352-358.
Pierce, J. G., Bahl, O. P., Cornell, J. S., and Swaminathan, N. (1971a), *J. Biol. Chem. 246,* 2321-2324.
Pierce, J. G., Liao, T. H., Howard, S. M., Shome, B., and Cornell, J. S. (1971b), *Recent Prog. Horm. Res. 27,* 165-212.
Rathnam, P., and Saxena, B. B. (1974) *J. Biol. Chem. 250,* 6735-6746.
Rodbard, D. (1974), *Clin Chem. 20,* 1255-1270.
Sairam, M. R., and Li, C. H. (1973), *Biochem. Biophys. Res. Commun. 54,* 426-432.
Shome, B., and Parlow, A. F. (1973), *J. Clin. Endocrinol. Metab. 39,* 618-621.
Shome, B., and Parlow, A. F. (1974), *J. Clin. Endocrinol. Metab. 39,* 203-205.
Stevens, V. C. (1976), *Excerpta Med. Int. Congr. Ser. No. 403,* 379-385.
Swaminathan, N., and Bahl, O. P. (1970), *Biochem. Biophys. Res. Commun. 40,* 422-427.
Swaminathan, N., Rasor, J., and Braunstein, G. D. (1978) *Immunochemistry* (in press).
Vaitukaitis, J. L., Braunstein, G. D., and Ross, G. T. (1972), *Am. J. Obstet. Gynecol. 113,* 751-758.
Ward, D. N., Reichert, L. E., Liu, W. K., Nahn, H. S., Hsia, J., Lamkin, W. M., and Jones, N. S. (1973) *Recent Progr. Horm. Res. 29,* 533-561.

Unusual Organization of DNA Sequences in the Chicken†

Francine C. Eden* and John P. Hendrick

ABSTRACT: Single-copy and repeated sequences in the chicken genome were identified by measuring the reassociation kinetics of short fragments of DNA by hydroxylapatite chromatography. Eighty-seven percent of the DNA is single copy and 13% is repetitive. The repetitive DNA can be divided into two classes based on repetition frequency; about one-third and two-thirds of the sequences are repeated 15- and 1500-fold, respectively. The arrangement of single-copy and repeated sequences was determined by measuring the fraction of labeled DNA fragments bound to hydroxylapatite by repetitive duplexes as a function of the DNA fragment length. DNA sequence organization was also studied by electron microscopy of reassociated DNA. The results indicate that 40% of the chicken genome is arranged in an alternating pattern of single-copy and repetitive sequences, with single-copy regions about 4500 nucleotides in length. A smaller fraction of the genome has this same alternating arrangement but contains even longer single-copy regions. Almost one half of the DNA consists of very long single-copy regions, uninterrupted by repeated sequences for distances of at least 17 500 nucleotides. To measure the lengths of repeated sequences directly, long fragments of chicken DNA were reassociated until the repeated sequences had formed duplexes, but the single copy sequences had not yet done so. Single strands were digested away using S1 nuclease, and the lengths of repetitive duplexes were measured by gel-filtration chromatography. Most chicken repeated sequences are at least 2000 nucleotides long. These characteristics place the chicken genome in a category intermediate between the extremes of sequence organization previously described.

Exploration of DNA sequence organization in a number of eukaryotic genomes has revealed a regular alternating arrangement of repetitive and single-copy sequence elements. A short-period interspersion pattern, in which single-copy sequences 1000-1500 nucleotides in length are flanked by repetitive sequences 300 nucleotides in length, characterizes the genomes of a number of animals (Davidson et al., 1973; Graham et al., 1974; Angerer et al., 1975; Firtel and Hindle, 1975; Schmid and Deininger, 1975; Efstratiadis et al., 1976; Crain et al., 1976b; Smith and Boal, 1978) and plants (Zimmerman and Goldberg, 1977; Walbot and Dure, 1976). This pattern of sequence organization is represented widely in both vertebrate and invertebrate phyla (Davidson et al., 1975).

A few clear exceptions to this general pattern have been discovered. For example, the genomes of three insects, *Drosophila melanogaster* (Manning et al., 1975; Crain et al., 1976a), *Chironomus tentans* (Wells et al., 1976), and the honeybee *Apis mellifera* (Crain et al., 1976a) and the mold *Achlya* (Hudspeth et al., 1977) are characterized by long-period interspersion, in which single-copy regions are uninterrupted by repeated sequences for distances of many thousands of nucleotides. In the genome of *Drosophila melanogaster,* the repetitive sequences average 3500 nucleotides in length, and the average single-copy length exceeds 13 000 nucleotides (Manning et al., 1975). Although most of these examples of long-period interspersion occur in insects, the genomes of two other insect species, the silkmoth *Antheraea pernyi* and the housefly *Musca domestica,* contain short-period interspersion as the predominant pattern (Efstratiadis et al., 1976; Crain et al., 1976b). The occurrence of both long- and short-period interspersion even within the same insect order (*Drosophila* and *Musca* are both Diptera) suggests that patterns of DNA sequence arrangement may undergo rapid evolutionary change.

The occurrence of these contrasting patterns of DNA sequence organization has provided the impetus for a detailed investigation of other animal genomes. Many specific issues can be addressed by extending our present knowledge of DNA sequence organization. The broad phylogenetic occurrence of short-period interspersion and its striking similarity in diverse animal species suggest that this pattern arose shortly after the inception of metazoan life and confers some evolutionary advantage. However, relatively few individual species have been

† From the Laboratory of Biochemistry, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20014. *Received July 27, 1978.*

well characterized, and it remains possible that major classes of invertebrates and vertebrates contain other different interspersion patterns. It is important to establish whether conservation of patterns of DNA sequence organization or rapid alterations, as suggested by the situation in Dipteran insects, has been the predominant mode of evolution of DNA sequences.

Although each animal species studied so far can be unequivocally placed in the long- or short-period interspersed category, most present measurements indicate that in those genomes in which short-period interspersion predominates there is also a minor DNA fraction with more widely spaced interspersed repetitive sequences (Davidson et al., 1973; Graham et al., 1974; Angerer et al., 1975; Schmid and Deininger, 1975; Efstratiadis et al., 1976). Conversely, two of the three insects with long-period interspersion may also contain a minor fraction of DNA in the short-period interspersion pattern; *Drosophila melanogaster* is the only organism with long-period interspersion whose genome has been sufficiently characterized to completely exclude even a minor fraction of DNA in the short-period interspersion pattern (Crain et al., 1976a).

A very interesting feature of present knowledge about DNA sequence organization is that no pattern intermediate between long- and short-period interspersion has been previously described. We report here that the DNA of the chicken, *Gallus gallus,* falls into an intermediate category. About 50% of the genome contains interspersed repetitive and single-copy sequences. The interspersed single-copy sequences are about 4500 nucleotides in length. There appears to be a small interspersed fraction of DNA with even longer single-copy sequences. The remainder of the chicken genome (at least one third) consists of single-copy sequences uninterrupted by repetitive sequences for a distance of at least 17 000 nucleotides. DNA in the short-period interspersion pattern was not detected.

Materials and Methods

Preparation of DNA. Chicken DNA was prepared by grinding frozen liver with dry ice in a Waring blendor, and DNA was purified as described previously (Britten et al., 1974). The optical hyperchromicity of purified DNA, from 60 to 100 °C, was 26% of the maximum optical density (260 nm), measured at 100 °C.

Sea urchin DNA was extracted from sperm as described (Britten et al., 1974).

Radioisotope Labeling of DNA. Chick embryo fibroblasts were grown in Dulbecco's modification of Eagle's medium with 10% fetal calf serum at 37 °C, in the presence of 50 μCi/mL [^{3}H]thymidine (New England Nuclear, 49.8 Ci/mmol), for 48 h. DNA was extracted from cells scraped from ten 100-mm diameter petri plates by a procedure described previously (Eden et al., submitted for publication). About 100 μg of purified DNA was obtained, with a specific radioactivity of 7.3 × 10^6 cpm/μg and a single-strand length of 11 000 nucleotides.

Alternatively, DNA was labeled in vitro by nick translation as described previously (Crain et al., 1976a).

Chicken DNAs labeled in vitro or in vivo, respectively, were indistinguishable with respect to the fraction bound to hydroxylapatite at C_0t 50 as a function of DNA fragment length.

Fragmentation of DNA and Measurement of Fragment Lengths. Unlabeled DNA fragments of desired lengths were prepared by shearing in a Virtis 45 or 60 homogenizer as described (Britten et al., 1974). Labeled DNA fragments of desired lengths were prepared using isokinetic alkaline sucrose gradients according to Noll (1967).

DNA fragment lengths were measured in duplicate determinations using isokinetic alkaline sucrose gradients in an SW 40 rotor. Every gradient contained two well-characterized sheared DNAs as internal length standards. The lengths of the sheared DNAs were determined by cosedimentation with restriction fragments of DNA.

DNA Reassociation and Hydroxylapatite Chromatography. DNA was reassociated in 0.12 M PB[1] at 60 °C or 0.4 M PB at 68 °C. Equivalent C_0t for the 0.4 M PB, 68 °C condition was calculated by multiplying C_0t by 4.9 (Britten et al., 1974). Hydroxylapatite chromatography was performed according to Britten et al. (1974). Radioactivity of bound and unbound fractions was determined by counting in 0.08 M PB and Instagel (Packard), in a Packard Tricarb liquid scintillation spectrometer. Some DNA reassociation data were normalized to exclude small amounts of residual foldback DNA and 5–10% unreactable DNA, prior to computerized data analysis.

Electron Microscopy. DNA samples were prepared for electron microscopy by a modified Kleinschmidt procedure (Davis et al., 1971), as previously described (Chamberlin et al., 1975). DNA was spread from 57% formamide, 0.1 M Tris (pH 8.5), 0.01 M EDTA, onto a hypophase of 17% formamide, 0.01 M Tris (pH 8.5), 1 mM EDTA. The hyperphase contained 33 μg/mL cytochrome *c* and 1 μg/mL DNA. Samples were picked up on parlodian-coated grids, stained with uranyl acetate (5 × 10^{-5} M), and rotary shadowed with platinum/palladium (80:20). Pictures were taken with a Siemens 102 electron microscope, using a 50-μm objective aperture and 60-kV accelerating voltage at a magnification of 15 000×. Micrographs were printed at a final magnification of 75 000× and DNA contour lengths were measured directly from the photographs using a Hewlett-Packard digitizer. Lengths in nucleotides were calculated from the relationship 1 μm of DNA = 3000 nucleotides.

S1 Nuclease Digestion. We used an S1 nuclease preparation isolated from *Aspergillus oryzae* and quantitated as previously described (Britten et al., 1976). Before digestion, DNA samples were reassociated in 0.3 M NaCl, 0.01 M Pipes (Sigma), pH 6.7, at 64 °C. Equivalent C_0t was calculated by multiplying C_0t by 2.3 (Britten et al., 1974). Reassociated DNA was adjusted to 0.025 M sodium acetate buffer (pH 4.3), 0.1 mM $ZnSO_4$, and 25 mM mercaptoethanol and digested with S1 nuclease at 37 °C for 45 min. Sufficient S1 nuclease was added to ensure complete digestion of single strands without digestion of mismatched duplexes (Britten et al., 1976).

Measurement of Optical Hyperchromicity of DNA. Optical hyperchromicity of DNA samples was determined using a Gilford Model 240 spectrophotometer equipped with a recorder and thermal programmer. Temperature and optical density at 260 nm were recorded continuously as the temperature was raised from 30 to 98 °C at a rate of 1 °C/min. The percent hyperchromicity was calculated by the formula:

$$H = \frac{\text{OD at 98 °C} - \text{OD at 60 °C}}{\text{OD at 98 °C}} \times 100$$

[1] Abbreviations used: PB, equimolar mono- and disodium phosphate buffer (pH 6.8); T_m, midpoint of a thermal denaturation curve at which 50% of the DNA has become single stranded; C_0t, DNA concentration in moles of nucleotides/L × time (in seconds); K, reassociation rate constant (M^{-1} s^{-1}); rms, root mean square deviation; EDTA, ethylenediaminetetraacetic acid; $NaDodSO_4$, sodium dodecyl sulfate; kb, kilobase; nt, nucleotide, ntp, nucleotide pair; OD, optical density; HAP, hydroxylapatite; Tris, 2-amino-2-hydroxymethyl-1,3-propanediol; Pipes, 1,4-piperazinediethanesulfonic acid.

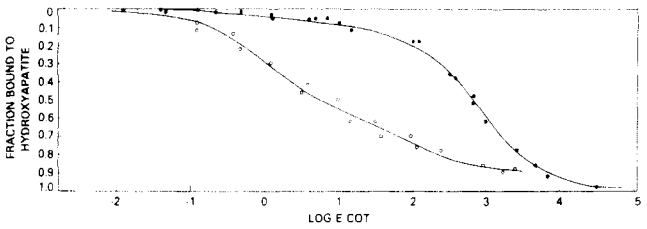

FIGURE 1: Reassociation kinetics of chicken DNA. (●) Total chicken DNA, labeled in vitro, was sheared to fragments 400 nucleotides in length. Foldback sequences were removed by incubation to $C_0t = 10^{-4}$ and isolation of the fraction that did not bind to hydroxylapatite. Labeled DNA fragments were reassociated with excess unlabeled chicken DNA (550 nucleotides in length) as described under Materials and Methods. Data have been normalized to exclude a small fraction of residual foldback DNA and 5% unreactable DNA. The reassociation curve was generated by computer analysis of the data points and represents a composite of three components: 88%, $K = 0.0011$; 5.5%, $K = 0.017$; and 7.5%, $K = 1.66$. Root mean square deviation = 1.5%. (○) Chicken repetitive DNA was isolated by successive incubations to repetitive C_0t (C_0t = 50, 20, 10) and isolation of the hydroxylapatite-bound fraction after each incubation. After in vitro labeling and removal of foldback sequences, the repetitive [^{3}H]DNA (390 nucleotides in length) was reassociated with a 30 000- to 60 000-fold excess of unlabeled total DNA 550 nucleotides in length. The curve drawn was derived by computer analysis and represents two components: 55%, $K = 0.95$; and 33%, $K = 0.01$. Root mean square deviation = 2.7%. The remaining 12% of the DNA may represent contaminating single-copy sequences in the isolated repetitive fraction.

Results

Single-Copy and Repetitive Sequence Elements in Chicken DNA. Figure 1 shows the reassociation kinetics of short fragments of chicken DNA measured by hydroxylapatite chromatography. The reaction is dominated by a single-copy component which constitutes 87% of the total DNA. The remaining 13% of chicken DNA consists of repetitive sequences. These results are in substantial agreement with those already reported (DeJiminez et al., 1974; Arthur and Straus, 1978). To obtain an accurate determination of sequence representation in the small fraction of chicken DNA that is repetitive, repeated sequences were isolated, labeled in vitro by nick translation, and reassociated with an excess of total DNA (Figure 1). We could distinguish two repetition frequency components comprising 33 and 55% of the repetitive DNA or 5.5 and 7.5% of total DNA, respectively. The characteristics of these components are summarized in Table I. About two-thirds of chicken repetitive sequences have an average repetition frequency of 1500-fold; the minor repetitive component has a much lower repetition frequency, 10- to 20-fold. Both repetition frequency components can also be identified in the reassociation curve of total DNA. The low repetition frequency or "slow" component is more difficult to distinguish kinetically from single-copy sequences in total DNA. Since measurements of sequence interspersion rely on a clear discrimination between repetitive and single-copy sequences, subsequent measurements refer principally to the arrangement of the major rapidly reassociating "fast" repetitive component relative to single-copy DNA.

The reassociation curve of total DNA is characterized by the absence of detectable reassociation between a C_0t of 10^{-4} and 10^{-2}, indicating that chicken DNA contains very little satellite or highly repeated DNA, which would be expected to reassociate in this range of C_0t. The chicken genome can be adequately described for present purposes using three frequency classes of DNA, fast and slow repetitive sequences, and single-copy DNA (Table I).

The size of the chicken genome was calculated from the reassociation rate constant of single-copy DNA. We used as a standard the reassociation rate constant of the sea urchin single-copy component (0.00125 M^{-1} s^{-1}), corresponding to a haploid genome size of 0.88 pg. A small correction was made for the fragments 550 nucleotides in length used here compared to fragments 450 nucleotides in length employed in the sea urchin measurement (Graham et al., 1974). The haploid genome size of the chicken was calculated to be (0.00125 M^{-1} s^{-1}/0.0010 M^{-1} s^{-1}) × 0.88 pg = 1.1 pg.

TABLE I: Reassociation Kinetic Analysis of Chicken DNA.

component[a]	fract of total DNA	K (M^{-1} s^{-1})	av repetition[b] frequency
fast	0.075	1.66	1500
slow	0.055	0.017	15
single copy	0.870	0.0011	1

[a] The fraction of each repetitive component and its reassociation rate constant, K, were determined from analysis of the reassociation curve of total DNA. These values are in substantial agreement with those derived from the reassociation of an isolated repetitive fraction with excess total DNA. [b] Determined by comparing the rate of reassociation of each component in total DNA with the single-copy rate. These values may represent averages of many components not individually identified.

Interspersion of Single-Copy and Repetitive Sequences. Interspersion refers to an alternating arrangement of single-copy and repetitive DNA sequences. We initiated a study of sequence interspersion in chicken DNA with measurements of the type that have provided the most substantial evidence for sequence interspersion in those animal genomes studied so far (Davidson et al., 1973; Graham et al., 1974; Angerer et al., 1975; Schmid and Deininger, 1975; Efstratiadis et al., 1976). The method relies on the fact that if DNA fragments of a given length bear both a repetitive and a single-copy sequence the whole fragment will bind to hydroxylapatite, even if only the repetitive sequence has reassociated. The fraction of fragments of a given length bearing a repetitive sequence element can be directly measured, and the fraction of the genome interspersed and the predominant lengths of single-copy sequences can be derived from this measurement.

Using the reassociation rate constant of each component in the chicken genome (Table I), we calculated the C_0t at which the majority of repetitive sequences will have reassociated but single-copy sequences will not yet have done so. Interspersion of single-copy and repetitive sequences in chicken DNA was measured at $C_0t = 50$, which includes 99% of the fast and 46% of the slow repetitive sequences, with less than 5% reassociation of single-copy sequences.

Figure 2 shows the fraction of labeled fragments bound to hydroxylapatite at C_0t 50 when labeled fragments of various lengths were reassociated with excess unlabeled total chicken DNA. The fraction bound to hydroxylapatite increased dramatically as the tracer fragment length was increased from 470 to 4500 nucleotides, indicating the presence of an interspersed fraction of chicken DNA. As the fragment length was further increased from 4500 to 17 500 nucleotides, there was some increase in hydroxylapatite binding with a curve of much lower slope. These data indicate that about 40% of the chicken genome consists of interspersed single-copy and repetitive sequences, with a predominant length of single-copy sequences of about 4500 nucleotides. A minor component of the chicken genome seems to contain even longer interspersed single-copy sequences.

From the observation that 45% of the chicken DNA is not bound to hydroxylapatite at $C_0t = 50$ even when very long

TABLE II: Absence of Shear Breakage of Long DNA Fragments during Hydroxylapatite Chromatography.

expt	length of DNA fragments[a] (nt) before HAP chromatogr	length of DNA fragments[a] (nt) after HAP chromatogr	fract of total[b] DNA bound to HAP
1	1770	1630	0.28
2	2500	1900	0.35
3	5400	4920	0.42

[a] [^{3}H]DNA fragments were reassociated with a 40 000-fold excess of unlabeled DNA (280 nt) to C_0t = 50. Prior to hydroxylapatite chromatography, an aliquot was removed for determination of [^{3}H]DNA length. An aliquot of the DNA not bound to HAP was used to measure the length of DNA strands after chromatography. DNA lengths were determined in duplicate using alkaline sucrose gradients. [b] Fraction of the labeled DNA bound to HAP, corrected for zero time binding measured in a parallel incubation without unlabeled DNA.

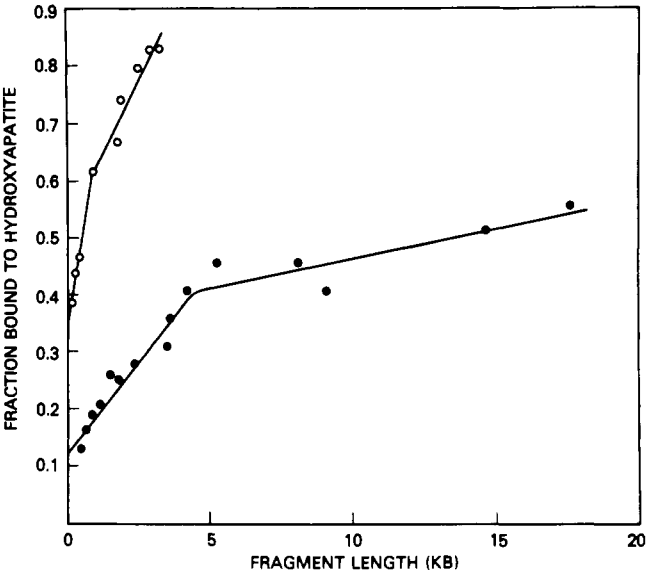

FIGURE 2: Fraction of chicken DNA containing repetitive sequences as a function of DNA fragment length. Labeled DNA fragments of various lengths were reassociated to C_0t = 50 with a 30 000- to 60 000-fold excess of unlabeled chicken DNA fragments 550 nucleotides in length. An aliquot was removed and counted for calculation of recovery, and the remainder was loaded onto 0.5 mL of hydroxylapatite at 60 °C in 0.12 M PB. The bound fraction was eluted with 0.12 M PB at 98 °C. To minimize shear breakage of DNA, no air pressue was used during hydroxylapatite chromatography: (●) chicken DNA: (○) labeled sea urchin DNA fragments of various lengths reassociated with excess unlabeled sea urchin DNA (450 nucleotides in length) to C_0t = 20. [Data reproduced from Graham et al. (1974) with the kind permission of the authors].

fragments are used, it appears that almost one-half of the chicken genome consists of single-copy regions whose lengths exceed 17 500 nucleotides. To confirm this result it is necessary to demonstrate that the failure of almost one-half of the longer fragments to bind to hydroxylapatite at C_0t = 50 does not result from shearing of the long DNA fragments during hydroxylapatite chromatography; that is, that the fraction not bound to hydroxylapatite does not consist of short, broken ends deriving from larger, duplex-containing structures. This was tested by measuring the lengths of the DNA fragments before and after hydroxylapatite chromatography. Table II shows the fraction of labeled DNA bound to hydroxylapatite at C_0t = 50 determined using DNA fragments whose lengths were measured before and after hydroxylapatite chromatography. There was only a slight decrease in the length of the DNA fragments after passage over hydroxylapatite, so we can rule out the possibility that the large fraction of the chicken genome not bound to HAP at C_0t = 50 is an artifact of the analytical procedure.

The data shown in Figure 2 could represent an underestimate of the fraction of the chicken genome that is interpersed if some of the DNA formed tangles or aggregates during reassociation that were not eluted from the hydroxylapatite columns. We can also exclude this possibility, since the recovery of radioactive DNA exceeded 95% for each column.

Extrapolation of the initial, rapidly rising portion of an interspersion curve to zero fragment length provides an independent estimate of the fraction of the genome that is repetitive. The data for interspersion in the chicken genome extrapolate to an intercept at 12%, in good agreement with the fraction of repetitive sequences (13%) determined by reassociation kinetics (Table I).

The presence of an interspersed fraction with single-copy sequences 4500 nucleotides in length in chicken DNA differs significantly from the predominant pattern of sequence organization in most animal genomes. This can be readily appreciated by referring to Figure 2, in which the more typical short-period interspersion pattern is represented by data for sea urchin DNA (Graham et al., 1974).

Preparative Isolation and Characterization of Interspersed and Noninterspersed Fractions of the Chicken Genome. To confirm the presence of interspersed and noninterspersed fractions within the chicken genome, [^{3}H]DNA fragments 5400 nt in length were reassociated with excess short, unlabeled DNA fragments to C_0t = 50. A parallel, control mixture contained ^{3}H-labeled DNA and heterologous, carrier DNA, incubated for the same time in the same volume, and was used to correct for zero time binding. At C_0t = 50, 54% of the [^{3}H]DNA incubated with chicken driver DNA and 12% of the control mixture bound to hydroxylapatite. Thus, the corrected binding at C_0t = 50, (54% − 12% = 42%), is in close agreement with that expected for fragments 5400-nt long (Figure 2). The fraction bound to hydroxylapatite (interspersed fraction) and the fraction not bound to hydroxylapatite (noninterspersed fraction) were isolated, and their content of single-copy and repetitive sequences was determined by reassociation of the ^{3}H-labeled sheared fragments with total DNA. The reassociation kinetics (Figure 3) indicate that the interspersed fraction contains both repetitive and single-copy sequences, whereas the noninterspersed fraction is principally single copy. Quantitation of single-copy and repetitive sequences in each fraction was carried out by computer analysis of the reassociation data according to Pearson (1977).

About 75% of chicken repetitive DNA should have reassociated at C_0t = 50 (calculated from Table I), constituting 10% of the genome. Thus, the interspersed fraction should contain 10/54 = 18.5% repeated DNA. The analysis shown in the upper portion of Table III is aimed at comparing the data with this prediction. During the first three trials, all the parameters were left free and the computer determined that the data are best fit by a composite curve having two components, a repetitive component (K = 0.95) comprising 10.6% of the DNA and a single-copy component (K = 0.00076) comprising 74% of the DNA. We then fixed the amount of repetitive DNA at values ranging from 0 to 20% and reevaluated the data. The relationship of the rms deviation of the data points from the curve drawn at each fixed value of the fraction of repetitive DNA is shown in Table III. The data are sufficient to specify the presence of 9–14% repeated DNA, with the apparent lowest rms deviation at about 12% repeated DNA. Thus, re-

TABLE III: Sequence Representation in Chicken Interspersed and Noninterspersed Fractions.[a]

fract of chicken genome	trial[b]	repetitive component				single copy component		final fract unreassoc	rms
		fract	K	fract	K	fract	K		
interspersed	1	0.093	1.0	0.014	0.83	0.740	0.00076	0.053	0.028
	2			0.106	0.99	0.739	0.00076	0.050	0.028
	3					0.771	0.001	0.080	0.040
	4			[0.05][c]	0.99	0.756	0.00088	0.066	0.032
	5			[0.125]	1.01	0.734	0.00072	0.048	0.028
	6			[0.150]	12.3	0.729	0.00069	0.043	0.030
	7			[0.200]	12.6	0.740	0.00078	0.055	0.035
noninterspersed	1			[0.025]	0.00095	0.780	0.00081	0.130	0.028
	2			[0.050]	0.00081	0.760	0.00081	0.120	0.028
	3			[0.100]	0.00082	0.710	0.00081	0.120	0.028

[a] ^{3}H-labeled interspersed and noninterspersed fractions of chicken DNA were prepared as described in the legend to Figure 3. [b] Reassociation kinetics of ^{3}H-labeled interspersed or noninterspersed DNA with total DNA were measured by hydroxylapatite chromatography. Data were analyzed by a least-squares computer program as described (Pearson et al., 1977). [c] Values enclosed in brackets were held constant during computer analysis.

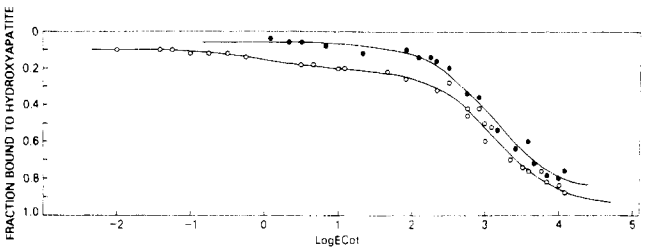

FIGURE 3: [^{3}H]DNA fragments 5400 nt in length (specific radioactivity 7.6×10^6 cpm/μg) were reassociated with a 40 000-fold excess of unlabeled DNA (280 nt in length) to $C_0t = 60$. The fractions that did and did not bind to hydroxylapatite (54 and 46% of the labeled DNA, respectively) were isolated, sheared to fragments 700–800 nucleotides in length, and reassociated with a 10- to 1000-fold excess of unlabeled total chicken DNA. The reassociation kinetics of the fraction bound to hydroxylapatite (interspersed fraction, open circles) and the fraction not bound to hydroxylapatite (noninterspersed fraction, closed circles), respectively, with total DNA, are shown. Reassociation kinetics were assayed by hydroxylapatite chromatography, and the data were analyzed by computer.

peated DNA constitutes $12/(12 + 74) = 14\%$ of the interspersed fraction.

A similar analysis applied to the noninterspersed fraction is shown in the lower portion of Table III. The data are best fit by a single component with a single-copy rate constant of 0.00081 M^{-1} s^{-1}, constituting 82% of the DNA. The remainder of the DNA consists of foldback (6%) and unreactable DNA (12%). Attempts to fit the data with two-component curves including 2.5, 5, and 10% repeated DNA failed, since the repetitive components were accommodated but given the same rate constant as single-copy DNA. Thus, we conclude that less than 2.5% repetitive DNA is contained in the noninterspersed fraction.

Sequence Interspersion Studied by Electron Microscopy. Since direct confirmation of short-period interspersion in the DNA of some animals has been obtained by observation of reassociated DNA in the electron microscope (Chamberlin et al., 1975; Deininger and Schmid, 1976; Manning et al., 1975), we used this method to determine whether a small fraction of DNA in the short-period interspersion pattern is also present in the chicken genome but was undetected using the methods described above.

The majority of reassociated repetitive duplexes are formed between single strands representing different members of the same repetitive family. Single-copy sequences adjacent to reassociated repetitive duplexes are usually nonhomologous and can be visualized as four single-stranded tails attached to a duplex region. These four-ended structures are diagnostic of sequence interspersion, and their presence leads directly to two conclusions: (a) the spacing of repetitive sequences is shorter than the fragments used, and (b) the individual repetitive sequence elements are also shorter than the fragments used. This second conclusion derives from the fact that fragments terminating with a repetitive sequence will be unable to form four-ended structures.

We prepared DNA fragments too short to detect interspersion of the longer periodicity typical of the majority of chicken DNA but appropriate for visualization of short-period interspersion if it is present. DNA fragments 2200 nucleotides in length were reassociated to $C_0t = 50$, bound to hydroxylapatite, and eluted with 0.4 M PB at 60 °C to preserve reassociated structures. The fraction bound to hydroxylapatite (29%) was spread for electron microscopy from sufficient formamide to extend single strands fully while preserving duplex regions intact (Chamberlin et al., 1975). Based on a hydroxylapatite thermal chromatogram of reassociated chicken repetitive DNA (data not shown), we calculate that the 57% formamide used during spreading provides a criterion 15 °C below the T_m of chicken repetitive duplexes. Thus, even very short interspersed repetitive duplexes should be sufficiently stable, under the conditions used, for visualization in electron micrographs.

Structures in electron micrographs were measured and placed into categories according to the number of ends (Table IV). Only molecules entirely within the frame of the micrographs were included in the analysis. The vast majority of structures were two ended and contained no visible loops or branches. Ten percent of the DNA mass was scored as three-ended structures. Nine percent of the DNA mass was not classified because of tangles or dirt which obscured the number of ends. Nine percent of the DNA was classified as four-ended structures; only 16 four-ended structures were observed among the 569 molecules scored. From this we conclude that DNA in the short-period interspersion pattern is very rare or absent in the chicken genome.

Since most structures seen in the electron micrographs were two ended, we also infer that most fragments of chicken DNA 2200 nucleotides in length which bear repetitive sequences terminate with a repetitive sequence element. This provides an indication that most repetitive sequence elements in chicken DNA are at least 2000 nucleotides in length.

TABLE IV: Structures Present in Electron Micrographs of Reassociated Chicken DNA.

class of structure	no. of molecules	fract. of molecules	av length[a] (ntp)	total length in class (ntp)	fract. of total DNA
2 ended	499	0.877	2251	1.12×10^6	0.72
3 ended	34	0.06	4684	0.16×10^6	0.10
4 ended	16	0.03	8987	0.14×10^6	0.09
uninterpretable[b]	20	0.04	7190	0.14×10^6	0.09

[a] Lengths given are contour lengths calculated by using the magnification factor (75 000×) and the relationship 1μ of DNA = 3000 ntp. Since the contour lengths of double and single strands of the same length differ slightly under these spreading conditions and the majority of structures were molecules in which single and double strands could not be clearly distinguished, these lengths are approximate but reliable within 20%. [b] Structures were classed as uninterpretable if the number of ends could not be clearly distinguished or if a portion of the molecule was obscured by dirt.

TABLE V: Effect of DNA Fragment Length on Hyperchromicity of Chicken DNA Reassociated to C_0t = 20.

DNA fragment length (nt)	native hyperchrom (%)	fract bound to HAP at C_0t = 20	hyperchrom of HAP bound fract (%)	length of fragments bound to HAP (nt)
430	24	0.095	16	520
2200	26	0.200	13	2200

Optical Hyperchromicity of Partially Reassociated Chicken DNA Fragments. When short, repeated sequences are interspersed with longer single-copy sequences, structures formed by reassociation of long DNA strands to repetitive C_0t contain short duplexes flanked by longer, single-stranded regions. This situation can be detected by rapidly decreasing hyperchromicity of the reassociated DNA as the length of the fragments is increased. Table V shows the hyperchromicity of the fraction of chicken DNA bound to hydroxylapatite at C_0t = 20 using fragments about 500 and 2000 nucleotides in length, respectively, compared to the hyperchromicity of native DNA of the same lengths. The short fragments regained 67% of native hyperchromicity at C_0t = 20, and fragments 2000 nucleotides long regained 50% of native hyperchromicity. Thus, we did not observe the rapid loss of hyperchromicity typical of genomes in which the short-period interspersion pattern predominates.

Measurement of the Length of Repetitive DNA Sequences Using S1 Nuclease. To obtain direct information about the lengths of repetitive sequences in chicken DNA, fragments 3300 nucleotides in length were denatured and reassociated to a C_0t = 50. Single strands were digested with S1 nuclease under standard conditions (Britten et al., 1976). Variability of S1 nuclease preparations imposes the requirement for digestion of a well-characterized genome with the same S1 nuclease preparation if valid conclusions about the length distribution of repetitive sequence elements are to be drawn. We used the sea urchin genome for this purpose, since it has already been shown to contain a major fraction of repetitive sequence elements 300 nucleotides in length, as well as some longer repeated sequences (Britten et al., 1976). S1 digested chicken or sea urchin repetitive duplexes were isolated by hydroxylapatite chromatography, and duplex lengths were determined by chromatography on a column of Sepharose CL-2B (Figure 4).

Distinctly different distributions of repetitive sequence lengths were determined for reassociated chicken and sea urchin DNA, respectively. The prominent fraction of repetitive sequences 300 nucleotides in length, typical of sea urchin DNA, was not present in reassociated chicken DNA. Most

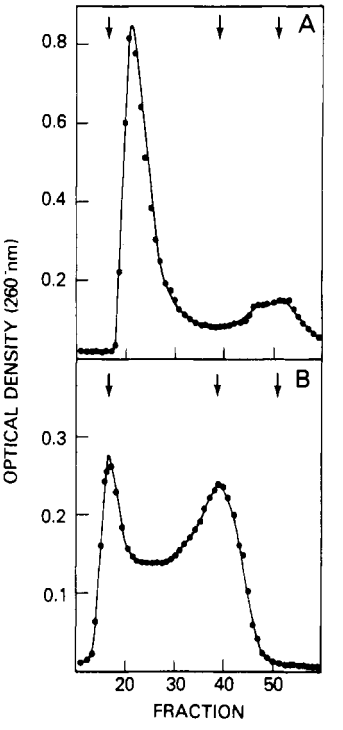

FIGURE 4: Sepharose CL-2B profiles of S1-resistant repetitive duplexes of chicken and sea urchin DNA. (A) Fragments of chicken DNA, 3300 nucleotides in length, were reassociated to C_0t = 50 and digested with S1 nuclease, as described under Materials and Methods. S1-resistant duplexes (17%) were collected on hydroxylapatite in 0.12 M PB at 60 °C, eluted with 0.4 M PB at 60 °C, and chromatographed on a Sepharose CL-2B column (Pharmacia) (3 × 25 cm) in 0.12 M PB at 60 °C. The positions of the void volume, of a 300 ntp sheared DNA marker, and of the included volume are indicated by arrows. (B) Sea urchin DNA fragments 2900 nucleotides in length were reassociated to C_0t = 20, treated with S1 nuclease, using the same DNA and enzyme concentration as described for chicken DNA in (A), collected on hydroxylapatite (14% bound), and chromatographed on the same Sepharose CL-2B column.

chicken repetitive sequences chromatographed near the void volume, indicating a minimum duplex length of 2000 nucleotide pairs. The length of the major fraction of chicken repetitive sequences (fractions 19–26, Figure 4A) was confirmed by electrophoresis in a 1.4% agarose gel, in which the mobility of chicken repetitive DNA indicated a broad range of lengths from 2000 to 4000 nucleotide pairs (data not shown). The small peak of optical density at the position of the included volume of the Sepharose column profile of chicken repetitive DNA does not have the absorbance spectrum of DNA, and represents a contaminant in the DNA preparation of unknown origin. Precise quantitation of the relative amounts of long and short repetitive sequence elements in chicken DNA has not yet

been accomplished. However, it is clear that the distribution of repetitive sequence lengths in chicken DNA indicates a predominance of long repetitive sequences with only a minor fraction of short repeated sequence elements 300 nucleotides in length.

Discussion

Previous reports have presented conflicting interpretations of data relating to DNA sequence organization in the chicken, emphasizing short (DeJiminez et al., 1974) or long (Arthur and Strauss, 1978) period interspersion as the predominant pattern. More complete investigation of chicken DNA has provided a description intermediate between these two extremes.

Our current knowledge of DNA sequence organization in the chicken can be briefly summarized. From the reassociation rate constant of the single-copy fraction of total DNA, we calculate that the haploid genome size is 1.1 pg (Table 1). The majority of the chicken genome is single-copy sequence; only 12–13% of the DNA is repetitive. The repetitive fraction can be divided into two repetition frequency classes with average repetition frequencies of 1500- and 15-fold, respectively. Most of the S1-resistant repetitive duplexes formed at C_0t = 50 are at least 2000 nucleotides in length.

Single-copy and repetitive sequences are interspersed with a dominant single-copy sequence length of about 4500 nucleotides. About 40% of the chicken genome is interspersed in this pattern. A minor component containing interspersed single-copy sequences 4500 to 17 500 nucleotides in length may also be present. More than one third of the chicken genome consists of single copy regions at least 17 500 nucleotides in length. We were unable to detect even a minor fraction of the chicken genome consisting of repetitive sequences 300 nucleotides in length interspersed with single-copy sequences 1000 nucleotides in length.

Some features of this description distinguish chicken DNA from that of other animals studied so far. The most striking and unusual feature is the presence of an interspersed pattern of single-copy and repeated sequences with much longer single-copy sequences than described previously, three to five times longer than commonly found in animal genomes. The data obtained from electron micrographs indicate that most repetitive sequences in the interspersed fraction exceed 2000 nucleotides in length, and this is confirmed by the optical hyperchromicity of long fragments reassociated to C_0t = 20. The S1 nuclease measurement of repetitive sequence lengths also indicates a predominance of long repetitive sequences.

Chicken DNA shares some characteristics with other genomes bearing the long-period interspersion pattern, most notably the unusually small fraction of repetitive DNA and the predominance of long repetitive sequence elements. In addition, the long single-copy regions occupying about one-third of the chicken genome are reminiscent of those identified previously in *Drosophila* DNA (Manning et al., 1975).

It now appears that a variety of different patterns of sequence organization may be revealed as the DNA of more animal species is studied. At present, we are unable to place the DNA sequence organization of the chicken in perspective with respect to the DNA of other avian species, since little published information is available. Preliminary investigations have already shown, however, that the DNA of the ostrich, *Struthio camelus,* shares at least one distinctive feature with chicken DNA: the fraction of the genome that is repetitive is unusually small (Eden et al., submitted for publication). The unique features of genome organization in the chicken provide an indication that further investigations of the structure of avian DNA can be expected to enrich our present understanding of the functional meaning of DNA sequence organization. Because the genome of the chicken contains some elements characteristic of each of the extremes of sequence organization previously described in other eukaryotic genomes, it provides a unique opportunity to relate structure and function. The most interesting question is whether the majority of active structural genes are found in the interspersed portion of the genome or occur in the extremely long single-copy regions of chicken DNA. Studies currently in progress are aimed at answering this question.

Acknowledgment

We thank Dr. Dale Graham for the gift of S1 nuclease, Dr. Richard Kramer for preparation of restriction fragments of DNA used as length standards, Dr. Gilbert Smith for providing the facility for electron microscopy, and Marilyn Meyers for help with tabulation of structures in electron micrographs.

References

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

Arthur, R. R., and Straus, N. A. (1978), *Can. J. Biochem. 56,* 257.

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

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

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

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

Crain, W. R., Davidson, E. H., and Britten, R. J. (1976b), *Chromosoma 59,* 1.

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

Davidson, E. H., Galau, G. A., Angerer, R. C., and Britten, R. J. (1975), *Chromosoma 51,* 253.

Davis, R. W., Simon, M., and Davidson, N. (1971), *Methods Enzymol. 21D,* 413.

Deininger, P. L., and Schmid, C. W. (1976), *J. Mol. Biol. 106,* 773.

DeJiminez, E. S., Gonzalez, J. L., Dominguez, J. L., and Saloma, E. S. (1974), *Eur. J. Biochem. 45,* 25.

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

Firtel, R. A., and Kindle, K. (1975), *Cell 5,* 401.

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

Hudspeth, M. E. S., Timberlakes, W. E., and Goldberg, R. B. (1977), *Proc. Natl. Acad. Sci. U.S.A. 74,* 4332.

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

Noll, H. (1967), *Nature (London) 215,* 350.

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

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

Smith, M. J., and Boal, R. (1978), *Can. J. Biochem.* (in press).

Walbot, V., and Dure III, L. S. (1976), *J. Mol. Biol. 101,* 503.

Wells, R. W., Royer, H., and Hollenberg, C. P. (1976), *Mol. Gen. Genet. 147,* 45.

Zimmerman, J. L., and Goldberg, R. B. (1977), *Chromosoma 59,* 227.