white regions in this reverse contrast photograph represent clusters of silver grains, i.e., grain tracks. Single silver grains which occur occasionally in the background are barely visible in this print. The AR-10 film was exposed for 3.3 months, developed for 5 min, and photographed under bright-field illumination.

the beginning of the labeling interval and do not merge during that interval are due to single forks; tracks from eyes that are created during the interval and that do not merge result from two diverging forks; and tracks which result from the merger of two eyes during the interval are formed by two converging forks. Although techniques are available for sorting out these three kinds of tracks (i.e., one-fork tracks and two-fork diverging and converging tracks) by the use of two successive pulses of [3H]thymidine of two different specific activities (Huberman and Riggs, 1968; Prescott and Kuempel, 1972; Callan, this volume), these techniques require longer track lengths, which reduce the resolution below the levels we desire here. Furthermore, it is clear that the formation of a given track may be considerably more complex than is indicated in Figure 9.

We therefore chose a method of analysis in which the principle is to reduce the heterogeneity of the track population—ideally to a one fork-one track population. The technique is to examine the tracks produced by progressively smaller pulse intervals applied to asynchronous cultures in exponential growth. As the pulse interval approaches zero, the fraction of tracks containing an origin or a merger will decrease, and the fraction resulting from the action of one fork will increase. Hence, the mean track length divided by the pulse time will approach the mean fork rate as the pulse interval

Figure 9. Diagram illustrating how a single grain track can arise from one fork or from two forks which are either diverging or converging.
Figure 10. Distributions of grain track lengths and of center-to-center distances for various pulse times. Cells in surface cultures were labeled for the indicated time and radioautographs obtained as described under Methods. The center-to-center distances are the distances between the centers of adjacent grain tracks in a linear array. The class interval in all of the histograms is 6 kb.

approaches zero. Similarly, the mean center-to-center distance between adjacent tracks will approach the mean fork-to-fork distance, which is one-half the mean eye-to-eye distance in the asynchronous cell population.

Fork rate. The distributions of grain track lengths for pulse times of 5, 15, 30, and 75 min are given in Figure 10. When the mean track length for each distribution is divided by the pulse time, and this ratio plotted against the pulse time, the curve given in Figure 11A is obtained (the ratios obtained from five additional experiments are also plotted in this figure as open circles). As the pulse time is decreased from 75 min, the ratio increases more and more rapidly, reaching an average value of 2.6 kb min⁻¹ at the minimum pulse time of 5 min. Although a linear extrapolation to zero time yields a value of about 3.2 kb min⁻¹, the expected nature of the extrapolation (to be discussed later) is indicated by the dotted line in the figure. Hence, we are limited to the statement that the fork rate in cell culture nuclei is equal to or greater than 2.6 kb min⁻¹ fork⁻¹; or, stated in another way, the fork rate in these nuclei is equal to or greater than that in cleavage nuclei.

Eye-to-eye distance. The distribution of the center-to-center distances for each of the four pulse times is also given in Figure 10, and a plot of the mean center-to-center distance versus pulse time is given in Figure 11B. Extrapolation to zero time yields a value of about 20 kb, which represents the estimated mean fork-to-fork distance. Twice this value, or 40 kb, represents the mean eye-to-eye distance for the replicating eye forms scored in the radioautographs. This distance is only four times that observed in cleavage nuclei.

Clustered origins. According to the above estimates of fork rate and eye-to-eye distance, it would take less than 8 min to replicate a segment of DNA in which the spacing between adjacent origins equaled the observed mean eye-to-eye distance, a time that is about two orders of magnitude less than the $S$ phase for these nuclei. An explanation for this surprising result is obtained if it is assumed that the origins leading to the tracks which we score are clustered and that the distance between adjacent clusters is so large that center-to-center distances between tracks from different clusters are seldom scored.

A cluster is defined as a set of adjacent origins which are activated at approximately the same time and exhibit a mean spacing such that individual eyes derived from these origins merge to form a giant eye within an interval that is small
merge within the pulse interval. Some simple conclusions about the distribution of tracks can be formulated on the basis of these rules and the fact that the growing population of cells is asynchronous, so that clusters are being formed by activation and lost by merging at equivalent rates during the pulse interval. These conclusions can be summarized by the following equations (see Appendix, for derivations):

\begin{align*}
\text{a) for } t < d/2r, & & \text{b) for } t \geq d/2r,
\frac{l}{t} = r & & \frac{l}{t} = d/2 \\
\frac{c}{r} = d/2 & & c = d
\end{align*}

where \( t = \) pulse time; \( d = \) origin-to-origin spacing in a cluster; \( r = \) fork rate; \( l = \) mean track length; and \( c = \) mean center-to-center distance. In addition, we note that when \( t \geq d/2r \), the fraction, \( F \), of clusters formed during the pulse or present at \( t = 0 \) which we can score is given by

\[ F = 2d/(d + 2rt), \]

and hence decreases with increasing \( t \). By contrast, when \( t < d/2r \), this fraction is 100%. If the population contains clusters with different values of \( d \), then the smaller the \( d \) value the greater the bias against scoring the corresponding tracks, providing \( t > d/2r \). The relative frequencies of scored clusters will therefore change with increasing \( t \) so as to favor those with larger \( d \)'s, increasing both \( l \) and \( c \).

The theoretical curve given in Figure 11A was constructed according to the above considerations for a homogeneous population in which \( d = 40 \) kb and \( r = 2.6 \) kb \( \cdot \) min \(^{-1} \) \cdot fork \(^{-1} \). For pulse times less than \( d/2r \), or 7.7 min, the mean track length is \( rt \) and, when divided by \( t \), is equal to the fork rate, or 2.6 kb \( \cdot \) min \(^{-1} \). For pulse times \( \geq 7.7 \) min, this ratio of the mean track length to pulse time equals \( d/2t \), or \( 20/t \) for the case at hand, and decreases with increasing time as indicated in the figure. There is reasonable agreement between the theoretical curve and the observed ratio for pulse times of 15 min or less, but for larger pulse times the observed value is larger by a factor which increases with time. This discrepancy is expected for the following reasons.

The actual population is heterogeneous, with \( d \) values that are both smaller and greater than the mean value of 40 kb. As the pulse time is increased, the fraction of clusters that are scored decreases differentially, favoring clusters with larger \( d \) values according to equation (5). This will cause \( l/t \) for the scored population to be larger than \( d/2t \) for the homogeneous population by a factor which increases with pulse time. This expected bias also explains why the plot of the mean center-to-center

---

**Figure 11.** Estimates of the average fork rate and the mean fork-to-fork distance by extrapolation to zero pulse time. (A) The mean grain track length divided by the pulse time was plotted against the respective pulse time. Filled circles represent the data shown in Figure 10. Open circles derive from a set of independent experiments in which spinner rather than surface cultures were used; each point is derived from the measurement of 450 to 1400 grain tracks. (B) The mean center-to-center distances were taken from the distributions shown in Figure 10. The estimate of the average fork rate (A) and the mean fork-to-fork distance (B) by extrapolation to zero time is discussed in the text. The derivations of the equations (3a,b) and (4a,b) used to construct the theoretical curve given by the dotted line (see text) are given in the Appendix.
distance versus pulse time (Fig. 11B) does not simply rise linearly from 20 to 40 kb in the first 7.7 min and thereafter remain constant, as would be predicted from equations 4a, b for a homogeneous population with \( d = 40 \text{ kb} \) and \( r = 2.6 \text{ kb} \cdot \text{min}^{-1} \cdot \text{fork}^{-1} \).

Some idea of the distribution of \( d \) values within the population can be gained by examining the distribution of center-to-center distances that are scored after long pulses (Fig. 10). With a pulse time of 75 min, \( t \geq d/2r \) for all \( d \leq 390 \text{ kb} \), and hence, essentially all of the center-to-center distances shown in the histogram should represent \( d \) values, though not at the frequencies they exhibit in the actual population. Two major peaks, which exhibit mean values of 28 kb and 57 kb, are apparent in the 75-min distribution. These two peaks are also observed in the distribution of center-to-center distances scored after a 60-min pulse (Blumenthal, unpublished experiments). They suggest that two major classes of origin-to-origin distances exist in the cell culture nuclei.

We can estimate the frequencies of these two classes in the actual population by correcting each of the frequencies in the 75-min distribution according to equation (5) and then determining the sum of the corrected frequencies belonging to each class. The resulting sums indicate that the 28-kb class occupies one-half, and the 57-kb class, one-third, of the corrected distribution, a total of 80 to 85% in both classes.

Although the equivalent two peaks are not observed in the 30-min distribution, a broad flat peak extends over the equivalent region and could arise from the overlap of the two classes, particularly when it is expected that each will be scored at approximately equal frequencies after a 30-min pulse. Furthermore, we calculate that this broad peak includes 85% of the distribution after it is corrected according to equation (5), i.e., the fraction it occupies is equivalent to the sum of the two classes indicated by the 75-min distribution. At 5 min, \( t < d/2r \) for both classes, and hence, the center-to-center distances expected for these classes (and for all classes of higher order) should be less than their respective \( d \) values (eq. 4a). And at 15 min, the 28-kb class should be apparent, but the 57-kb class is expected to overlap with center-to-center distances from higher order classes (where \( t < d/2r \)) to produce a tail like that observed in the 15-min histogram.

It therefore appears that a large majority (ca. 85%) of the origin-to-origin distances in the DNA segments which yield the grain tracks that we score fall into two broad classes centered around 28 kb and 57 kb, and these occur in a ratio of approximately 1.5 to 1, respectively. If higher order classes are neglected, then these two classes will yield a mean origin-to-origin distance of 38 kb, just under the value of 40 kb obtained by extrapolation of the center-to-center distances (Fig. 11B). Furthermore, we note from equations 3b and 4b that \( I \) should be one-half of \( \frac{c}{t} \), if \( t \geq d/2r \). Since this condition is fulfilled for both classes when the pulse time is greater than 11 min, then the mean track lengths observed after pulses of 15, 30, and 75 min should be about one-half the corresponding mean center-to-center distances. The values given in Figure 10 are in reasonable agreement with this expectation. For the 5-min pulse, where \( t < d/2r \) for both classes and equations 3a and 4a apply, the expected ratio (0.41) is only a little less than the ratio of one-half that is observed.

There is an assumption that is implicit in the identification of the 28-kb and 57-kb classes as origin-to-origin distances. The distributions resulting from long pulses register eye-to-eye distances, and we have assumed that the origin-to-origin distances are closely approximated by these eye-to-eye distances. This is, of course, the case in the hypothetical clusters used to formulate equations (3) through (5), since the origin-to-origin spacing is uniform, and all origins are activated at the same time (Appendix). It is also true in cleavage nuclei, where the mean eye-to-eye distance is 9.7 kb and the mean origin-to-origin distance is 7.9 kb; in this case, the approximation is good because the period of activation is less than the time required for two converging forks to replicate the DNA between pairs of origins belonging to the shortest spacing class (i.e., 3.4 kb). We should expect the approximation to be valid in cell culture nuclei for classes of origin-to-origin distances equal to or greater than 26 kb if the period of activation for a cluster is less than 5 min. The assumption of such short activation periods appears reasonable in view of the general consistency of the 5- to 75-min distributions analyzed above. We therefore consider that the 28-kb and 57-kb values are equal to or slightly greater than the means for the first two classes of origin-to-origin distances in cell culture nuclei.

**Comparison of the origin spacings in cell culture nuclei and chromomere size.** The 28-kb value for the first class of origin spacings is remarkably close to estimates for the DNA content of the average chromomere in polytene chromosomes of *D. melanogaster* (26 kb to 32 kb per chromatid, Rudkin, 1972; a value of 30 kb will be used here). It is therefore tempting to postulate that the origin spacings are determined by some aspect of chromomeric structure in the interphase chromosomes.
The chromomeric structure can be thought of as a periodic pattern in which the repeated unit consists of a region of highly compacted chromatin (the chromomere) followed (or preceded) by a non-compacted region (the interchromomere). The chromomere contains most of the 30 kb in the average unit of *D. melanogaster*, leaving at most only ca. 2 kb in the interchromomere (Beermann, 1972). If we now imagine that only those origin-specific sequences that are contained in the interchromomers (i.e., in the noncompacted regions) are available for activation, then the intervals between two adjacent origins will divide into classes according to whether they contain one, two, or three, etc., chromomeres. We should then expect that these classes will be observed as classes of origin-to-origin distances that are integral multiples of 30 kb, if the DNA contents of the large majority of chromomeres are close to this mean value.

The data on the distribution of the DNA contents of chromomeres is not extensive, but Rudkin’s (1961) measurements of the total absorbancy at 257 nm for each of 60 chromomeric bands in polytene chromosomes suggests that the distribution exhibits the requisite characteristics. Thus, the DNA content of chromomeres, corresponding to the 60 bands, ranges from 7 kb to 92 kb, with a mean and standard deviation of 32 kb and 21 kb, respectively, and a median value of about 26 kb; only 11 of the 60, or 18%, exhibited values greater than 53 kb. The close correlation between the values for the first two members of the above series (i.e., 30 kb and 60 kb) with the mean values of 28 kb and 57 kb exhibited by the two observed classes is therefore in reasonable agreement with the chromomere model. Classes corresponding to higher orders in the series may be observed if the number of center-to-center distances measured after long pulses is increased above the 229 given in Figure 10.

A General Model for the Distribution of Origins in Cleavage and Cell Culture Nuclei

The chromomere model given in the preceding section to account for the observed distribution of origins in cell culture nuclei implies that origins are formed only in a fraction of the interchromomeres represented in the clusters that we score. This fraction can be estimated to be about 60 (±10)% in several ways. Here, as before in the case of cleavage nuclei, we can imagine that this frequency of origins among the repeated units is due to one of two extreme conditions. Either (1) every interchromomere contains an origin sequence and this sequence is activated with a probability equal to 60%, or (2) the probability of activation approaches one, and only 60% of the interchromomeres contain origin sequences.

We emphasize the second of these conditions because it suggests the following general model for explaining the distribution of origins in both types of nuclei. Imagine that the origin sequence consists of a small number of adjacent base pairs and that these sequences are, for the most part, randomly distributed in the genome with a mean spacing of the order of 1 kb—e.g., a specific sequence of five base pairs should exhibit a mean spacing of 4^5 = 1024 base pairs. The frequency of interchromomeres which will not contain such an origin sequence is expected to be 37%, if each contains 1 kb (Poisson distribution); i.e., 63% will contain one or more origin sequences. Since single and multiple activations within one interchromomere could not be differentiated by radioautography, an activation probability that approaches one applied to this distribution of activable sequences should yield the classes of origin-to-origin spacings that we detect.

To explain the distribution of origins in cleavage nuclei, we suppose that the chromatin of the interphase chromosomes is arranged in a periodic pattern analogous to the chromomeric pattern but with a repeated unit containing only 3.4 kb, divided between the compacted and noncompacted regions (note that such a periodic pattern need not itself be sequence determined—see Lerman, this volume). If, as in the chromomeric pattern, only the origin sequences will determine the frequency of the activable 3.4-kb units; e.g., a specific sequence of 5 base pairs will occur at least once in 45% of the units if the noncompacted region contains 0.6 kb. Again, an activation probability that approaches

---

3 (a) A value of 70% is obtained by dividing the 28 kb of the basic origin spacing by the estimated mean origin-to-origin distance of 40 kb. (b) If the presence of an origin in an interchromomere is randomly determined, then the fraction of interchromomeres containing an origin should equal the frequency of the 28-kb class of origin-to-origin distances, which was estimated to be about 80%. (c) If the clusters we score represent a random sample from the 129,000 kb of DNA containing the chromomeres (78% of 165,000 kb in the haploid genome; Rudkin, 1972), then the number of active interchromomeres can be calculated by dividing the 129,000 kb by the mean origin-to-origin distance of 40 kb. The result is 3200 active interchromomeres, which is 64% of the 5000 interchromomeres.

---

8 The values were calculated from the distribution of total absorbancies given by Rudkin (1961; Fig. 5 of his article), using his more recent calibration that one pg of DNA equals 2.6 μ of total absorbancy (pers. commun.).
The concept that the origin sequences are, for the most part, randomly distributed requires further elaboration. While it is likely that a specific sequence containing a small number of base pairs will be formed by chance arrangements of other selected sequences (e.g., the codons in a gene) in most regions of the genome, this is not possible in the regions that contain the highly repetitious satellite DNAs. These DNAs constitute some 18% of the genome in D. melanogaster and are mostly located in the centromeric heterochromatin (Peacock et al., this volume; Gall et al., this volume).

One must suppose that origin sequences in the segments containing the satellite DNAs are created by direct genetic selection. Although we are only now in the process of determining the distributions of origins in the satellite DNAs from cleavage nuclei, certain predictions can be made. Thus, the distance between origins in the segments of satellite DNAs from cleavage nuclei should be no greater than ca. 20 kb, in order that these segments be replicated within the 3- to 4-min S phase; unless the fork rate is significantly different from 2.6 kb \( \cdot \) min\(^{-1} \cdot \) fork\(^{-1}\)—a possibility that we consider unlikely. One would expect that such a minimum demand on the selection process is also operative in other regions of the genome. Thus, the random formation of origin sequences in these regions must be modified by direct genetic selection, so that no origin-to-origin distances exist which are too long to be replicated in the required time. Indeed, this requirement is one reason we do not favor the hypothesis that origin sequences occur every 3.4 kb but are randomly activated with a probability approaching one-half, since one would then expect to find a significant number of cases during the replication of each chromosomal DNA where the distance between adjacent origins was in the range 23 to 44 kb (e.g., the expected numbers per large chromosomal molecule containing 62,000 kb are: 71 at 7 \( \times \) 3.4 kb = 23.8 kb, 36 at 27.2 kb, 18 at 30.6 kb, 9 at 34 kb, 4 at 37.4 kb, 2 at 40.8 kb, and 1 at 44.2 kb; or a total of 141 per molecule).

It should be pointed out that heterochromatization of the regions next to the centromere (which contain most of the satellite segments) does not appear to occur in cleavage nuclei (Schultz et al., 1967, as cited in Mahowald, 1968; Fristrom, 1970; Chen, 1971), and hence, this kind of compacted structure should not act to occlude origin sequences within the satellite segments in these nuclei. However, centromeric heterochromatization does occur after the cleavage nuclei have migrated to the periphery of the egg and cell membranes have formed around them, and it evidently occurs in all subsequent somatic nuclei, certainly in the cell culture nuclei that we have used. Little is known about the structure of this centromeric heterochromatin during interphase, but we may reasonably suppose that the compacted regions within it are much larger than the chromomerres in the euchromatic arms and that therefore most of the origin sequences that are available in this region in cleavage nuclei (i.e., at least one every 20 kb) will, according to our model, be occluded in cell culture nuclei, and indeed, in most or all somatic cell nuclei. Hence, the model predicts that the origin-to-origin distances within the centromeric heterochromatin will be large in cell culture nuclei, perhaps extending to the maximum distance allowed for an S phase of 600 min and a fork rate of 2.6 kb \( \cdot \) min\(^{-1} \cdot \) fork\(^{-1}\), namely, 3100 kb. We shall amplify this concept in the following section on the determinants of S phase.

### Determinants of S Phase

The equivalence of fork rates in cleavage and cell culture nuclei is striking and indicates that the fork rate is not a variable that is used to alter the time required to replicate the molecules of chromosomal DNA. Callan (1972; this volume) reached a similar conclusion for cells with different S phases in the newt, Triturus. We extend this conclusion to make the simplifying assumption that the fork rate is constant—that once created, all forks in either of these two kinds of nuclei move at a rate of 2.6 kb \( \cdot \) min\(^{-1} \cdot \) fork\(^{-1}\) until they meet an opposing fork.

This assumption implies the absence of the termination sequences postulated by Huberman and Riggs (1968) to lie between each pair of adjacent origin sequences and to act as termini for fork movement. Such termini would require that every origin sequence in a chromosomal DNA molecule be activated to complete its replication. They are inconsistent with our results, which indicate that the distance between activated origins is greater in cell culture nuclei than in cleavage nuclei. Furthermore, if such termination sequences exist, then inversions, translocations, and deletions would frequently result in the absence of an origin sequence between two adjacent termini. One would expect this situation to be lethal; as the segment of DNA between these termini could not be replicated and, as a result, the daughter chromosomes could not separate without breakage.

If the fork rate is constant, then the replication time for a chromosomal DNA molecule must be determined by where and when the multiple origins are activated within the molecule. To simplify our
treatment of this problem, we consider the replication of only the largest chromosomal DNA molecules in *D. melanogaster*, i.e., the 62,000-kb molecule of autosomes 2 or 3 (Kavenoff and Zimm, 1973). Somewhat less than 20% of this molecule, or about 10,000 kb, is contained in the central region which forms the centromeric heterochromatin in somatic cells (Rudkin, 1964), and which we presume consists mostly of the satellite DNAs. The remainder, or about 52,000 kb, is about equally divided between the two arms which form the ca. 2000 chromomeres (Bridges, 1939, 1941a, b, 1942). To a first approximation, we can take the time required to complete the replication of this molecule to equal the estimated S phase of the respective nuclei (3 to 4 min in cleavage nuclei and 600 min in cell cultures).

A 3- to 4-min S phase in cleavage nuclei is satisfactorily explained by our data which indicates that essentially all origins are activated during a very short period early in S (ca. 0.4 min) and that the length of S is determined by a maximum spacing between origins of about 20 kb which is replicated by two converging forks at a rate of 2.6 kb ⋅ min⁻¹ ⋅ fork⁻¹. According to our model, the factors required for activation are present in sufficient amounts to cause the probability of activation to approach one very early in S, so that essentially all available origin sequences (52,000 kb/7.9 kb = 6600 available sequences in two arms and at least 10,000 kb/20 kb = 500 available sequences in the central satellite region) are activated over the short time period of ca. 0.4 min. Once replicated, some mechanism (as yet unspecified) must prevent an origin sequence from being activated until the next S phase.

Our data are not sufficient to explain the 600-min S phase exhibited by cell culture nuclei, and we must make use of observations from other sources for this purpose. In accordance with the chromomeric model given in the preceding section, we presume that the clusters we score derive from the two arms where the DNA is in the chromomeric pattern. Whether each arm acts as a single cluster or as a set of clusters will depend upon whether activation occurs at different times for different parts of an arm, and/or upon the maximum distances between adjacent interchromomeres that contain origin sequences.

We argue that the activation probability approaches one during a short interval early in S for all origin sequences within all interchromomeres in the two arms on the basis of the following two lines of evidence. It is a common observation that the so-called “continuous label” pattern, in which practically all of each chromomeric arm in a single polytene nucleus is radioautographically labeled, is found after the administration of short pulses of [³H]thymidine (e.g., 15 min; Howard and Plaut, 1968; see also review by Rudkin, 1972). If different clusters within an arm were activated at times which differ by more than the pulse time, then this pattern could result only if most forks in a cluster were active during a time interval that is equal to or larger than the difference in time between the activation of the first and last clusters. However, our data indicate that the large majority of forks in a cluster are active only during short time periods. Thus, we estimate that about 85% of the origin-to-origin distances fall into the 28-kb and 57-kb classes, and it takes 5.4 and 11 min, respectively, to replicate the segments of DNA between origins in the two classes. Since activation is essentially synchronous within a cluster (i.e., occurs in less than 5 min), then all clusters should be activated during a time period of the order of 15 min or less; assuming, of course, that the data from the two sources can be coupled in this way. Given this assumption, we conclude that the chromomeric arms are not divided into clusters on the basis of different times of activation.

We then ask whether the arms may be divided into clusters by large distances between adjacent activated interchromomeres and, more specifically, whether any of these distances approaches the 3100 kb that is allowed by a 600-min S phase. If we assume that, for the most part, the activated interchromomeres (i.e., those containing one or more origin sequences) are randomly distributed among the 2000 interchromomeres in the arms with a probability per interchromomere of 0.6 (the value estimated in the preceding section), then the expected number of cases in which two active interchromomeres are separated by one, two, three, etc., chromomeres (or by an average of 30 kb, 60 kb, 90 kb, etc.) is: for one (or 30 kb), 720 cases; for two (or 60 kb), 290; for three (or 90 kb), 115; for four (or 120 kb), 46; for five (or 150 kb), 18; for six (or 180 kb), 7; for seven (or 210 kb), 3; and for eight (or 240 kb), 1. All of these average distances would probably be scored within a cluster (e.g., the longest center-to-center distance we measured, after the 75-min pulse, was 490 kb) and would require less than 50 min for replication. If there were a high proportion of the larger chromomeres in the higher order spacings, then some might be large enough not to be scorable and, therefore, to function as spacers between clusters. However, the distribution of the DNA contents of chromomeres referred to previously (Rudkin, 1961; see also reviews by Beermann, 1972; Rudkin, 1972) and the topography of dense and light bands given by Bridges’ maps (1939, 1941a,b, 1942) make it highly improbable that such spacers could be longer than 1000 kb, or require more than 200 min for their replication.
Hence, the picture we have constructed by coupling the results from our radioautographic measurements with the data on polytene chromosomes indicates (a) that all of the units of DNA replication in the chromomeric arms are activated at the beginning of S, (b) that most units complete their replication quite early in S, and (c) that even the few longest units, and therefore the entire arms, should complete their replication within the first one-third of S. This would mean that replication during the last part of S should be confined to the ca. 10,000 kb within the centromeric heterochromatin. This conclusion is consistent with the radioautographic observations of Barigozzi et al. (1966) and Dolfi (1971) on metaphase chromosomes from embryonic and cultured cells labeled during S phase. Their results indicate that DNA replication continues in the centromeric heterochromatin of the two large autosomes (as well as in the heterochromatic regions of the other chromosomes which includes most of the Y chromosome) late in the S phase, well after replication of the arms is complete.

The questions that remain concern the distribution of origin sequences within the heterochromatic DNA that are available for activation and when these sequences become available. The extreme situation in which no sequences in the heterochromatic DNA are activable can evidently be ruled out in these cells, since the time required to replicate the 10,000 kb in this region by two converging forks emanating from the two closest activated interchromomeres in the arms would be 1900 min or about three times the S phase. Thus the minimum number of origins required for the heterochromatic DNA is two (or three), separated from each other and from the closest activable interchromomeres in the arms by 3100 kb, and all activated at the beginning of S.

If activation of sequences within the heterochromatich DNA occurs later in S phase (as is true in some cases, but has not yet been studied in Drosophila), then to conform to our general model, we must suppose that such origins arise by changes of structure during S which make origin sequences available to activation by creating a small number of noncompacted regions within the heterochromatin. In either case, it is the structure of the heterochromatin which is presumed to control the duration of the S phase in cell cultures. Furthermore, in both cases we expect that origin-to-origin distances within this heterochromatin will be very large in comparison to that in the arms, and this is a testable prediction.

Concluding Remarks

In these concluding remarks we wish to emphasize and generalize the salient points of the preceding results and arguments. The replication of the genome in the short time available in cleavage nuclei requires a very large number of activable origin sequences. We suppose that it is the absence of heterochromatization of the satellite segments, and the presumed lack of chromomeres in the remainder of the DNA, that allows such a large number of origin sequences to be available for activation. The periodic spacing of 3.4 kb, or multiples thereof, between active origins is puzzling, but may indicate that even in these nuclei the origin sequences are divided into activable and non-activable classes by some periodic packing of the interphase chromatin. There must, in any case, be some mechanism which prevents the reactivation of an origin sequence immediately following its replication.

When cells are formed from cleavage nuclei, chromomeres are assumed to be created for the purpose of regulating the transcripitive activity of the genome (see review by Beermann, 1972). The chromomeric pattern is sequence determined (Judd and Young, this volume), and we presume that it is essentially constant in the interphase nuclei of all somatic cells in D. melanogaster. Although the distribution of active origins that we observe in cell cultures can be nicely explained by the occlusion of origin sequences resulting from the formation of chromomeres, clearly this mechanism cannot be used to explain differences in S phase for different kinds of somatic cells. Indeed, a prediction of this model is that most of the active origins in all somatic cells of D. melanogaster will exhibit the same distribution that we have observed for cultured cells. Nor can we expect that the differences in S phase are due to differences in fork rate, when such diverse nuclei as we have examined exhibit the same fork rate.

Hence, we imagine that it is the heterochromatization of the regions containing the satellite DNAs that provides the mechanism for controlling the duration of S phase in somatic cells. This postulate requires that the structure of the heterochromatin regions can be varied—from cell type to cell type and/or within a cell during S phase—so as to change the distribution of activable sequences. We presume that the primary result of this variation is to change the topography of the compacted and noncompacted segments within the region. By occluding most of the origin sequences within the compacted regions, the long origin-to-origin distances required for control of S phase can be created; and by adding or subtracting a few noncompacted regions, the length and timing of the replication of such segments can be sensitively controlled. Perhaps it is the kind, amount, and distribution of the various satellite DNAs in the
centered DNA that determines the nature of this structural control.

Acknowledgments

This work was supported by research grants from the National Science Foundation and the National Institutes of Health. A.B.B. was a fellow of the Damon Runyon Memorial Fund during the course of this work, and H.J.K. is a trainee in the Medical Scientist Training Program of the U. S. Public Health Service.

Appendix

Derivations of equations (3), (4), and (5). Consider a cluster in which the origins are activated simultaneously and separated by a distance d. Divide the distance d into an even number of equal intervals of length Δ, so as to define n = d/2Δ states of the cluster such that the length of the segment between adjacent origins that has yet to be replicated is 2Δ, 2(2Δ), 3(2Δ), ..., (n/2)Δ in the different states. Since the cell population is asynchronous, all states are equally represented; hence the number of eyes, N, in each state is the same. The time required for a cluster to move from one state to the next is Δ/r, where r is the fork rate, and, therefore, in each time interval, Δ/r, N new eyes will be formed and N eyes will be lost by merger. Let the pulse time, t, equal n(Δ/r), where n is any positive integer.

We consider first the condition where t < d/2r, or n < d/2Δ. The tracks from the n • N new eyes created during the pulse will be formed from two diverging forks and will exhibit lengths equal to 2Δ, 2(2Δ), ..., n(2Δ), and the total length of all these tracks will be N • n(1 + 1/2Δ). The remaining tracks are formed from clusters existing at the beginning of the pulse (except for the clusters where the length of the segments to be replicated equals d, these clusters being included in the first group). Of these, nN are formed by two converging forks and will exhibit lengths equal to 2Δ, 2(2Δ), ..., n(2Δ), or a total track length that also equals N • n(n + 1/2Δ). The rest of the tracks will each have a length equal to nΔ, as they are formed by single forks. Two of these forks will emanate from each of N[(d/2Δ) - n] eyes so that the 2N[(d/2Δ) - n - 1] tracks will exhibit a total length of 2NΔ(Δ/Δ) - n - 1). The sum of the total lengths for each of the three classes of tracks divided by the sum of the number of tracks in the three classes will give the mean track length, 〈l〉, which equals nNd/(d - 2Δ), or because we can make 2Δ ≪ d, equals nΔ. Since n • Δ = rt, then 〈l〉 = rt, equation (3a) in the text. In a similar but simpler manner it can be shown that the mean center-to-center distance, 〈d〉 = (d/2) + rt when t < d/2r, i.e., equation (4a) in the text.

When t ≥ d/2r, or n ≥ d/2Δ, all tracks from clusters existing at the beginning of the pulse will be formed by two converging forks. There will be N[(d/2Δ) - 1] of these tracks which exhibit a total length of N(d - 2Δ)/4Δ. Of the clusters formed during the pulse, only those in which the eyes do not merge during the pulse will form scorable tracks (see text), and these will arise from two diverging forks. The number of these tracks and their total length will likewise be N[(d/2Δ) - 1] and N(d - 2Δ)/4Δ, respectively. Hence the mean track length, 〈l〉, in this condition will be the second of these terms divided by the first, or d/2. This is equation (3b) in the text, and equation (4b) can be derived in the same manner. Equation (5) derives from the fact that the number of eyes leading to scorable tracks, or 2N[(d/2Δ) - 1], is a fraction of the sum of the number of eyes existing at the beginning and those created during the pulse, or N[(d/2Δ) + n]. This fraction, F, is therefore d/2d - 2Δ)/(d + 2Δ), or since we can make 2Δ ≪ d, is d/2 + 2Δ. Since nΔ = rt, then 〈l〉 = 2d/(d + 2rt), which is equation (5).


