Heidner, E. J., Ladner, R. C., and Perutz, M. F. (1976), *J. Mol. Biol. 104,* 707.
Iizuka, T., and Kotani, M. (1969a), *Biochim. Biophys. Acta 181,* 275.
Iizuka, T., and Kotani, M. (1969b), *Biochim. Biophys. Acta 194,* 351.
Ladner, R. C., Heidner, E. J., and Perutz, M. F. (1977), *J. Mol. Biol.* (in press).
Muirhead, H., Cox, J. M., Mazzarella, L., and Perutz, M. F. (1967), *J. Mol. Biol. 28,* 117.
Peisach, J. (1975), *Ann. N.Y. Acad. Sci. 244,* 187.
Perutz, M. F. (1970), *Nature (London) 228,* 726.
Perutz, M. F. (1976), *Br. Med. Bull. 32,* 195.
Perutz, M. F., Fersht, A. R., Simon, S. R., and Roberts, G. C. K. (1974), *Biochemistry 13,* 2174.
Rosemeyer, M. A., and Huehns, E. R. (1967), *J. Mol. Biol. 25,* 253.
Rougee, M., and Brault, D. (1975), *Biochemistry 14,* 4100.
Sono, M., Smith, P. D., McCray, J. A., and Asakura, T. (1976), *J. Biol. Chem. 251,* 1418.
Takano, T. (1977a), *J. Mol. Biol. 110,* 533.
Takano, T. (1977b), *J. Mol. Biol. 110,* 569.
Warshel, A. (1977), *Proc. Natl. Acad. Sci. U.S.A. 74,* 1789.
Yip, Y. K., Waks, M., and Beychok, S. (1972), *J. Biol. Chem. 247,* 7237.

Analysis of Rat Repetitive DNA Sequences†

William R. Pearson, Jung-Rung Wu, and James Bonner*

ABSTRACT: Parameters of repetitive sequence organization have been measured in the rat genome. Experiments using melting, hydroxylapatite binding, and single strand specific nuclease digestion have been used to measure the number, length, and arrangement of repeated DNA sequences. Renaturation and melting or S1 nuclease digestion of 1.0 kbp DNA fragment show about 20% of rat DNA sequences are 3000-fold repeated. Renatured duplexes from 4.0 kbp DNA fragments display two repetitive size fractions after nuclease digestion. About 60% of the repeated sequences are 0.2–0.4 kbp long while the remainder are longer than 1.5 kbp. The arrangement of the repeated sequences has been measured by hydroxylapatite fractionation of DNA fragments of varying lengths bearing a repeated sequence. Repeated DNA sequences are interspersed among 2.5 kbp long nonrepeated sequences throughout more than 70% of the rat genome. There are approximately 350 different 3000-fold short repeated sequences in the rat interspersed among 600 000 nonrepeated DNA sequences.

Evidence is accumulating that one of the mechanisms for differential gene expression is the sequence specific regulation of RNA transcription. Recent work on a variety of organisms has demonstrated an almost universal highly ordered pattern of repetitive sequence organization in DNA (see Davidson et al., 1975a, for review). The proximity of subsets of repeated sequences to transcribed and translated DNA sequences supports a regulatory function. These structural observations on sequences near functional DNA sequences suggest that repeated sequences may play an important role in DNA sequence recognition during the regulation of transcription.

Repeated DNA sequences display strikingly similar, highly ordered structures across a wide range of organisms from insects to mammals (Davidson et al., 1973a,b; Bonner et al., 1973; Davidson et al., 1975a,b; Graham et al., 1974; Angerer et al., 1975; Chamberlin et al., 1975; Goldberg et al., 1975; Schmid and Deininger, 1975; Efstratiadis et al., 1976). With a few exceptions (Manning et al., 1975; Crain et al., 1976), repeated sequences 0.2–0.4 kbp long are interspersed among 1.0–2.0 kbp nonrepeated single copy sequences in more than 65% of the DNA of all organisms studied. This organization was predicted by Britten and Davidson (1969; also Davidson and Britten, 1973) in a model which suggested that repeated DNA sequences can coordinately control the expression of adjacent single copy genes.

While the similarity of organizational patterns across the evolutionary spectrum is striking evidence for repeated DNA function in gene expression, experiments demonstrating specific subsets of repeated sequences adjacent to transcribed and translated genes provide even stronger support for the regulatory function of repeated sequences. Experiments on large nuclear transcripts in the rat (Holmes and Bonner, 1974b) and sea urchin (Smith et al., 1974) show that most large nuclear RNA is transcribed from DNA containing interspersed repeated sequences. More recent experiments suggest that a select subset of repeated sequences is near transcribed and translated sequences. Gottesfeld has reported (Gottesfeld et al., 1976) that a fraction of DNA from chromatin with increased transcriptional activity contains not only a subset of single copy DNA but also a subset of all repeated sequences. Experiments using sea urchin messenger RNA have shown 80% of the translated sequences are adjacent to repeated DNA sequences (Davidson et al., 1975b) and that these adjacent sequences are a subset of the repetitive sequence population (Klein et al., in preparation). Selected repeated sequences are adjacent to single copy DNA sequences which are transcribed into message and translated.

In this paper we describe experiments which measure the structural parameters of repeated DNA sequence organization in a mammal, the rat. We have made three basic measurements. First, the fraction of the DNA which contains repeated sequences has been determined. Second, we have measured the

† From the Division of Biology, California Institute of Technology, Pasadena, California 91125. *Received July 29, 1977.* This work was supported in part by U.S. Public Health Service Training Grant GM 00086 and in part by U.S. Public Health Service Research Grants GM 13762 and GM 20927.

0006-2960/78/0417-0051$01.00/0

length and thus the number of the repeated sequences. Third, we have studied the distribution of repeated sequences in single copy DNA.

Materials and Methods

Preparation of DNA. DNA used in these experiments was prepared from two sources. Unlabeled DNA was isolated from a rat Novikoff ascites cell line grown intraperitoneally and transferred at 7-day intervals. Cells removed from the rats were frozen and stored at −80 °C. For the DNA preparation, cells were thawed and a crude nuclear pellet isolated by the method of Dahmus and McConnell (1969). The crude nuclei were resuspended in 0.1 M Tris, 0.1 M EDTA, pH 8.5, and lysed with the addition of 20% sodium dodecyl sulfate to a concentration of 2%. The viscous lysate was heated to 60 °C for 15 min and then digested with Pronase (Calbiochem, grade B) at 50 μg/mL until the DNA went into solution (usually 2 h). The DNA solution was then extracted with an equal volume of 1:1 phenol:IAC[1] (24:1 chloroform:isoamyl alcohol), the aqueous phase removed, and the interphase reextracted with buffer and phenol:IAC. The aqueous phases were pooled and reextracted with phenol:IAC followed by 2–3 extractions with IAC. The DNA was then precipitated with 2.5 volumes of 95% ethanol at room temperature and spooled. The DNA was dissolved in 10 mM Tris, 10 mM EDTA, pH 8.5, overnight at 4 °C, then brought to 0.1 M Tris, 0.1 M EDTA, pH 8.5, and digested with 50 μg/mL RNase A (Worthington) for 90 min (after 10 min of RNase preincubation at 95 °C) after which 200 μg/mL preincubated Pronase (10 min, 37 °C) was added. After a second 90-min incubation at 37 °C the DNA was extracted three times with IAC, precipitated with ethanol, and spooled. The DNA was redissolved and spun at 27 000 rpm in a Beckman SW27 rotor to remove undissolved material and then reprecipitated and spooled.

^{3}H-labeled DNA was prepared from Novikoff hepatoma cells grown in suspension culture in flasks on a shaker. The cells were grown on a modified Swimm's 67 medium (Plageman and Swimm, 1966; Plageman, personal communication) and had a doubling time of 12 h. Cells were labeled with [^{3}H]thymidine by the addition of 0.1 μCi/mL (Schwartz, 46 mCi/μmol) four times at 12-h intervals. Under these conditions, we were able to isolate DNA with specific activities around 100 000 cpm/μg.

After growth to a concentration of 1×10^6 cells/mL the cells were centrifuged at 2000*g* for 2 min and resuspended in lysis buffer, 10 mM Tris, 10 mM NaCl, 1.5 mM $MgCl_2$, 1 mM $CaCl_2$, 1 mM triethanolamine, pH 7.4 (B4; Plageman, 1969), pelleted at 2 K for 5 min, resuspended in B4, and allowed to stand for 5 min. The cells were lysed with 10 strokes of glass homogenizer and the nuclei spun down at 2000*g* for 5 min. This crude pellet was then extracted as described above for the unlabeled ascites DNA.

Preparation of DNA Fragments. DNA fragments of various sizes were prepared by shearing in a Virtis 60 homogenizer (Britten et al., 1974). The 4.0-kbp fragments were prepared by shearing at 7500 rpm and 0.8–1.1-kbp fragments were sheared at 30 000 rpm for 45 min; 0.35 kbp DNA was prepared by shearing at 50 000 rpm in 66% glycerol as described by Britten et al. (1974). Broader distributions of fragment lengths for interspersion studies were prepared by shearing at 2000–40 000 rpm for 30 s to 5 min and individual sizes fractionated from preparative alkaline sucrose gradients. After shearing, all preparations were filtered and passed over Chelex-100 (Bio-Rad) and then precipitated from 0.3 M sodium acetate with 2.5 volumes of 95% ethanol.

Sizing DNA Fragments. Single-stranded DNA fragment lengths were determined by sedimentation through alkaline sucrose gradients. Isokinetic sucrose gradients (Noll, 1967) were formed in SW41 tubes in 0.1 N NaOH using V_{max} of 10.4 mL: C_{flask} = 16% w/v, C_{res} = 43% w/v. Gradients were centrifuged from 16 to 24 h at 40 000 rpm. All tubes contained two markers of known molecular weight and samples were run at least two times. Molecular weights were calculated from sedimentation rates using the Studier (1965) equations.

S1 Nuclease Renaturation and Digestion. DNA samples to be digested with S1 nuclease were incubated in 0.3 M NaCl, 0.01 M Pipes, pH 6.8, at 65 °C (Smith et al., 1975; Britten et al., 1976). The equivalent C_0t was calculated using a factor of 2.31 to correct for 0.3 M Na^+ concentration. Long (>0.5 kbp) DNA samples were alkaline denatured to ensure complete denaturation.

After incubation, samples to be nucleased were diluted with a volume of 0.05 M sodium acetate, 0.2 mM $ZnSO_4$, pH 4.2, and dithiothreitol was added to 5 mM. The final reaction mix was 0.15 M NaCl, 0.025 M sodium acetate, 5 mM Pipes, 0.1 mM $ZnSO_4$, 5 mM dithiothreitol, pH 4.4. The pH was checked in each experiment. The S1 nuclease preparation used was the gift of Dr. F. Eden and D. Painchaud and has been extensively characterized (Britten et al., 1976). The standard enzyme to DNA ratio used was 15 μL/mg for measurements of the fraction repetitive in the genome. DNA samples were incubated with S1 nuclease for 45 min at 37 °C. The reaction was terminated by addition of 2.0 M PB to a final concentration of 0.12 M and the sample passed over hydroxylapatite. For some experiments, the fraction bound to hydroxylapatite was eluted with 0.5 M PB. For other experiments the fraction bound to hydroxylapatite in 0.12 M PB was denatured at 100 °C and eluted with 0.12 M PB. The size distribution of S1 nucleased duplexes was measured on a 110 × 1.3 cm column of agarose A-50 (Bio-Rad). The gel bed was poured around a support of 5-mm glass beads (Britten et al., 1974). Samples were chromatographed in 0.12 M PB using $^{32}PO_4$ as an inclusion marker.

DNA Renaturation. Samples which were not to be digested by S1 nuclease were incubated in 0.12 M PB at 60 °C or in 0.48 M PB at 70 °C. After incubation, samples were frozen in dry ice–ethanol. Samples were thawed and diluted to 0.12 or 0.14 M PB and passed over hydroxylapatite at 60 °C. The fraction bound was eluted after thermal denaturation at 100 °C in all experiments except the duplexes described in Tables III–V. The fraction and rate parameters for the renaturation curves were calculated using a nonlinear least-squares fitting program (Pearson et al., 1977).

Melting. DNA samples were melted in 0.12 M PB in a Gilford Model 2400 spectrophotometer equipped with a Model 2527 thermal cuvette. Samples were melted at a rate of 0.5 °C/min and the A_{260} automatically sampled at 0.5 °C intervals. Hyperchromicity was calculated from the formula

$$H = \frac{A_{260(98°C)} - A_{260(60°C)}}{A_{260(98°C)}}$$

after subtraction of the buffer absorbance at each temperature.

Results

The Renaturation of Rat DNA. Repeated sequences can be characterized from the kinetics of renaturation of short

[1] Abbreviations used: IAC, isoamyl alcohol:chloroform, 1:24 v/v; EDTA, ethylenediaminetetraacetic acid; RNase, ribonuclease; PB, phosphate buffer; kbp, kilobase pair; HAP, hydroxylapatite; Pipes, piperazine-*N*,*N*′-bis(2-ethanesulfonic acid); DIG, digestion estimate.

TABLE I: Renaturation of 0.35-kbp Rat DNA Fragments.

Component	Fraction	Rate	$C_0t_{1/2}$	Repetition frequency
	A. Unconstrained Fit (No Parameters Fixed, Goodness of Fit, 3.0%)			
1	0.0883	>100	0.01	>350 000
2	0.177	1.06	0.9434	3500
	(±0.015)	(±0.39)		
3	0.191	0.00381	262.5	13
	(±0.098)	(±0.00308)		
4	0.537	0.000292	3424	1
	(±0.079)	(±0.000132)		
	Final fraction unreacted: 0.0067 ± 0.0453			
	B. Constrained Fit (Single Copy Rate Fixed at 0.00034 for 2.9 pg of Genome)			
1	0.0882	>100	0.01	>300 000
2	0.176	1.10	0.909	3235
	(±0.014)	(±0.39)		
3	0.181	0.00421	238	12
	(±0.042)	(±0.00211)		
4	0.525	0.0034	2941	1
	(±0.051)			
	Final fraction unreacted: 0.0298 ± 0.0194			
	C. Constrained Fit (Single Copy and Repetitive Rates Fixed)			
1	0.0729	>100	0.01	>300 000
2	0.176	2.97[a]	0.337	8727
	(±0.011)			
3	0.213	0.00412[a]	243	12
	(±0.020)			
4	0.499	0.00034	2941	1
	(±0.030)			
	Final fraction unreacted: 0.0391 ± 0.0165			

[a] Fixed repetitive rate constants calculated from the fit of the data in Figure 2.

DNA. Figure 1 shows the renaturation of 0.35-kbp long fragments. The fraction of the fragments containing duplex was determined by hydroxylapatite binding. Least-squares computer fits to these data are summarized in Table I. The data display three qualitatively distinct components: a single slow component renaturing between C_0t 200–20 000; a more rapidly renaturing component binding between C_0t 0.02–200; and a very rapidly reannealing component bound by the earliest times in the data.

The slowly renaturing component containing 55–65% of the DNA is single copy nonrepeated DNA. Table I shows two different fits of the data analyzing this component. The first fit, a "free fit", allows all parameters to vary to find the best fit. This fit does not terminate but will continue beyond these parameters to a parameter set with more than 100% reaction (Pearson et al., 1977). In the second fit, the rate constant for the single copy component was fixed at a number appropriate for the genome size of the rat. For a rat genome size of 2.9 pg (Sober, 1968), the single copy rate constant is fixed at 0.00034 L mol^{-1} s^{-1} corresponding to a $C_0t_{1/2}$ of 2900. This fit terminates properly and we can calculate the fraction in the single copy component and the fractions and repetition frequencies of the repetitive components.

The renaturation from C_0t 0.02 to 200 is due to the reaction of sequences repeated from 10 to 100 000 times in the rat genome. About 25–30% of the 0.35-kbp long fragments contain duplex in this region of the curve. The DNA fragments bound to hydroxylapatite contain single-stranded tails, and the real fraction of repeated sequences in the genome is lower. Due to the scatter in the data and the fraction of the genome involved, accurate estimation of the repetition frequencies for the different components is difficult. This is reflected in the high parameter error estimates.

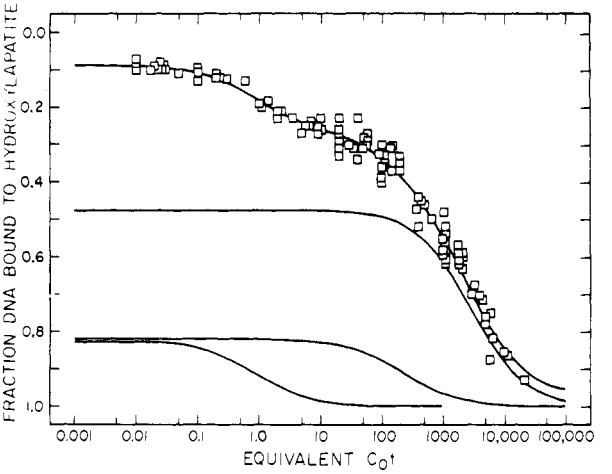

FIGURE 1: Renaturation of 0.35-kbp rat DNA fragments. These data have been collected by a number of investigators (Holmes and Bonner, 1974a,b; Gottesfeld et al., 1976). Denatured 0.35-kbp DNA was renatured in 0.12 M PB at 60 °C or in 0.48 M PB at 70 °C. The fraction single stranded was measured by hydroxylapatite chromatography. Equivalent C_0t is the C_0t (L mol^{-1} s^{-1}) times a salt concentration factor (Britten et al., 1974). The solid lines show the least-squares fit of the data using a single copy rate fixed at 0.00034 for a genome size of 2.9 pg (Sober, 1968). The actual coefficients for this fit are shown in Table IB.

For better resolution of the repetitive components a repetitive fraction of the genome was isolated. ^{3}H-labeled 0.35-kbp long fragments were renatured to C_0t 100 and the duplexed repetitive sequences separated on hydroxylapatite. This repetitive fraction was then hybridized with a 100–1000-fold excess of whole 0.35-kbp long fragments. The data are shown in Figure 2. This analysis of the repetitive fraction provides a

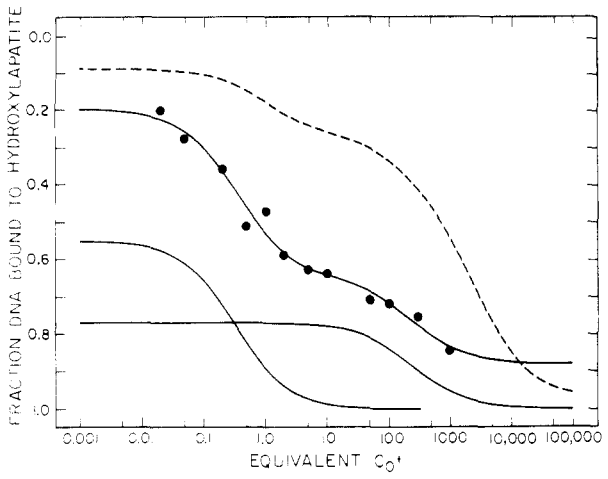

FIGURE 2: Renaturation of selected repetitive sequences. 0.35 kbp ^{3}H-labeled rat DNA was renatured to C_0t 100 and the double strand containing fraction (30%) separated on hydroxylapatite. This fraction was driven by unfractionated 0.35 kbp unlabeled DNA. The mass ratio of driver to tracer DNA was 100 from C_0t 0.001 to 0.5 and 1000 from C_0t 1.0–1000. The best least-squares fit to the data gives 0.45 of the DNA reannealing with a rate of 2.97 L mol^{-1} s^{-1} and 0.23 reannealing with a rate of 0.00412; 0.12 of the DNA had not reached at the highest C_0t's. These rate values were used for the third fit in Table I. The line drawn through the data displays this fit. The two repetitive components are also shown. The line above the data is the whole rat DNA fit from Figure 1.

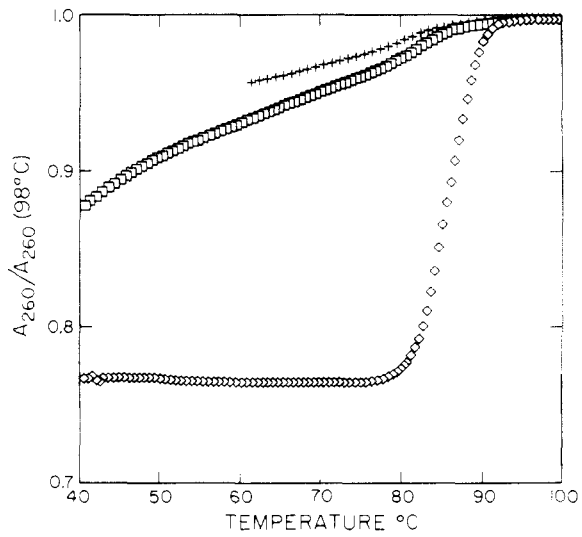

FIGURE 3: Melt of rat DNA fragments renatured to C_0t 5. DNA was sheared to 1.0 kbp, denatured, and renatured to C_0t 5 in 0.12 M PB. The sample was then melted in a spectrophotometer equipped with a thermal cuvette. The temperature was raised at a rate of 0.5 °C/min to 98 °C. After melting, the samples were cooled to 60 °C and remelted. (□) the 1.0-kbp DNA renatured to C_0t 5; (+) remelt of the DNA; (◇) native DNA.

more accurate estimate of the repetitive reaction rate constant. About half of the repetitive reaction takes place with a rate constant of 3.0 L mol^{-1} s^{-1}, corresponding to a repetition frequency of 10 000. The remainder renatures with a rate constant of 0.004 L mol^{-1} s^{-1}, corresponding to a repetition frequency of 10. These values can be used in the least-squares fit of the whole rat data and provide a third interpretation of the kinetic fractions in the rat genome in Table I. There is little difference in the goodness of fit criterion for each of the three fits, and we will use the second interpretation (B) as our best model for rat DNA renaturation.

Ten percent of rat DNA has renatured by the earliest times shown on the curve. These sequences may contain very highly repeated (>300 000-fold) DNA sequences and foldback self-complementary sequences. Again, some of this fraction is due to single-stranded tails on duplexes bound to HAP.

This paper will concentrate on the repeated fraction of the genome renaturing before C_0t 50. It is difficult to study the low (10–50-fold) repeated fraction because of substantial contamination by single copy sequences. These sequences have been classed with nonrepeated fractions for the purposes of this paper. The rapidly reannealing sequences are included in most of our analysis. There is evidence (Wu et al., in preparation) that the foldback complementary repeated sequences are moderately repeated sequences similar to the other repeated sequences in the genome.

Melting Experiments. The number and length of the repeated sequences can be calculated from measurements on the fraction of the genome in true duplex after renaturation to repetitive C_0t's. We have used optical melting and S1 nuclease digestion to measure the fraction of the genome containing repeated sequences and the length of the duplexed repeats.

To measure the repeated fraction of rat DNA optically, 1.0 kbp long fragments were renatured to C_0t 5 and C_0t 50 and melted in a spectrophotometer. Figure 3 shows a sample melt, which includes a native DNA standard and a remelt of the melted DNA. The remelt shows the contribution of single-strand tails and very rapidly reannealing sequences to the hyperchromicity of the duplex. Table II summarizes the measurements on C_0t 5 and C_0t 50 duplexes. These values are very sensitive to corrections for single-stranded hyperchromicity. In addition, some of the apparent single-stranded hyperchromicity is due to renaturation of rapidly reannealing sequences indicated by the transition above 80 °C. The hyperchromicity in the transition, 0.015 to 0.02, is consistent with 5% rapidly reannealing DNA. When this 5% fraction is added to the values in Table II, we find 18% of the genome is in duplex at C_0t 5 and 24% is in duplex at C_0t 50.

TABLE II: Melts of 1.0-kbp Rat DNA Fragments Renatured to C_0t 5 and C_0t 50.

Sample	Hyperchromicity[a]	T_m (°C)	Fraction[b] native
C_0t 5A	0.0800	77.0	0.1818
B	0.0687	75.5	0.1305
C	0.0586	75.0	0.0764
D	0.0677	76.0	0.1259
Average fraction native 0.1287 ± 0.0431			
C_0t 50A	0.0788	74.5	0.1764
B	0.0821	75.5	0.1914

[a] Hyperchromicity measured from 60 to 98 °C. [b] Fraction native is corrected for 0.04 hyperchromicity after remelt and uses 0.2600 for the native hyperchromicity. The calculation is: (Hyper − 0.04)/(0.2600 − 0.04).

An average length for repeated duplexes can be measured by melting DNA fragments of different length containing duplexes (Graham et al., 1974). DNAs of different lengths were renatured to C_0t 50 and duplex containing strands isolated on hydroxylapatite. The results of the optical melt of the duplex containing fragments are summarized in Table III. The hyperchromicity of each strand length is a function of the average length of the strand in duplex. Longer strands have lower hyperchromicity because the duplex region is a smaller fraction of the molecule (Davidson et al., 1973a,b). These measurements indicate the average length of the duplex is 0.4 kbp.

These data also provide an estimate of the amount of re-

TABLE III: Melts of Rat DNA Fragments Bearing Repeated Sequences.

Input fragment length (kbp)	0.35	1.1	1.6	Native
Fraction bound to HAP	0.26	0.44	0.47	
Hyperchromicity	0.175	0.110	0.090	0.260
Fraction duplex[a]	0.646	0.375	0.292	
Duplex length[b]	0.226	0.413	0.467	
Fraction of genome in duplex[c]	0.168	0.165	0.137	

[a] Fraction duplex is calculated as in Table II for fraction native. [b] Duplex length is the product of the fraction duplex and the input fragment length. [c] Fraction genome in duplex is the product of the fraction duplex and the fraction bound to HAP (hydroxylapatite).

TABLE IV: Nuclease Digestion of 1.0-kbp DNA Fragments.

C_0t	Fraction containing duplex before digestion[a]	Fraction duplex after S1 digestion[b]
0.05	0.215 ± 0.017 (n = 2)	0.080 ± 0.046 (n = 3)
0.5	0.307 ± 0.028 (n = 2)	0.134 ± 0.032 (n = 2)
5.0	0.282 ± 0.032 (n = 3)	0.156 ± 0.031 (n = 5)
50.0	0.403 ± 0.051 (n = 5)	0.239 ± 0.010 (n = 4)

[a] Fraction of DNA bound to hydroxylapatite before S1 nuclease digestion. [b] Fraction of DNA bound to hydroxylapatite after 45-min digestion with S1 nuclease (15 μL/mg).

TABLE V: A-50 Fractionation of Repeated DNA Fragments.

C_0t	Fraction S1 resistant	Fraction excluded from A-50
0.05	0.0368	0.484
	0.0558	0.683
5.0	0.150	0.453
	0.140	0.489
	0.171	0.481
	0.167	0.553
	0.177	0.534

petitive sequences in the genome. Although not all of the duplex molecules can be removed from hydroxylapatite by elution with the 0.5 M phosphate buffer used in these experiments, if the fraction retained has the same hyperchromicity as the duplex fraction eluted, about 17% of the DNA is duplex at C_0t 50.

Nuclease Digestion of Repeated Sequences. An independent measure of the fraction duplex is provided by digestion with the single strand specific nuclease S1. Fragments 1.0-kbp long were renatured to C_0t's from 0.05 to 50 and duplexed molecules fractionated on hydroxylapatite with and without S1 nuclease digestion (Table IV). The difference between binding with and without nuclease is due to the absence or presence of single-stranded tails. Using the S1 criterion, 8% of the DNA contains repeats renaturing by C_0t 0.05 while 16% is repetitive at C_0t 5 and 24% at C_0t 50.

To make certain that the enzyme was behaving properly, the fraction of duplex as a function of enzyme to DNA ratio was investigated at C_0t 5. The 16% of the DNA in duplex at C_0t 5 is not a function of enzyme concentration over the range of 15 μL/mg (15 μL/mg is the standard concentration used for DIG = 0.78; Britten et al., 1976) and is not affected by a twofold increase in the salt concentration. Lowering the enzyme concentration or raising the salt concentration should stabilize poorly matched duplexes. We do not believe the enzyme is digesting slightly mismatched regions as these changes do not affect our results.

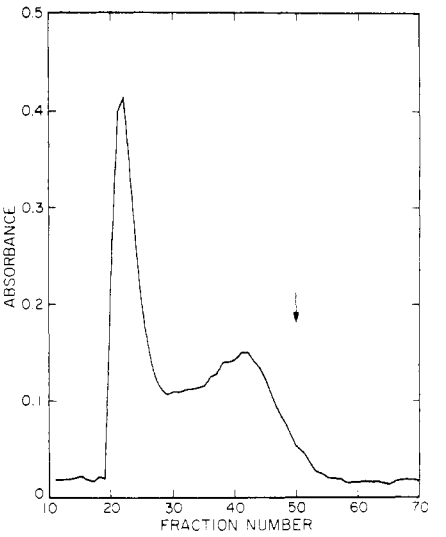

FIGURE 4: Profile of rat repeated DNA duplexes on agarose A-50. DNA sheared to 3.0 kbp was denatured and renatured to C_0t 5, digested with S1 nuclease, and bound to hydroxylapatite. The double-stranded fraction (15%) was eluted with 0.5 M PB and chromatographed on Biol-Gel agarose A-50.

Fragments from nuclease digestion can be used to determine the length of the duplex by direct sizing. Duplexes from 3.0-kbp DNA fragments isolated on hydroxylapatite after nuclease digestion were fractionated by size on Bio-Gel agarose A-50. Figure 4 shows a typical elution pattern for this DNA renatured to C_0t 5. The included and excluded material was sized on alkaline sucrose gradients and by electron microscopy. The material in the included peak is 0.2–0.3-kbp long while the excluded material is longer than 1.5 kbp. Table V summarizes hydroxylapatite and A-50 fractionation data at C_0t 0.05 and C_0t 5.

To make certain that long duplexes did not contain single-stranded tails, duplexes fractionated on A-50 were melted. All fractions show more than 90% of native hyperchromicity. The melting temperatures range from 2.5 °C below native for the excluded material to 15 °C below for the 0.2-kbp fragments.

The melts of both the long and short repetitive fragments display a number of distinct melting regions. These regions can be seen most clearly in the melt of the long A-50 excluded material where there are components melting with T_m's of 70, 82, and 86 °C. The short fragments also show components with T_m's of 65, 75, and 83 °C.

Interspersion of Repeated Sequences. The data on melting and nuclease digestion of long DNA fragments presented earlier indicate that many short repeated sequences are surrounded by single copy DNA. In this section we present two measures of the properties of interspersed repeated sequences. The first experiment gives a qualitative measure of the fraction of the genome containing interspersed repeats; the second gives quantitative data on the fraction of the DNA interspersed, the length of the interspersed single copy sequences, and the length of the repeated sequences.

When 0.3, 1.5, and 3.0-kbp long DNA fragments are renatured and the duplex containing fraction is separated on hydroxylapatite, the repeated fraction appears to increase with the length of the renatured fragments. Figure 5 shows the renaturation of these three fragment lengths. Thirty percent of 0.3-kbp long fragments are bound at C_0t 50 but 70% of 3.0 kbp DNA fragments are retained on hydroxylapatite at this C_0t.

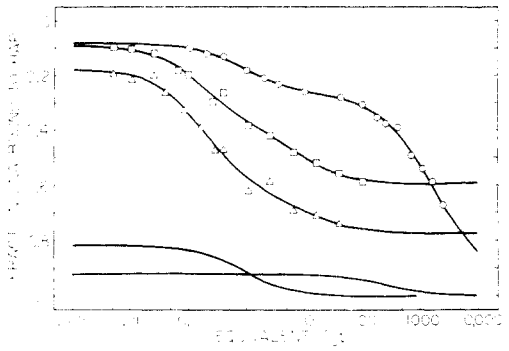

FIGURE 5: Renaturation of different DNA fragment lengths. DNA sheared to 0.3, 1.5, and 3.0 kbp determined by alkaline sucrose sedimentation was denatured and renatured. The samples were passed over hydroxylapatite and the double strand containing fraction eluted at 100 °C with 0.12 M PB. The fraction single stranded as a function of equivalent C_0t is plotted. (○) The 0.3-kbp DNA; (□) 1.5-kbp fragments; (△) 3.0-kbp.

The increased binding is due to single-stranded tails on short duplexes. These tails were seen as loss of hyperchromicity in the melting experiments and loss of hydroxylapatite binding after S1 nuclease digestion.

The precise fraction of the DNA containing short interspersed repeats and their spacing can be measured by renaturing trace quantities of varying length DNA fragments with a vast excess of unlabeled short DNA. Whole rat DNA sheared to 0.35 kbp was used to drive labeled fragments from 0.25 to 8.0 kbp in length. The fraction bound to hydroxylapatite at zero time ($<C_0t$ 0.005), C_0t 5, and C_0t 50 was measured. Figures 6A and 6B show the binding of the DNA at C_0t 5 and C_0t 50 corrected for the zero time binding fraction. Figure 6C shows the zero-time binding used to correct the data. The lines plotted show the range of curves which can be used to fit the data. The curves show that more than 80% of the 8.0-kbp DNA contains repetitive sequences renaturing by C_0t 50. The spacing between repeats (the interspersion period) is about 2.3–2.7 kbp and does not change significantly between C_0t 5 and C_0t 50. At C_0t 5, 57% of the DNA is bound at a fragment length of 2.3 kbp; 73% of the DNA is bound at a fragment length of 2.7 kbp at C_0t 50. The fraction of the DNA in repeated duplexes is the Y intercept of the data, about 14% at C_0t 5 and 22% at C_0t 50. A length estimate for the repeated duplex is shown at the X-axis intercept and is about 0.84 kbp at C_0t 5 and 1.3 kbp at C_0t 50.

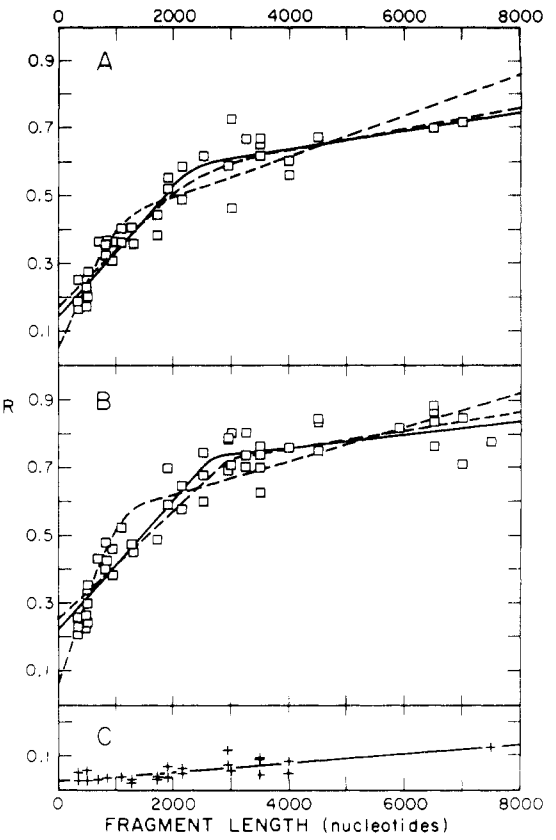

FIGURE 6: The fraction of rat DNA sequences containing a repeated DNA sequence as a function of fragment length. Labeled DNA fragments of various sizes were driven by whole rat DNA to C_0t 5 and C_0t 50. Binding to hydroxylapatite without driver at very short times (zero-time binding) was also measured. The fraction of the fragments containing duplex F was measured by hydroxylapatite chromatography. Values of F were corrected for the fraction of zero-time binding sequences Z. The zero-time binding correction was calculated from a linear fit of the data in C. The values plotted in A and B are $R = (F - Z)/(1.0 - Z)$. The solid line drawn to the data represents the best least-squares fit to the data. The dashed lines provide alternative interpretations of the data. (A) Corrected binding of fragments driven to C_0t 5. Solid line: best fit with fraction repetitive (f_{rep}) = 0.140, interspersion period (L_{int}) = 2.315 kbp, fraction interspersed at short period (f_{int}) = 0.573, repetitive sequence length (L_{rep}) = 0.835 kbp, fraction bound by 8.0 kbp ($f_{8.0}$) = 0.745. One dashed line shows the fit when the fraction repetitive is fixed at 0.17 while the other shows the best fit of the data when the interspersion distance is set to 1.0 kbp (1000 nucleotides). (B) Corrected binding of fragments driven to C_0t 50. Solid line: best fit with f_{rep} = 0.219, L_{int} = 2.680 kbp, f_{int} = 0.730, L_{rep} = 1.27 kbp, f_{8000} = 0.836. The dashed lines show the best fits when the fraction repetitive is set to 0.25 and the interspersion period is set to 1.0 kbp (1000 nucleotides). (C) Binding of fragments incubated without driver or for very short times (zero-time binding). The line drawn to the data was used to correct the data in A and B.

Discussion

We have presented measurements on the repetition frequency, fraction of the genome, and length of repeated DNA sequences in rat DNA. In this section we will summarize the different lines of evidence which support a consistent model for repeated sequence organization in the rat genome.

The kinetic data from the hydroxylapatite C_0t curve are consistent with the reported genome size for the rat (2.9 pg, Sober, 1968). A variety of combinations of rate constants and repetitive and single copy fraction quantities describe the data equally well. This is indicated by the standard error of the parameters in Table I. Some of this uncertainty is due to the difficulty of getting good termination data for the reaction. A change from a final fraction unreacted from 5% to 1% changes the apparent rate constant for single copy reaction from 0.00032 to 0.00024 L mol^{-1} s^{-1}. In addition, the presence of a tenfold repeated fraction makes accurate determination of the single copy rate constant more difficult.

The slave "mini-C_0t" reaction (Britten et al., 1974) provides more accurate information about the moderately repeated fraction. Isolation of the fraction excludes many of the sequences repeated less than ten times, however. The mini-C_0t analysis certainly suggests there are two definite frequency classes of repeated DNAs.

Measurements on Repeated Sequence Elements. The number, size, and spacing of the repeated sequences have been measured by melting, nuclease digestion, and interspersion experiments. Each of these experiments measures one number well and provides other information which should be consistent. In this section we will try to collate the data from each of the experiments to find agreement on the fraction of the genome which is repetitive and on the length of the repeated sequences.

The Fraction Repetitive. Melting presents the best data on

the fraction repetitive. Nuclease digestion may destroy mismatched duplexes which are considered repeated by standard hydroxylapatite fractionation. The interspersion data do not give precise values for the fraction repetitive or length of the repeated sequences. Melting data are difficult to interpret because of the large contribution of collapsed single-stranded tails when only a small fraction (<20%) of the DNA is duplexed. We have tried to control the problem by immediately remelting the samples and subtracting the hyperchromicity due to single strands. Some of the single-stranded hyperchromicity is due to rapidly reannealing sequences, however, and the simple calculation:

$$\frac{\text{single strand hyperchromicity} - \text{duplex hyper.}}{\text{single strand hyper.} - \text{native hyper.}}$$

gives 13% duplex at C_0t 5 and 19% at C_0t 50. The single-stranded hyperchromicity corrections were made at 60 °C; this excludes contributions from very short duplexes which melt below 50 °C. A 5% rapidly renaturing fraction subtracted out by the single-stranded correction must be added to the 13% and 19% to find the total fraction in duplex: 18% at C_0t 5 and 24% at C_0t 50.

These numbers agree closely with the nuclease values presented in Table IV. The enzyme ratios used in these experiments selectively clip single-stranded tails and do not attack slightly mismatched duplexes (Britten et al., 1976). Our studies with a range of enzyme ratios and salt concentrations suggest that poorly matched duplexes are exposed to the same criterion with nuclease digestion as with optical melts and hydroxylapatite fractionation. The interspersion curves at C_0t 5 and C_0t 50 are also consistent with 18% of the genome in duplex at C_0t 5 and 24% in duplex at C_0t 50.

Repetitive Sequence Length. The best measure of repetitive sequence lengths comes from nuclease digestion studies. These studies display the real distribution of sizes on agarose A-50. About 40–50% of the repeated DNA sequences are excluded from A-50 while the remainder are included with a length of 0.2–0.3 kbp.

These lengths are similar to results from the sea urchin (Davidson et al., 1973a,b; Britten et al., 1976), *Aplysia* (Angerer et al., 1975), a host of marine invertebrates (Goldberg et al., 1975), and an insect (Efstratiadis et al., 1976). The distribution of excluded/included material is variable. This distribution may be much more sensitive to enzyme concentration than the fraction of DNA bound to hydroxylapatite and in some cases excluded material may include short sequences attached by poorly matched, nuclease susceptible regions. The variability may also be due to complementary long and short sequences. Short duplex regions on long molecules may have a much higher sensitivity to slight changes in incubation conditions.

If 60% of the mass of the repeated sequences are 0.275-kbp long and the remainder 1.5-kbp long, the weight-average length is $(0.6 \times 0.275) + (0.4 \times 1.5) = 0.765$ kbp. Under ideal conditions the duplex melting experiment would provide the weight-average number while the interspersion data would provide a number-average estimate. The number from the melting experiments ranges from 0.225-kbp duplex length on 0.35-kbp fragments to 0.47-kbp duplex on 1.6-kbp fragments. This may be due to the relatively short (0.3, 1.1, and 1.6 kbp) DNA fragment lengths used in the experiment. On the average, even the longest fragment length, 1.6 kbp, could only form an 0.8-kbp duplex (Smith et al., 1975) when two long repetitive sequences renature. Long repetitive sequences would not account for their full fraction of hyperchromicity until fragments longer than 3.0 kbp were reacted. Such reactions are technically difficult; 3.0-kbp long fragments form duplexes which cannot be eluted from hydroxylapatite with 0.5 M phosphate and must be denatured.

The interspersion long tracer/short driver experiments provide a number-averaged repetitive sequence length near 1.0 kbp. It is difficult to interpret these values; the extrapolation of the interspersion data to zero fraction bound is dependent on assumptions about the interspersion of the long fragments.

A consistent model for repeated sequences must explain the 0.35-kbp hydroxylapatite binding data. According to these data, 24% of the fragments contain duplex at C_0t 5 and 31% contain duplexes at C_0t 50. This contrasts with 18% duplex at C_0t 5 and 24% duplex at C_0t 50 measured by melting and nuclease digestion. This implies 75% of the strands bound on hydroxylapatite are duplexed, in agreement with the melting data in Table III (65% in duplex). If 0.6 of the repeated sequences are about the same length as the renatured fragments (0.3 kbp from A-50 fractionation), 50% of those fragments should be duplexed due to random overlap (Smith et al., 1975; Britten and Davidson, 1976). The same argument gives 1.35/1.5 or 90% of the 1.5-kbp long fragments in duplex. This implies $0.6 \times 0.5 + 0.4 \times 0.9 = 0.66$ of the strands should be duplexed. The agreement between the 75% fraction duplex after hydroxylapatite binding and the 66% expected from the repetitive sequence length is good. The more repeated sequences will be more than 50% covered but we may have overestimated the fraction of short repeats. In any case the hydroxylapatite data are in excellent agreement with our estimates for the fraction and length of the repetitive duplex.

These estimates for sequence length are also similar to measurements made on 1.0–3.0-kbp fragments renatured and visualized in the electron microscope (Bonner et al., 1973; Wilkes et al., 1978). Three lengths of DNA were stripped of rapidly renaturing sequences and renatured to C_0t 50. The duplexes formed from 0.9-kbp fragments had a number-averaged length of 0.33 kbp and a weight average of 0.54 kbp. Duplexes formed from 1.7 kbp strands had number and weight average lengths of 0.38 and 0.68 kb, while 0.45- and 0.99-kbp duplexes were formed by 2.5-kbp strands. The increase in average duplex length with fragment length is due to the long repeated sequences. Longer fragments can form longer structures identifiable as duplexes in the electron microscope. Duplexes formed by 2.5-kbp fragments have a weight average close to the 0.77 kbp predicted from our nuclease experiments. Duplexes formed by shorter fragments display the length of the short repeated fraction 0.2–0.4 kbp. Electron microscopy experiments by Chamberlin et al. (1975) find the short repetitive sequence length in *Xenopus* is 0.35 kbp, in good agreement with our findings. The fragment size used in the *Xenopus* experiments precluded measurements on long duplex sequences.

While the data on the fractions of long and short repetitives (Table V) are variable, we believe they reflect distinct lengths of repetitive sequences in rat DNA. Our numbers suggest the rat genome can be separated into the repetitive and single copy classes shown in Table VI. Seventy percent of the DNA is single copy, 25% repeated from 100 to 100 000-fold and 5% contains sequences in duplex before C_0t 0.05. The 16% in duplex by C_0t 5 can be divided into a long fraction with 6.4% of rat DNA and a short 0.2–0.3-kbp fraction with 9.6% of the DNA.

The long sequences differ from shorts not only in size but also in fidelity of matching. The melting data show the same melting temperature dependence on length reported by Da-

TABLE VI: Sequence Distribution in Rat DNA.

Component	Fraction[a]	Size distribution	Repetition	Complexity (nucleotides)	No. of sequences
1	0.02[b]		>100 000	530	
2	0.21[c]	0.4 > 1.5 kbp	3000	1.85×10^5	<50
3	0.10[d]	0.6 (0.2-0.4 kbp)	12.5	2.12×10^7	350
4	0.67	0.75 (2.5 kbp) 0.25 >6.0 kbp	1	1.79×10^9	530 000 <70 000

[a] Fraction of DNA is our best estimate of the true fraction duplex. These values differ from those in Table I because of single-stranded tails bound to hydroxylapatite. [b] This fraction assumes that 0.67 of the 0.08 of the duplex bound before C_0t 0.05 is 3000-fold repeated sequence (Wu et al., in preparation). [c] This fraction includes 0.06 from the binding before C_0t 0.05. [d] This is a rough estimate. None of the nuclease digestion studies separated the low repeat fraction from single copy DNA. From the hydroxylapatite data of Table I, 0.18 of the genome consists of slow repeats. We have assumed 0.55 of that DNA is true duplex.

vidson et al. (1973a,b), and Goldberg et al. (1975). The T_m differences between long and short repeated sequences are not simply due to the differences in length. The $650/n$ equation expressing dT_m as a function of nucleotide length n (Britten et al., 1974) would only predict a 3 °C increase in T_m for 0.2 kbp sequences while the difference is 15 °C. This corresponds to 12% mismatch in short repeats assuming 1.0% mismatch per °C decrease in T_m (Britten et al., 1974). The long and short fragment melting curves show more than one transition, however. There are long sequences with low precision regions and high precision short repeated sequences.

Interspersion of Repeated Sequences in Rat DNA. Almost every one of the results we have presented shows qualitative evidence for interspersion of short repeated sequences with single copy DNA in the rat. If repeated sequences were not interspersed, we could not measure the duplex length by digesting away adjacent single strands. This is not a trivial result; similar experiments by Crain et al. (1976) including S1 nuclease digestion and melting failed to detect interspersed repetitive sequences in *Drosophila*.

The best data on repeated sequence interspersion come from the short driver/long labeled tracer experiments shown in Figure 6. The range of least-squares fits which can be imposed on the data is indicated in the figure. Changes in the fraction repetitive move the inflection point for the short interspersion period. When the nuclease and melting data are used to constrain the fit to 16% and 25% repetitive at zero single strand tail length the short interspersion periods are 2.4 and 2.7 kbp at C_0t 5 and C_0t 50. The similarity of interspersion periods at those two C_0t's suggests different frequency repeats are probably not arranged differently. At these interspersion periods, 57% and 73% of the DNA are organized as short interspersed repeats while another 15% and 10% of the DNA are interspersed with a period longer than 6.0 kbp.

Electron microscopy experiments mapping short duplex fragments on very long (10–50 kbp) DNA have also been used to determine the interspersion period in the rat (Bonner et al., 1973; Wilkes et al., 1978). These data cannot show the same division of interspersion periods shown in Figure 6 but show a large class of spacings between 1.0 and 3.0 kbp.

These data are in excellent agreement with measurements on human DNA (Schmid and Deininger, 1975) and are not significantly different from measurements on *Xenopus* (Davidson et al., 1973a,b), sea urchin (Graham et al., 1974), or *Aplysia* (Angerer et al., 1975). The data in Figure 6 are open to a range of interpretations. A model of interspersion in the rat which included three interspersion periods, at 0.8, 4.0 and >6.0 kbp, would fit the data equally well.

Conclusions

Table VI presents a model for the number and length of the different sequence fractions in rat DNA. Nearly 70% of the DNA is single copy; the exact fraction is obscured by the low frequency (10-fold) repeated fraction of the DNA seen in Figures 1 and 2. We estimate the slow repeated fraction contains 10% of the DNA. The major repetitive component is repeated about 3000-fold and includes 20% of the DNA. From S1 nuclease digestion and agarose chromatography, we can measure the length of the repeated sequences. Sixty percent of the moderately repetitive sequences are 0.3-kbp long and 40% are longer than 1.5 kbp. From this length distribution we calculate there are 50 different families of sequences 1.5-kbp long repeated 3000-fold and 350 different 0.3-kbp sequences with the same repetition frequency.

The interspersion measurements also allow us to calculate the length and number of the single copy sequences. There are 530 000 single copy sequences 2.5-kbp long and another 70 000 (or fewer) sequences longer than 6.0 kbp. There are enough short repeated sequences (1 050 000) to bound each of the short single copy sequences.

The highly structured interspersion of the short repeated DNA sequences, the universality of this pattern of organization, and the proximity of expressed sequences to subsets of repeated sequences all suggest an important functional role for repeated sequences. Britten and Davidson (1969) have presented a model for regulatory function of repeated DNA sequences which is consistent with a wide range of structural and functional features of DNA sequences discovered since the model was proposed.

The distribution and function of the long repeated DNA sequences are unclear. Some of this DNA must contain the known repetitive gene families such as the ribosomal and histone genes (Brown and Sugimoto, 1974; Birnstiel et al., 1974; for discussion, see Galau et al., 1976) and must be distinct from short repeated sequences. Long repeated DNA may contain other sequences which are shared with short repeated sequences. Some repeated sequences may sometimes be bounded by single copy sequences and other times be bounded by other repeated sequences as part of a long repeated sequence. Although we have begun to explore the sequence relationships between long and short repeated DNA sequences (Wu et al., 1977), measurement of sequences shared by these two size classes is technically difficult. Many short fragments may be derived from long repeated sequences because of mechanical shear. Purification of long sequences is easier but it is difficult to design experiments which are insensitive to 5–10% con-

tamination by short repeated sequences. If sequence overlap can unambiguously be demonstrated, it may provide evidence for the "integrator" gene sequence postulated by Britten and Davidson (1969).

Acknowledgment

We would like to thank Dr. Francine Eden for helpful discussions and the kind gift of the S1 nuclease preparation, Mr. Bill Buchanan for his assistance, and Mr. James Posakony for his technical assistance and careful reading and valuable comments on the manuscript.

References

Angerer, R. C., Davidson, E. H., and Britten, R. J. (1975), *Cell 6,* 29–39.

Birnstiel, M., Telford, J., Weinberg, E., and Stafford, D. (1974), *Proc. Natl. Acad. Sci. U.S.A. 71,* 2900–2904.

Bonner, J., Garrard, W. T., Gottesfeld, J. M., Holmes, D. S., Sevall, J. S., and Wilkes, M. M. (1973), *Cold Spring Harb. Symp. Quant. Biol. 38,* 303–310.

Britten, R. J., and Davidson, E. H. (1969), *Science 165,* 349–357.

Britten, R. J., and Davidson, E. H. (1976), *Proc. Natl. Acad. Sci. U.S.A. 73,* 415–419.

Britten, R. J., Graham, D. E., and Neufeld, B. R. (1974), *Methods Enzymol. 29E,* 363–418.

Britten, R. J., Graham, D. E., Eden, F. C., Painchaud, D. M., and Davidson, E. H. (1976), *J. Mol. Evol. 9,* 1–23.

Brown, D. D., and Sugimoto, K. (1974), *Cold Spring Harb. Symp. Quant. Biol. 38,* 501–505.

Chamberlin, M. E., Britten, R. J., and Davidson, E. H. (1975), *J. Mol. Biol. 96,* 317–333.

Crain, W. R., Eden, F. C., Pearson, W. R., Davidson, E. H., and Britten, R. J. (1976), *Chromosoma 56,* 309–326.

Dahmus, M. E., and McConnell, D. J. (1969), *Biochemistry 8,* 1524–1534.

Davidson, E. H., and Britten, R. J. (1973), *Q. Rev. Biol. 48,* 565–613.

Davidson, E. H., Graham, D. E., Neufeld, B. R., Chamberlin, M. E., Amerson, C. S., Hough, B. R., and Britten, R. J. (1973a), *Cold Spring Harb. Symp. Quant. Biol. 38,* 295–301.

Davidson, E. H., Hough, B. R., Amenson, C. S., and Britten, R. J. (1973b), *J. Mol. Biol. 77,* 1–23.

Davidson, E. H., Galau, G. A., Angerer, R. C., and Britten, R. J. (1975a), *Chromosoma 51,* 235–259.

Davidson, E. H., Hough, B. R. Klein, W. R., and Britten, R. J. (1975b), *Cell 4,* 217–238.

Efstratiadis, A., Crain, W. R., Britten, R. J., Davidson, E. H., and Kafatos, F. (1976), *Proc. Natl. Acad. Sci. U.S.A. 73,* 2289–2293.

Galau, G. A., Chamberlin, M. E., Hough, B. R., Britten, R. J., and Davidson, E. H. (1976), in Molecular Evolution, Ayala, F. J., Ed., Sunderland, Mass., Sinauer Association, Chapter 12, pp 200–224.

Goldberg, R. B., Crain, W. R., Ruderman, J. V., More, G. P., Barnett, J. R., Higgins, R. C., Gelfand, R. A., Galau, G. A., Britten, R. J., and Davidson, E. H. (1975), *Chromosoma 51,* 225–251.

Gottesfeld, J. M., Bagi, G., Berg, B., and Bonner, J. (1976), *Biochemistry 15,* 2472–2483.

Graham, D. E., Neufeld, B. R., Davidson, E. H., and Britten, R. J. (1974), *Cell 1,* 127–137.

Holmes, D. S., and Bonner, J. (1974a), *Biochemistry 13,* 841–848.

Holmes, D. S., and Bonner, J. (1974b), *Proc. Natl. Acad. Sci. U.S.A. 71,* 1108–1112.

Manning, J. E., Schmid, C. W., and Davidson, N. (1975), *Cell 4,* 141–155.

Noll, H. (1967), *Nature (London) 215,* 360–363.

Pearson, W. R., Davidson, E. H., and Britten, R. J. (1977), *Nucleic Acids Res. 4,*1727–1735).

Plageman, P. G. W. (1969), *Biochim. Biophys. Acta 182,* 46–56.

Plageman, P. G. W., and Swimm, H. E. (1966), *J. Bacteriol. 91,* 2317–2326.

Schmid, C. W., and Deininger, P. L. (1975), *Cell 6,* 345–358.

Smith, M. J., Hough, B. R., Chamberlin, M. E., and Davidson, E. H. (1974), *J. Mol. Biol. 85,* 103–126.

Smith, M. J., Britten, R. J., and Davidson, E. H. (1975), *Proc. Natl. Acad. Sci. U.S.A. 72,* 4805–4809.

Sober, H. A. (1968), In Handbook of Biochemistry, Cleveland, Ohio, Chemical Rubber Co., p H-58.

Studier, F. W. (1965), *J. Mol. Biol. 11,* 373–390.

Wilkes, M. M., Pearson, W. R., Wu, J. R., and Bonner, J. (1978), *Biochemistry* (following paper in this issue).

Wu, J. R., Pearson, W. R., Posakony, J. W., and Bonner, J. (1977), *Proc. Natl. Acad. Sci. U.S.A. 74,* 4382–4386.