Eur. J. Biochem. 115, 235–239 (1981)

The Major Components of the Mouse and Human Genomes

2. Reassociation Kinetics

Philippe SORIANO, Gabriel MACAYA, and Giorgio BERNARDI

Laboratoire de Génétique Moléculaire, Institut de Recherche en Biologie Moléculaire, Paris

(Received October 27, 1980)

The reassociation kinetics of DNA fragments obtained from the major components of the mouse and human genomes (recently isolated in our laboratory) have been investigated. It has been found that the relative amounts of interspersed repeated and unique sequences strikingly differ in the different major components of each genome and in the corresponding major components of the two genomes. Furthermore, within each major component, the interspersed repeated and unique sequences do not differ in dG + dC contents. These findings lead to the general conclusion that the sequence organization of mammalian genomes is not uniform in different chromosomal regions and that it exhibits remarkable variations in different mammals.

Main-band DNA from warm-blooded vertebrates can be resolved by density gradient centrifugation techniques into three or four families of fragments of different dG + dC contents [1–3]. These major DNA components are similar in their buoyant densities and relative amounts in all species tested and are observed in DNA preparations ranging in M_r from 2×10^6 to over 200×10^6 [2, 3].

The four major components of the mouse and human genomes have been recently prepared in high yields, characterized in their basic properties and shown to be endowed with a very low compositional heterogeneity over a wide molecular weight range [4]. The availability of such components, which all together represent at least 85% and possibly 100% of the main bands of these DNAs, allowed us to investigate their kinetics of reassociation. The results so obtained are presented and discussed in this paper.

MATERIALS AND METHODS

DNA Preparations

Mouse and human DNAs and their major components were preparations obtained in our laboratory as described in the preceding paper [4].

Reassociation Kinetics

Escherichia coli and mouse DNA samples were digested with restriction endonucleases *Hae*III and *Hae*III + *Hpa*II, respectively; the weight-average single-strand length was 200–300 nucleotides, as determined by alkaline band sedimentation using the M_w vs s relationship of Prunell and Bernardi [5]. These M_w values were lower than expected on the basis of the results of gel electrophoresis of *Hae*III-degraded native DNA samples, possibly owing to the introduction of single-stranded breaks during enzymatic digestion. Human DNA samples and some mouse DNA samples were sheared (at 100 μg/ml in 0.05 M potassium phosphate, 60% glycerol) to a weight-average single-stranded length of 300–400 nucleotides in a Virtis homogenizer (Gardiner, NY) as described by Britten et al. [6]. All samples were filtered through Chelex 100 (Bio-Rad) chromatographed on hydroxyapatite, transferred into 1 mM potassium phosphate by gel filtration on Sephadex G-25 (Pharmacia, Uppsala, Sweden) and adjusted to a concentration of 1 mg/ml in 0.2 M potassium phosphate and 0.1 mM EDTA.

In the range of c_0t values (c_0 being the initial DNA concentration, t the incubation time) higher than 1 M · s, DNA samples (2–4 μg) were sealed in 20-μl capillaries, heated for 10 min in boiling water, incubated at a temperature of $t_m - 25$ °C for various times, diluted 50–100-fold with 0.03 M potassium phosphate and simultaneously loaded on a small (0.1-ml volume, 5-mm diameter) hydroxyapatite column equilibrated with the same buffer and thermostated at 50 °C. In order to eliminate the formation of air bubbles, all buffers were saturated with helium before use.

In the low c_0t value range ($c_0t < 1$ M · s) dilute DNA solutions (approximately 2 μg/ml) in 0.05 M potassium phosphate and different KCl concentrations were sealed in glass ampoules, denatured as described above and incubated at $t_m - 25$ °C for various times. DNA solutions were adsorbed batchwise on about 0.1 ml of hydroxyapatite and the suspension was packed in a column thermostatted at 50 °C.

DNA elution from hydroxyapatite was performed with a linear potassium phosphate gradient (0.03–0.4 M), since this was found to provide better separation between denatured and reassociated DNA compared to sodium phosphate. The column effluent was monitored at 260 nm with a Zeiss MR1D microcuve placed in a PMQII spectrophotometer and each recorded chromatographic peak was then integrated using Simpson's rule. The integral absorbance value of the denatured DNA peak was corrected for hyperchromicity. This was found to be in the range of 15–20% at the passage of denatured DNA through the microcuvette. The relative amount of reassociated DNA was estimated as the percentage of material eluting above 0.20 M potassium phosphate.

Enzymes. Restriction endonucleases *Hae*III (EC 3.1.23.17), *Hpa*II (EC 3.1.23.24) and *Alu*I (EC 3.1.23.1).

Nomenclature. The nomenclature of DNA components used is described in a footnote to the preceding paper (4].

Note. All computer programs mentioned in this paper are available upon request.

The reassociation profiles, namely the curves of the fraction of reassociated DNA vs 'equivalent' c_0t (namely the c_0t values corrected to 0.18 M K^+ according to Britten et al. [6]) were resolved into second-order kinetic curves, using the following approach. A certain number of experimental values belonging in the slowest-reassociating class were used to estimate the $c_0t_{1/2}$ (corresponding to half-reassociation) and the relative amount of this kinetic class by a least-square method. The same procedure was applied to the following reassociating classes. The resulting kinetic curves were used to draw the overall profile. This step was reiterated by using different numbers of experimental values for each reassociating class. The solution retained in this trial-and-error procedure was the one fitting the data with a minimum number of kinetic classes and with the smallest total least-square deviation from the experimental values. While the decomposition of the overall reassociation profile into kinetic classes, as just outlined, provided a reliable estimate of the amount of both slowly reassociating and foldback DNA, the relative amounts of DNA reassociating at a fast or at an intermediate rate were those corresponding to the best fit with experimental data. These calculations were done using a Fortran program made for a Control Data 6600 computer.

Other Methods

Determinations of base composition and degradation by S1 nuclease were done as described by Thiery et al. [7] and Britten et al. [8], respectively.

RESULTS

Fig. 1 A shows the reassociation kinetic profiles obtained with total mouse DNA and its major components, and with *Escherichia coli* DNA. The reassociation results obtained with *E. coli* DNA and total mouse DNA were practically identical with those reported by Britten and Kohne [9] and by Cech and Hearst [10], respectively. The kinetic classes present in the light components were similar in amount to those in total mouse DNA, but were enriched in rapidly reassociating material and did not contain any very rapidly reassociating DNA; the relative amount of the latter and its absence in the major components indicate that it corresponds to satellite DNA. In contrast, the two heavy components were characterized by a very low level of foldback, fast and intermediate sequences.

In the case of human DNA (Fig. 1 B) the two light components showed slightly lower proportions of foldback and repeated (fast and intermediate) sequences, and higher proportions of slowly reassociating sequences compared to total DNA. The latter showed a reassociation curve very similar to that reported by Schmid and Deininger [11]. In contrast, the two heavy components showed considerably more foldback and repeated sequences relative to total DNA or the two light components.

Table 1 shows the relative amounts and the $c_0t_{1/2}$ values of the different kinetic classes in total DNA and its major components. The $c_0t_{1/2}$ values of the slowly reassociating sequences of the latter showed the decrease expected on the basis on their enrichment as obtained upon fractionation and were used to provide an independent estimate of the percentages of different DNA components; this was in good agreement with that derived from their yields (compare the results of Table 1 with those of Table 1 in the preceding paper

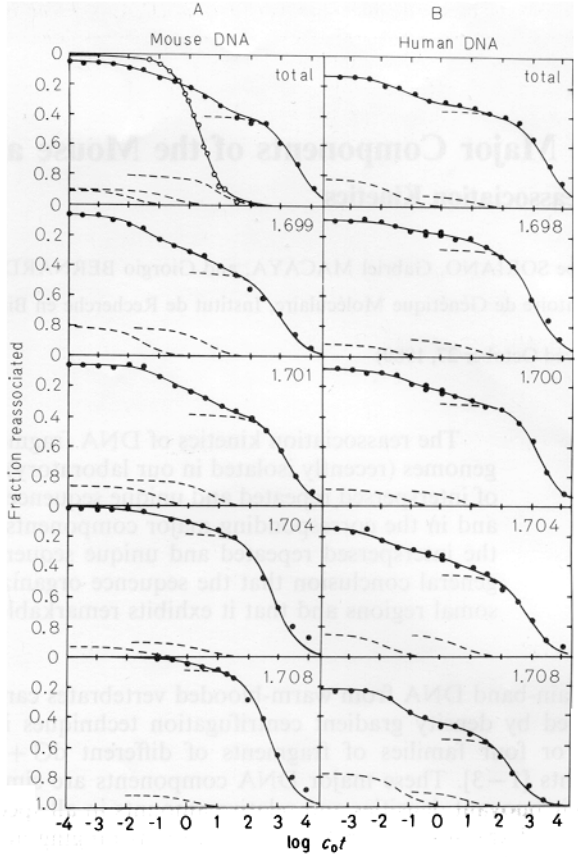

Fig. 1. *Reassociation kinetics of mouse and human DNAs and their major components.* Results obtained with *E. coli* DNA (O) are shown for the sake of comparison. The solid lines through the experimental points are the overall profiles resulting from the analysis of kinetic classes (----) described in Materials and Methods. c_0t values are given in mol nucleotide $\cdot l^{-1} \cdot s$

[4]). Incidentally, determining the amount of interspersed repeated sequences on isolated major components avoids mistaking satellite DNA for interspersed repeated sequences, an error which was made [12] in the case of the total bovine genome (see [13]).

The reassociation profiles of mouse DNA components, as obtained on sheared DNA (data not shown) were similar to those obtained on the sample degraded by *Hae*III + *Hpa*II and presented in Fig. 1 A. Likewise, the 1.708-g/cm^3 component degraded by *Alu*I, an enzyme having a restriction site, d (A-G$\downarrow$C-T), completely different from those of *Hae*III, d (G-G$\downarrow$C-C), and of *Hpa*II, d (C$\downarrow$C-G-G), to chain lengths close to those obtained with the latter enzymes showed the same reassociation profile as that presented in Fig. 1 A. These results indicate that using a restriction enzyme like *Hae*III to degrade mouse DNA and its components does not introduce any bias in the reassociation kinetics.

A characteristic feature of the hydroxyapatite chromatography of mouse DNA and its components is that the boundary between denatured and reassociated DNA is less sharp than for *E. coli* DNA; this is particularly true at c_0t values lower than those causing the reassociation of slowly reassociating DNA, and more so for the heavy than for the light components. These effects are enhanced in Fig. 1 by the fact that transmission and not absorbance of eluent was recorded. Separation between denatured and reassociated DNA was

Table 1. *Kinetic classes in total mouse and human DNAs and their major components*

DNA	Amount		Class	Amount of class		$c_0t_{1/2}$	
	mouse	human		mouse	human	mouse	human
	%			%		M · s	
Total	100	100	foldback	5	12		
			very fast[a]	8	–	0.02	
			fast	9	18	0.3	0.03
			intermediate	19	8	9	8
			slow	59	62	3000	2400
1.699 or 1.698 g/cm³	26	31	foldback	7	8	–	–
			fast	19	9	0.03	0.03
			intermediate	18	12	4	11
			slow	56	71	950	1350
1.701 or 1.700 g/cm³	37	37	foldback	6	8	–	
			fast	14	14	0.10	0.05
			intermediate	18	10	5	4
			slow	62	68	1600	1600
1.704 g/cm³	15	18	foldback	1	15	–	–
			fast	7	16	0.05	0.01
			intermediate	10	15	3.5	4
			slow	82	54	700	800
1.708 g/cm³	10	7	foldback	–	22	–	–
			fast	–	22	–	0.04
			intermediate	9	10	1.7	7
			slow	91	46	600	350

[a] This kinetic class appears to correspond to satellite DNA (see Results).

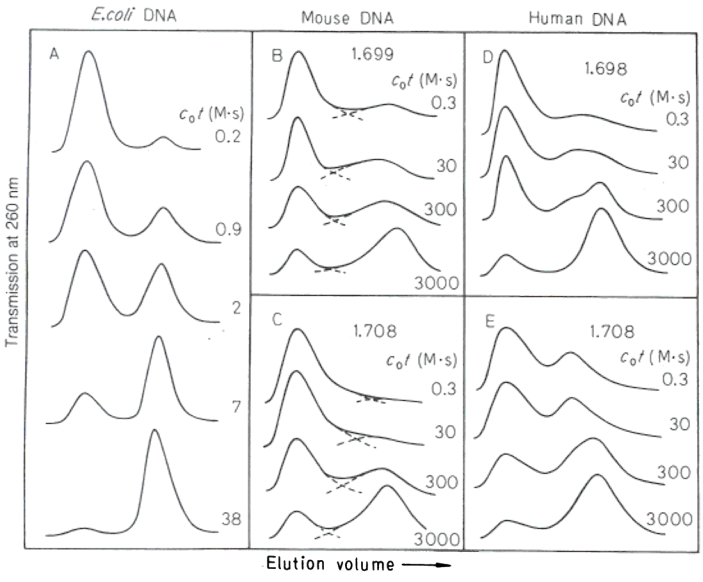

Fig. 2. *Gradient elution profiles from hydroxyapatite of* E. coli *DNA (A), mouse DNA 1.699-g/cm³ (B) and 1.708-g/cm³ (C), human DNA 1.698-g/cm³ (D) and 1.708-g/cm³ (E) components, reassociated to the different* c_0t *values indicated*

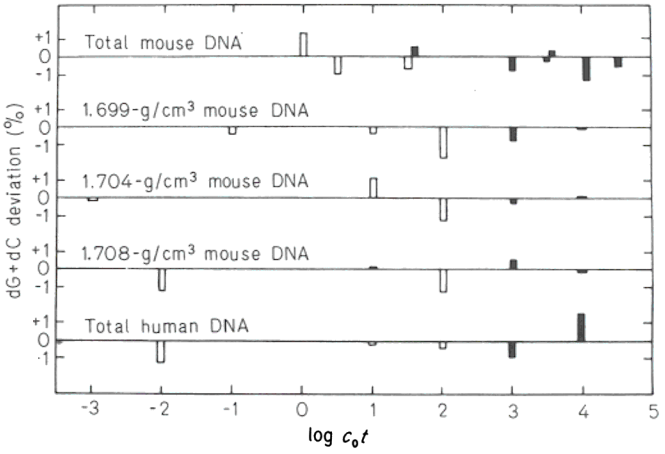

Fig. 3. *dG + dC contents of DNA fractions denatured (white bars) or reassociated (black bars) at different* c_0t *values.* Data are presented as difference histograms with the dG + dC contents of total mouse DNA, its major components, and total human DNA. Error is ± 1%

also not complete in the case of human DNA (Fig. 2). The chromatograms were, however, different from those obtained with mouse DNA and showed trimodal elution patterns in total DNA as well as in its four major components. In all cases, the intermediate peak disappeared at very high c_0t values. The results with human DNA are reminiscent of those of Hoyer and van de Velde [14].

The base composition of denatured and reassociated fractions as obtained from total mouse DNA and from its major components was found to be the same within experimental error (Fig. 3). DNA fractions from total DNA and from the 1.701-g/cm³ component, reassociated at a c_0t value corresponding to repeated sequences, were also analyzed after S1 nuclease treatment and removal of mononucleotides and oligonucleotides; their dG + dC content was found to be 41% and 40.1% respectively, in very close agreement with the contents (40.3% and 40.2%) determined on the native DNAs. Similar experiments showed that the dG + dC content of denatured and reassociated total human DNA, as obtained at different c_0t values, was the same within 1%.

DISCUSSION

The results obtained in the present work fall into three different categories.

The first set of data shows that the base composition of DNA fragments denatured or reassociated at different $c_0t_{1/2}$ values, as obtained from total mouse or human DNA and from mouse DNA components, did not show any significant difference (Fig. 3), contradicting claims that interspersed repeated sequences of human DNA are enriched in dG+dC [14]. These were, however, based on inferences from results only indirectly bearing on the issue under consideration rather than on actual analysis of DNA fragments. Such claims were also used to question the very existence of discrete density components in the mammalian genome [15,16]. Since the argument is crucially dependent upon an assumption disproved by the data of Fig. 3, it is not worthwhile discussing it in detail.

The second main point established by the reassociation kinetics data is that particular sub-sets of slowly reassociating DNA fragments are enriched in individual major components, as indicated by the fact that the $c_0t_{1/2}$ value of the slowly reassociating class is lower in individual components compared to total DNA. On the one hand, this stresses the fact that a real separation of different unique sequences has been achieved by density gradient fractionation. On the other it allows one to obtain estimates of the relative amount of each major component (Table 1). These are in reasonable agreement with those obtained from the yields [4].

A comparison of the $c_0t_{1/2}$ values as obtained for the unique sequences of isolated major components with those obtained on total DNAs indicates that the sum of the former is higher than the latter (Table 1). Even if this comparison is of a semiquantitative nature, there is definitely an apparent discrepancy between the two sets of data in the case of both genomes. This may be due to two reasons which are not mutually exclusive. First, the major components, as prepared are not 100% pure and therefore display anomalously high $c_0t_{1/2}$ values owing to the fact that they contain unique sequences belonging to neighboring major components; this may be particularly true for the two light components which are very close in buoyant density and the most abundant in the genomes; such an explanation has been shown to be correct by experiments [17] showing that while the hybridization of a mouse β-globin cDNA probe mainly take place on the 1.701-g/cm^3 component, some hybridization also occurs on the 1.699-g/cm^3 component. Second, low-copy families of repeated sequences may distribute themselves in different components. Under the low stringency ($t_m - 25$ °C) conditions used in the present work, such slowly reassociating sequences as are present in individual components would only reassociate with themselves, whereas in total DNA they will cross-hybridize with similar sequences present in different components; this explanation has also been shown to be correct (F. Rakotobe, J. Doly and G. Bernardi, unpublished results).

The two main results obtained in the present work are that the relative amounts of interspersed and unique sequences differ (a) in the different major components of the mouse and the human genomes and (b) in the corresponding components of the two genomes, these differences being particularly striking for the heavy components. The first result leads to the conclusion that sequence organization, in terms of interspersed repeated and unique sequences, is not uniform all over the mammalian genomes. Recalling that the DNA fragments making up the major components of these genomes derive from very large chromosomal segments [4], this result indicates a regional variation in sequence organization along mammalian chromosomes. Even if our results only concern the relative amounts of the two types of sequences, it is evident that they also imply a different kind of interspersion. In the case of the mouse genome, for instance, the relative amount of interspersed repeated sequences is so low in both heavy components that is possible that these components have a long-period interspersion pattern of the *Drosophila* type [18]; in contrast, the light components might well have a short-period interspersion pattern of the *Xenopus* type [19]. Current investigations will provide information on the interspersion patterns of different major components.

The second result leads to another conclusion, namely that sequence organization, again in terms of interspersed repeated and unique sequences, is different in corresponding major components. (The large amounts of foldback sequences in human DNA raise the possibility that the relative amounts of repeated and unique sequences may be underestimated or overestimated if they are overrepresented or underrepresented, respectively, near the inverted repeats of foldback DNA. However, even in the extreme case of only one kind of sequence being present in foldback DNA, the difference between the repeated and unique sequences in the heavy components of mouse and human DNA remains large.) Recalling that the corresponding major components of mammalian genomes show little change in buoyant densities and in relative amounts, and that they appear to contain the same genes (unpublished results), it is rather unexpected that they differ so much in the relative amounts of interspersed repeated and unique sequences. Again, these findings should be supplemented by information on the interspersion patterns. The available results already stress, however, the complexity of genome organization in mammalian genomes and suggest that the similarity of relative amounts of interspersed repeated sequences and the widespread distribution of *Xenopus* interspersion patterns [20] hide large differences in the sequence organization of eukaryotic genomes.

The results of the gradient elution of reassociating DNA from different major components of mouse and human genomes not only stresses once more the large differences in the relative amounts of kinetic classes of the 1.708-g/cm^3 (Fig. 2C and E) but also indicates differences in components like the lightest ones, where the relative amounts of kinetic classes do not differ dramatically (Fig. 2B and D; see also Table 1). Further studies are needed to define such differences precisely; it is very probable, however, that they reflect differences in the secondary structure of the reassociated interspersed repeated sequences; if this is so, the results would suggest a higher degree of homology of those sequences in mouse compared to human DNA.

We wish to thank Dr Hugues Blanc for writing the computer program, Ms Martine Brient for typing this manuscript, Mr Philippe Breton for the drawings.

REFERENCES

1. Filipski, J., Thiéry, J. P. & Bernardi, G. (1973) *J. Mol. Biol. 80*, 177–197.
2. Thiéry, J. P., Macaya, G. & Bernardi, G. (1976) *J. Mol. Biol. 108*, 219–235.
3. Macaya, G., Thiéry, J. P. & Bernardi, G. (1976) *J. Mol. Biol. 108*, 237–254.
4. Cuny, G., Soriano, P., Macaya, G. & Bernardi, G. (1981) *Eur. J. Biochem. 115*, 227–233.

5. Prunell, A. & Bernardi, G. (1973) *J. Biol. Chem. 248*, 3433–3440.
6. Britten, R. J., Graham, D. E. & Neufeld, B. R. (1974) *Methods Enzymol. 29E*, 363–418.
7. Thiéry, J. P., Ehrlich, S. D., Devillers-Thiéry, A. & Bernardi, G. (1973) *Eur. J. Biochem. 38*, 434–442.
8. Britten, R. J., Graham, D. E., Eden, F. C., Painchaud, D. M. & Davidson, E. H. (1976) *J. Mol. Evol. 9*, 1–23.
9. Britten, R. J. & Kohne, D. E. (1968) *Science (Wash. DC) 161*, 529–540.
10. Cech, T. R. & Hearst, J. E. (1976) *J. Mol. Biol. 100*, 227–256.
11. Schmid, C. W. & Deininger, P. L. (1975) *Cell, 6*, 345–358.
12. Davidson, E. H., Galau, G. A., Angerer, R. C. & Britten, R. J. (1975) *Chromosoma (Berl.) 51*, 253–259.
13. Macaya, G., Cortadas, J. & Bernardi, G. (1978) *Eur. J. Biochem. 84*, 179–188.
14. Rinehart, F. P., Wilson, V. L. & Schmid, C. W. (1977) *Biochim. Biophys. Acta, 474*, 69–81.
15. Vizard, D. L., Rinehart, F. P., Rubin, C. M. & Schmid, C. W. (1977) *Nucleic Acids Res. 4*, 3753–3768.
16. Rinehart, F. P., Deininger, P. L. & Schmid, C. W. (1978) *Biochim. Biophys. Acta, 520*, 21–37.
17. Cuny, G., Meunier-Rotival, M., Soriano, P. & Bernardi, G. (1978) in *Genetic Engineering* (Boyer, H. W. & Nicosia, S., eds) pp. 109–115, Elsevier/North-Holland, Amsterdam.
18. Manning, J. E., Schmid, C. W. & Davidson, N. (1975) *Cell, 4*, 141–155.
19. Davidson, E. H., Hough, B. R., Amenson, C. S. & Britten, R. J. (1973) *J. Mol. Biol. 77*, 1–23.
20. Davidson, E. H., Galau, G. A., Angerer, R. C. & Britten, R. J. (1975) *Chromosoma (Berl.) 51*, 253–259.

P. Soriano and G. Bernardi, Laboratoire de Génétique Moléculaire,
Institut de Recherche en Biologie Moléculaire du Centre National de la Recherche Scientifique, Université de Paris VII
Tour 43, 2 Place Jussieu, F-75221 Paris-Cedex-05, France

G. Macaya, Universidad de Costa Rica, Ciudad Universitaria Rodrigo Facio,
San José, Costa Rica