Proc. Natl. Acad. Sci. USA
Vol. 80, pp. 1816–1820, April 1983
Biochemistry

The distribution of interspersed repeats is nonuniform and conserved in the mouse and human genomes

(repeated sequences/mammalian genome/major components)

PHILIPPE SORIANO*, MICHÈLE MEUNIER-ROTIVAL, AND GIORGIO BERNARDI†

Laboratoire de Génétique Moléculaire, Institut Jacques Monod, 2 Place Jussieu, 75005 Paris, France

Communicated by Maxine F. Singer, December 21, 1982

ABSTRACT **We investigated the genomic distribution of mouse and human repeated sequences by assessing their relative amounts in the four major components into which these genomes can be resolved by density gradient centrifugation techniques. These components are families of fragments that account for most or all of main-band DNAs, range in dG+dC content from 37% to 49%, and are derived by preparative breakage from long DNA segments (>300 kb) of fairly homogeneous composition, the isochores. The results indicate that the short repeats of the B1 family of mouse and of the *Alu* I family of man are most frequent in the heavy components, whereas the long repeats of the *Bam*HI family of mouse and of the *Kpn* I family of man are mainly present in the two light components. These results show that the genomic distribution of repeated sequences is nonuniform and conserved in two mammalian species. In addition, we observed that the base composition of two classes of repeats (60% dG+dC for short repeats; 39% dG+dC for long repeats) is correlated with the composition of the major components in which they are embedded. Finally, we obtained evidence that not only the short repeats but also the long repeats are transcribed, these transcripts having been found in mouse poly(A)$^+$ mRNA.**

Studies of reassociation kinetics for a variety of eukaryotic DNAs have indicated two types of organization of interspersed repeated sequences—namely, the extremely widespread *Xenopus* pattern and the rare *Drosophila* pattern (see ref. 1 for a recent brief review). These studies have been performed on unfractionated DNA preparations, and it is generally assumed that the observed patterns extend throughout the genome. This question can be examined experimentally in the case of DNAs from warm-blooded vertebrates. In fact, we have demonstrated that these DNAs can be fractionated into a characteristic set of three or four major components by preparative density gradient techniques (2–4). In the case of mouse, the components have buoyant densities of 1.699, 1.701, 1.704, and 1.708 g/cm^3 and represent 22, 34, 26, and 7% of DNA, respectively; those of human DNAs have densities of 1.698, 1.700, 1.704, and 1.708 g/cm^3 and relative amounts of 29, 33, 22, and 9%, respectively. These major components account for all or nearly all of main-band DNA (excluding satellites and minor components) and are derived by preparative breakage from three or four corresponding families of long DNA segments of at least 300 kilobases (kb), the isochores, which are fairly homogeneous in base composition and may be responsible for chromosomal banding patterns (5).

Recently, we have examined the reassociation kinetics of total DNA and of the four major components of mouse and human DNAs. The results showed very significant differences and a high species specificity in the relative amounts of the various kinetic classes—particularly of the interspersed repeated sequences—among the four components (6). Nonuniformity and species specificity in the distribution of repeated sequences also have been shown in the four major components of the chicken genome (7).

The recent advances in our knowledge of repeated sequences and the development of specific probes for them (see refs. 8 and 9 for reviews) provide another approach to the study of the distribution of these sequences. Using probes for the *Bam*HI family of long interspersed repeats in mouse, we have found this family to be localized almost entirely in the two light-density components of the mouse genome (10). We now report the distribution of additional families of interspersed repeated sequences in the mouse and human genomes. The results show that the relative amounts of short repeats (the B1 family of mouse and the *Alu* I family of man) are highest in the heavy components, whereas long repeats (the *Bam*HI family of mouse and the *Kpn* I family of man) are mainly present in the light components. Short repeats and long repeats also differ in their dG+dC contents (60% in the first case and 39% in the second) and show a general correlation with the composition of the components in which they occur. Finally, we have obtained evidence that the long repeats from the mouse and human genomes are transcribed.

MATERIALS AND METHODS

mRNA Preparation. Livers from fasted 4- to 5-mo-old BALB/c mice were washed in 50 mM Tris, pH 7.4/25 mM KCl/3 mM $MgCl_2$/1 mM dithiothreitol/250 mM sucrose, ground in a tissue press, and homogenized in 2 vol of buffer with a Dounce homogenizer (Kontes). The homogenate was centrifuged for 10 min at 700 × *g*, and the nuclear pellet was washed twice. The supernatants were pooled and centrifuged for 10 min at 17,000 × *g*, and the resulting postmitochondrial supernatant was treated with Triton X-100 and sodium deoxycholate (final concentrations, 1% and 0.5%, respectively) in the presence of 250 mM KCl. The supernatant was layered on discontinuous sucrose gradients as described (11); after centrifugation, polyribosomal pellets were resuspended in a small volume of 50 mM Tris, pH 7.4/0.3 M NaCl/5 mM $MgCl_2$. An aliquot was analyzed on linear 10–40% sucrose gradients in the same buffer and centrifuged for 40 min in a Beckman SW-41 rotor at 4°C and 39,000 rpm to check the polyribosome size distribution; a maximum absorbance at about 8–10 ribosomes was routinely obtained. Polyribosomal mRNA was prepared by extraction with phenol/chloroform/isoamylalcohol, 50:48:2 (vol/vol), followed by

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviations: bp, base pairs; kb, kilobases.

* Present address: Centre de Biophysique Moléculaire, 1A Avenue de la Recherche Scientifique, 45045 Orléans, France.
† To whom reprint requests should be addressed.

chromatography on oligo(dT)-cellulose (type T3; Collaborative Research, Waltham, MA).

Total (nuclear and cytoplasmic) RNA from HeLa cells was prepared by extraction as above, followed by precipitation at 4°C overnight with 2 M LiCl in the presence of 4 M urea. Poly(A)$^+$ RNA was isolated by oligo(dT)-cellulose chromatography.

Recombinant DNA. cDNA recombinant plasmids were prepared (12) and colony hybridization was performed (13) as described. Blur 8, a plasmid containing a human *Alu* I sequence (14); pCaα6-5, a plasmid containing a 2.8-kb *Kpn* I interspersed repeated sequence from African green monkey (15); and Mm31 and Mm35, two plasmids containing B1 repeated sequences (16), were kindly supplied by Warren Jelinek, Maxine Singer, and Georgi Georgiev, respectively. pH-2^d-3, a plasmid containing a cDNA coding for a histocompatibility antigen (17), was provided by François Brégégère.

DNA Components. The major components of mouse DNA were prepared in our laboratory by Jordi Cortadas, Gerard Cuny, and Alain Huyard, essentially as described (5) except that density gradient centrifugation was done in Cs_2SO_4/bis(acetatomercurimethyl)dioxane (18) instead of Cs_2SO_4/Ag^+. The sedimentation coefficient of the starting DNA preparation was 50 S, as determined by band sedimentation in Vinograd 30-mm Kel-F single-sector cells with a Spinco model E analytical ultracentrifuge equipped with a monochromator, a photoelectric scanner, and a multiplexer (19). By using the S vs. M_r relationship of Eigner and Doty (20), the S value corresponds to a molecular weight of 66×10^6 or 100 kb. The major components of human DNA were preparations already described (5). In this case, the S value of the starting DNA preparation was 32 S, corresponding to a molecular weight of 22×10^6 or 33 kb. Finally, the chicken DNA preparation used to prepare the major components in which the CR1 repeats were localized (unpublished data) had a sedimentation value of 47 S, corresponding to a molecular weight of 57×10^6 or 86 kb (21). In all cases, the shape of the sedimentation band was fairly symmetrical, indicating no anomalous size distribution in the DNA sample. After preparative density gradient centrifugation, the isolated DNA components showed S values that were equal or only slightly (about 10%) lower than that of the starting DNA.

DNA and RNA Blot Hybridization. Restriction endonucleases were purchased from New England BioLabs or prepared by L. Mallet and A. Meier in our institute. Restriction digests, DNA transfer to nitrocellulose filters, and hybridization conditions were as described (22, 23); final washings of the filters were done at 65°C in 0.2 M NaCl/10 mM sodium phosphate, pH 6.8/1 mM EDTA.

Polyribosomal poly(A)$^+$ mRNA was denatured with glyoxal and fractionated on vertical 1.2% agarose gels as described (24). RNAs were blotted onto nitrocellulose filters as described (25). Prehybridization was performed at 42°C in 50% formamide/1 M NaCl/50 mM sodium phosphate, pH 6.8/5 mM EDTA containing sheared denatured salmon sperm DNA at 250 μg/ml, 0.1% $NaDodSO_4$, and 0.2% each of Ficoll, polyvinylpyrrolidone-40, and bovine serum albumin. Hybridization was carried out in the same buffer containing ^{32}P-labeled nick-translated probe ($1–5 \times 10^8$ cpm/μg; $1–5 \times 10^5$ cpm/ml). Filters were washed at 42°C in 50% formamide/1 M NaCl/50 mM sodium phosphate, pH 6.8/5 mM EDTA/0.1% $NaDodSO_4$, then in 0.4 M NaCl/20 mM sodium phosphate, pH 6.8/2 mM EDTA/0.1% $NaDodSO_4$, and lastly in 20 mM NaCl/1 mM sodium phosphate, pH 6.8/0.1 M EDTA/0.1% $NaDodSO_4$. Autoradiography was performed on preflashed Fuji RX Films at −70°C with intensifying screens.

RESULTS

Characterization and Hybridization Patterns of Mouse cDNA Plasmids Containing Repeated Sequences. A library of cDNA clones was constructed from mouse liver polyribosomal poly(A)$^+$mRNA and screened for the presence of repeated sequences by colony hybridization with ^{32}P-labeled mouse DNA under conditions not allowing hybridization of single-copy sequences; 37 plasmids out of the 5,000 screened showed a positive hybridization indicative of repeated sequences. These plasmids were shown to belong in two distinct groups based on their hybridization patterns on *Eco*RI and *Bam*HI digests of DNA and on mRNA. The first group, formed by nine plasmids, could be further subdivided into groups of six and three plasmids; the second group, comprising 28 plasmids, was subdivided into groups of 12 and 16 plasmids. Four plasmids (pMR1–4) that contained the longest inserts and were typical of each group were subjected to further analysis.

Fig. 1 shows the hybridizations of the four ^{32}P-labeled plasmids on nitrocellulose filters of restriction digests of total mouse DNA and its major components. pMR1 gave a prominent band, 1.3 kb, accompanied by lighter bands and a strong smear on *Eco*RI digests of total DNA and of the two light components; essentially no hybridization was seen on the nitrocellulose filters of two heavy components. Two bands, 4 kb and 0.5 kb, and a smear on high molecular weight fragments were seen on filters of *Bam*HI digests of total DNA. pMR2 produced a strong band of 2.9 kb, some lighter ones, and a smear in the high molecular weight range on filters of *Eco*RI digests of total DNA and on filters of the two light components. As in the previous case, practically no hybridization was detected on filters of the two heavy components. A strong 4-kb band, several faint bands, and a high molecular weight smear were seen on filters of *Bam*HI digests of total DNA. When *Eco*RI digests were examined, pMR3 produced a hybridization smear that was intense on filters of the heavy components and relatively weak bands, 6 kb and 2.5 kb, which were best visible on filters of the 1.701 and 1.704 components. Only a smear could be seen when pMR3 was hybridized on filters of *Bam*HI digests of total DNA. Similar results were obtained with pMR4, except that *Eco*RI bands were 6 kb and 3.5 kb in length in this case.

Restriction maps of the four plasmids are presented in Fig. 2*A*. Fragments containing repeated sequences were identified by their positive hybridization with ^{32}P-labeled total DNA, a procedure leading to the establishment of the upper limits of repeated sequence sizes. These were found to be equal to 370,

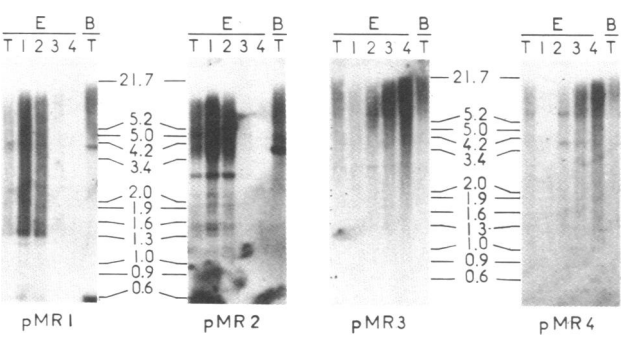

FIG. 1. Autoradiograms of mouse cDNA clones containing repeated sequences (pMR1–4) hybridized on nitrocellulose filters of *Eco*RI (E) and *Bam*HI (B) digests from unfractionated mouse DNA (lanes T) and its major components, 1.699 (lanes 1), 1.701 (lanes 2), 1.704 (lanes 3), and 1.708 (lanes 4). All DNA samples were loaded in equal amounts (1 μg) on 0.6% agarose gels and transferred, after electrophoresis, onto nitrocellulose. Size markers were *Eco*RI/*Hin*dIII fragments of phage λ DNA; sizes are shown in kb.

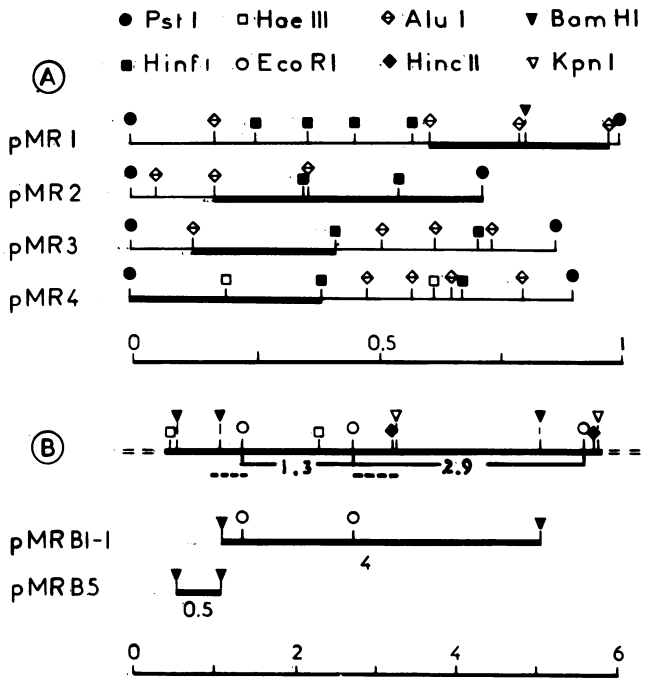

FIG. 2. Restriction maps. (*A*) Cloned mouse cDNAs containing repeated sequences (heavy lines). (*B*) Mouse genomic 5.6-kb *Bam*HI repeated sequences (10). ---, Regions homologous to the repeated sequences contained in pMR1 (left-hand segment) and pMR2 (right-hand segment). Simplified maps of the 4-kb and 0.5-kb *Bam*HI repeats as found in plasmids pMRB1.1 and pMRB5, respectively (10), are also shown. Size scales are shown in kb.

550, 290, and 390 base pairs (bp) for pMR1, 2, 3, and 4, respectively. The sizes of cloned inserts were larger, ≈1,000, 720, 870, and 910 bp, respectively, indicating that the starting poly(A)$^+$ polysomal mRNAs contained unique sequences in addition to repeated sequences.

Cross-hybridization experiments between the four plasmids were negative under conditions of relatively high stringency (0.218 M Na^+ and 65°C) (not shown). pMR1 and pMR2 hybridized with pMRB1.1, a plasmid containing a member of the 4-kb *Bam*HI family (10), but not with Mm31 and Mm35, two plasmids containing B1 sequences (16). pMR3 and pMR4 hybridized with Mm31 and Mm35, respectively, but not with pMRB1.1. A more detailed analysis, by hybridization on restriction fragments from plasmids pMRB1.1 and pMRB5 (10) revealed that pMR1 hybridized with both the 0.5-kb and the 4-kb *Bam*HI fragments but showed only a weak hybridization with the 1.3-kb *Eco*RI fragment. This indicated that the sequence homology included part of the right-hand side of the 0.5-kb *Bam*HI fragment and the left-hand side of the adjacent 4-kb *Bam*HI fragment up to the *Eco*RI site (Fig. 2*B*, left broken line). Other clones of the pMR1 group gave no hybridization with the 0.5-kb *Bam*HI fragment and varying intensities with the 1.3-kb *Eco*RI fragment. In contrast, pMR2 hybridized with the 0.54-kb *Eco*RI/*Kpn* I fragment of pMRB1.1, showing that the sequence homology is located on another part of the 4-kb *Bam*HI fragment (Fig. 2*B*, right broken line).

Hybridization patterns of the B1 probes Mm31 and Mm35 on nitrocellulose filters of total DNA and its major components were identical with those obtained with pMR3 and pMR4, respectively. In addition, pMR4 hybridized on filters of the repeated sequence carried by pH-2^d-3, a plasmid containing a cDNA coding for a histocompatibility antigen (17); this repeated sequence occurs within the 525-bp 3′ untranslated region of the mRNA (26, 27).

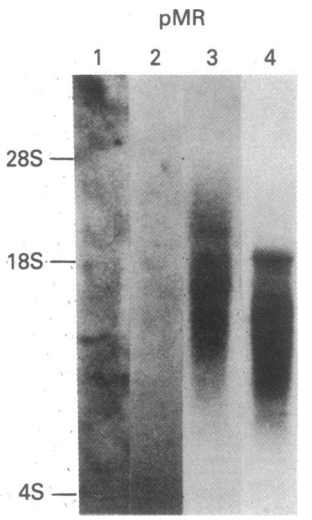

FIG. 3. Autoradiograms of mouse cDNA clones containing repeated sequences (pMR1–4) hybridized on polyribosomal poly(A)$^+$ mRNA (20 μg) fractionated on a 1.2% agarose gel after denaturation by glyoxal. Size markers were 28S and 18S ribosomal RNAs and 4S tRNA from mouse.

Hybridization of the four plasmids with mRNAs was very weak in the case of pMR1 and pMR2 and strong in the case of pMR3 and pMR4; the RNA species hybridized by pMR3 and pMR4 were different (Fig. 3).

Localization of Human Repeated Sequences. pCaα6-5, a plasmid containing a 2.8-kb *Kpn* I interspersed repeated sequence from African green monkey (which also hybridizes to the human *Kpn* I fragments visible after ethidium bromide staining), and Blur 8, a plasmid containing an *Alu* I repeat, were hybridized with total human DNA digested with *Kpn* I or *Hin*dIII. pCaα6-5 strongly hybridized to the bands visible after ethidium bromide staining and, in addition, gave a strong hybridization smear. Both the bands and the smear were strongest on the light components (Fig. 4). In contrast, Blur 8 gave a hybridization smear which was more intense on nitrocellulose filters of the heavy components than on filters of the light ones.

In cross-hybridization experiments between species, it was found that neither pCaα6-5 nor Blur 8 hybridized with mouse DNA under the condition of stringency (t_m − 20°C) used. In contrast, the mouse B1 probes hybridized very weakly with human DNA, in particular on nitrocellulose filters of the heavy

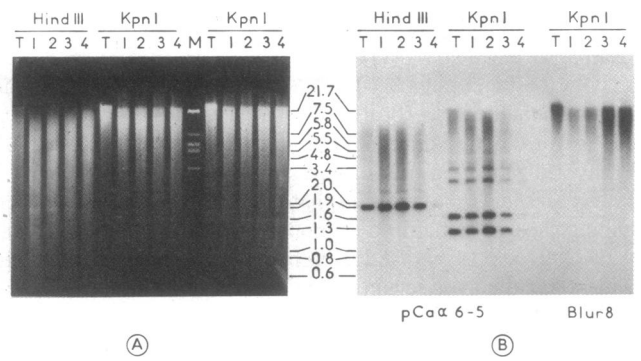

FIG. 4. (*A*) Gel electrophoresis on 0.6% agarose gel of *Hin*dIII and *Kpn* I digests of unfractionated human DNA (lanes T) and its major components 1.698 (lanes 1), 1.700 (lanes 2), 1.704 (lanes 3), and 1.708 (lanes 4) loaded in equal amounts (1 μg). Size markers (lane M) were *Eco*RI and *Eco*RI/*Hin*dIII fragments of phage λ DNA; sizes are shown in kb. (*B*) Autoradiograms of cloned, repeated sequences pCaα6-5 and Blur 8 hybridized on nitrocellulose transfers from the gel shown in *A*.

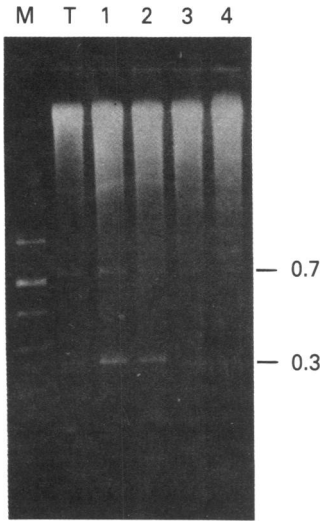

FIG. 5. Gel electrophoresis on 0.6% agarose gel of *Eco*RI digests of unfractionated human DNA (lane T) and its major components (lanes 1–4) loaded in equal amounts (2 μg). Size markers (lane M) were *Alu* I fragments of plasmid pBR322; sizes are shown in kb.

components, where a smear could be detected only after an extended period of autoradiography. This finding is consistent with the moderate homology between B1 and *Alu* I sequences. The absence of hybridization by the human *Alu* I probe with mouse DNA under our conditions is likely to reflect the lower concentration of B1 sequences in mouse compared to *Alu* I sequences in human DNA [these have been estimated as 4×10^5 sequences per haploid genome (28) versus 1×10^5 for the B1 sequences (29)].

Hybridization of pCaα6-5 with HeLa cell total RNAs and poly(A)$^+$ RNA yielded only a weak smear (not shown). Finally, the 680- and 340-bp *Eco*RI fragments corresponding to the α repeats (30) were found to be present in the three lightest components (Fig. 5) and absent from the heaviest one.

Base Composition of the Repeated Sequences and of the Major Components Embedding Them. Table 1 summarizes data for dG+dC content of five repeated sequences and the corresponding content for the DNA component(s) in which they are localized. As shown, the dG+dC content of the short repeats of mouse, man, and chicken is high, 60% and 55%, whereas that of the two long repeats *Bam*HI of mouse and *Kpn* I of monkey is low, 39%. Correspondingly, the major components in which these repeats are localized are high and low in dG+dC, respectively. It should be noted that the size and the size distributions of the isolated major components do not allow the base composition of the repeats to influence, to a significant extent, the fractionation of the DNA fragments carrying them.

Table 1. dG+dC contents of repeated sequences and of the major components containing them*

Repeated sequence	dG+dC, %	Ref.	Major component(s)†	dG+dC†, %
B1 (mouse)	60	31	1.704; 1.708	43; 49
Alu I (human)	60	31	1.704; 1.708	43; 49
*Bam*HI (mouse)	39	10	1.699; 1.701	38; 40
Kpn I (monkey)	39		1.698; 1.700; 1.704	37; 38.5; 43
CR1 (chicken)	48	32	1.708	47

* Base compositions of repeated sequences are from the references indicated, except for the *Kpn* I sequence from pCaα6-5, which was determined in the present work as described (5). In the case of B1 and *Alu* I sequences, the contribution of dA+dT-rich tails has not been taken into account. Base compositions of major components are from refs. 5 and 7. The localizations of repeats are from this work [confirming previous data (10) in the case of *Bam*HI repeats] and from ref. 21 for CR1 repeats.

† Underlined values indicate components (and their dG+dC contents) mainly embedding the repeats.

DISCUSSION

As shown by the results of Fig. 1, each of the four families of interspersed repeats in mouse cDNA clones is characterized by a specific pattern of hybridization on nitrocellulose filters of restriction fragments of mouse DNA and by a specific distribution among the four major components. As far as pMR1 and pMR2 are concerned, it is evident that they correspond to two subfamilies of the *Bam*HI family (10). Their lack of cross-hybridization is accounted for by the fact that different segments of the *Bam*HI repeats are represented in the clones (Fig. 2*B*), a reason also accounting for the fact that other cDNA clones of the pMR1 group showed no hybridization with the 0.5-kb *Bam*HI fragment or hybridizations of different intensities with the 1.3-kb *Eco*RI fragment, or both. Interestingly, one of the members of the pMR2 family has been found in an actin genomic clone (λMac62; ref. 12).

The other two families of mouse repeated sequences are very different from the previous ones and share common properties. These repeated sequences appear to belong to the B1 family, as indicated by the fact that the two B1 plasmids Mm31 and Mm35 hybridized with pMR3 and pMR4, respectively. In addition, they hybridize with one 6-kb *Eco*RI fragment and with another *Eco*RI fragment that is 2.5 kb in the case of pMR3 and 3.5 kb in the case of pMR4. These fragments, mainly present on components 1.701 and 1.704, appear to correspond to families of long repeated sequences containing the repeats of pMR3 and pMR4, respectively. The finding of hybridization bands and the lack of cross-hybridization of individual B1 sequences might be due to the relatively high stringency of washing used in this work (0.218 M Na$^+$ and 65°C).

The results obtained by hybridizing plasmids pCaα6-5 and Blur 8 with restriction digests of human DNA show an overall similarity with those just described for the two pairs of mouse plasmids. The first plasmid hybridized with *Hin*dIII and *Kpn* I repeated sequence fragments detected by ethidium bromide staining and was mainly present in the lighter components; much weaker hybridization bands, mostly at the same position as the strongest ones found in the light components, also could be detected in the 1.708 component. In all components, except 1.708, hybridization bands were accompanied by a hybridization smear. The hybridization pattern shown by Blur 8 was quite different in that only a smear was present without detectable bands but with increasing intensity from the light to the heavy components. Finally, the α-repeats were found to hybridize on *Eco*RI fragments from the three lightest components. This result provides information on the genomic distribution of the sequences and confirms that, in spite of the fact that human α repeats can exist as tandem arrays (up to 40-mers) of 170-bp units with specific chromosomal localization at the centromeres (33), some members of this family are also embedded in nonrepeated sequences (34).

As far as transcription of human and mouse repeated sequences is concerned, the present results confirm the fact that repeats of the B1 family in mouse are transcribed (35), like the repeats of the human *Alu* I family (36). On the other hand, our data provide novel information on the transcription of the mouse repeats of the *Bam*HI family. These were found to be present in cDNA clones obtained from polyribosomal poly(A)$^+$ mRNA and appear to present a tissue specificity, because brain mRNA contains a higher proportion of *Bam*HI transcripts than does

liver mRNA (data not shown). The different restriction maps of the cDNA clones are likely to be due to the existence of a wide variety of repeats in the *Bam*HI family and to the fact that the cloned repeats may be only partial copies of the repeats present in the poly(A)$^+$ polyribosomal mRNA, but they also might be explained by different splicings of more similar or identical primary transcripts. Finally, the rather weak transcription detected by hybridizing the pCaα6-5 plasmid with HeLa cell RNA parallels the results obtained with the *Bam*HI family. It should be pointed out, however, that while the latter results may be accounted for by the underrepresentation of repeat transcripts in total HeLa cell RNA, the former are in apparent conflict with our finding of long repeats in one-fourth of all plasmids containing repeats. However, this may be due to a number of factors, such as an underrepresentation of some mRNAs in cDNA clones owing to differences in the secondary structures of mRNAs and an underestimate of the plasmids containing short repeats by the colony hybridization with labeled total mouse DNA.

The main conclusions reached on the basis of the present work can be summarized as follows. Of the two families of interspersed repeated sequences studied here, the long repeats appear to be almost entirely localized in the two light components in mouse. In the case of man they extend to the 1.704 component and can be detected in small amounts even in the 1.708 component. The almost exclusive or predominant localization of long repeats in the light components, which represent together about 60% of mouse and human main-band DNAs, indicates that long repeats must be rare in the chromosomal segments corresponding to the heavy isochores. In contrast, the short repeats increase in relative amounts when going from the lightest to the heaviest components; this distribution is also found for the short-repeat family CR1 of chicken (unpublished data). Because the two heavy components only represent about one-third of main-band DNA in both mouse and man and even less in chicken, the number of short repeats in the light components is still considerable, but their concentration is much lower than in the heavy components. The report of a positive hybridization of the Blur 8 probe with over 90% of the clones from a human genomic library (28) is not in contradiction with our results because the human fragments from the library had an average size of 18 kb and were tested under conditions in which a single *Alu* I sequence, representing 1.6% of the fragment, could be detected. An important feature of the genomic distribution of both long and short repeats just described is its high conservation in evolution. Such conservation is paralleled by the sequence homology found in the short repeats of mammalian and chicken genomes (31, 32). It should be stressed that the evolutionary conservation of the genomic distribution of long- and short-repeat families studied here is not seen when the distribution of all interspersed repeats in the major components is investigated by reassociation kinetics (6, 7). The reason for this is that the long- and short-repeat families under consideration vary in relative amounts in the DNAs of different species and are accompanied by other families of repeats as judged from published estimates. Another observation of general interest is that the long and short repeats differ in their base composition, the long ones being lower in dG+dC than the short ones in the different species tested, matching the base compositions of the major components embedding them (Table 1). Finally, it is very likely that the long repeats are less represented in mouse mature transcripts and in human total transcripts than are the short repeats. The repeat length, the frequency, the genome distribution, the base composition, and the representation in transcripts are a number of important differences between the two main families of repeats of mammalian genomes, for which the names SINES and LINES have been proposed (8). Further implications of the present results, will be discussed in detail elsewhere.

We thank Drs. F. Brégégère, G. Georgiev, W. Jelinek, and M. Singer for the cloned probes used in this work and Dr. M. Singer for useful discussions.

1. Britten, R. J. (1981) in *Chromosomes Today*, eds. Bennett, M. D., Bobrow, M. & Hewitt, G. M. (Allen & Unwin, Winchester, MA), Vol. 7, pp. 9–23.
2. Filipski, J., Thiéry, J. P. & Bernardi, G. (1973) *J. Mol. Biol.* **80**, 177–197.
3. Thiéry, J. P., Macaya, G. & Bernardi, G. (1976) *J. Mol. Biol.* **108**, 219–235.
4. Macaya, G., Thiéry, J. P. & Bernardi, G. (1976) *J. Mol. Biol.* **108**, 237–254.
5. Cuny, G., Soriano, P., Macaya, G. & Bernardi, G. (1981) *Eur. J. Biochem.* **115**, 227–233.
6. Soriano, P., Macaya, G. & Bernardi, G. (1981) *Eur. J. Biochem.* **115**, 234–239.
7. Olofsson, B. & Bernardi, G. (1983) *Eur. J. Biochem.* **130**, 241–245.
8. Singer, M. F. (1982) *Cell* **28**, 433–434.
9. Singer, M. F. (1982) *Int. Rev. Cytol.* **76**, 67–112.
10. Meunier-Rotival, M., Soriano, P., Cuny, G., Strauss, F. & Bernardi, G. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 355–359.
11. Ramsey, J. C. & Steele, W. J. (1976) *Biochemistry* **15**, 1704–1712.
12. Dodemont, H. J., Soriano, P., Quax, W. J., Ramaekers, F., Lenstra, J. A., Groenen, M. A. M., Bernardi, G. & Bloemendal, H. (1982) *EMBO J.* **1**, 167–172.
13. Grunstein, M. & Hogness, D. S. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 3961–3965.
14. Rubin, C. M., Deininger, P. L., Houck, C. M. & Schmid, C. W. (1980) *J. Mol. Biol.* **136**, 151–167.
15. Grimaldi, G., Queen, C. & Singer, M. F. (1981) *Nucleic Acids Res.* **9**, 5553–5568.
16. Krayev, A. S., Kramerov, D. A., Skryabin, K. G., Ryskov, A. P., Bayev, A. A. & Georgiev, F. P. (1980) *Nucleic Acids Res.* **8**, 1201–1215.
17. Brégégère, F., Abastado, J. P., Kvist, S., Rask, L., Lalanne, J. C., Garoff, H., Cami, B., Wiman, K., Larhammat, D., Peterson, P. A., Gachelin, G., Kourilsky, P. & Dobberstein, B. (1981) *Nature (London)* **292**, 78–81.
18. Cortadas, J., Macaya, G. & Bernardi, G. (1977) *Eur. J. Biochem.* **76**, 13–19.
19. Prunell, A. & Bernardi, G. (1973) *J. Biol. Chem.* **248**, 3433–3440.
20. Eigner, J. & Doty, P. (1965) *J. Mol. Biol.* **12**, 549–580.
21. Cortadas, J., Olofsson, B., Meunier-Rotival, M., Macaya, G. & Bernardi, G. (1979) *Eur. J. Biochem.* **99**, 179–186.
22. Southern, E. M. (1975) *J. Mol. Biol.* **94**, 51–69.
23. Soriano, P., Szabo, P. & Bernardi, G. (1982) *EMBO J.* **5**, 579–583.
24. McMaster, G. K. & Carmichael, G. C. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 4835–4838.
25. Thomas, P. S. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 5201–5205.
26. Cami, B., Brégégère, F., Abastado, J. P. & Kourilsky, P. (1981) *Nature (London)* **291**, 673–675.
27. Steinmetz, M., Frelinger, J. G., Fisher, D., Hunkapiller, T., Pereira, D., Weissman, S. M., Uehara, H., Nathenson, S. & Hood, L. (1981) *Cell* **24**, 125–134.
28. Tashima, M., Calabretta, B., Torelli, G., Scofield, M., Maizel, A. L. & Saunders, G. F. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 1508–1513.
29. Georgiev, G. P., Ilyin, Y. V., Chmeliauskaite, V. G., Ryskov, A. P., Kramerov, D. A., Skryabin, K. G., Krayev, A. S., Lukanidin, E. M. & Grigoryan, M. S. (1981) *Cold Spring Harbor Symp. Quant. Biol.* **45**, 641–654.
30. Wu, J. C. & Manuelidis, L. (1980) *J. Mol. Biol.* **142**, 363–386.
31. Schmid, C. W. & Jelinek, W. R. (1982) *Science* **216**, 1065–1070.
32. Stumph, W. E., Kristo, P., Tsai, M.-J. & O'Malley, B. W. (1981) *Nucleic Acids Res.* **9**, 5383–5397.
33. Manuelidis, L. (1979) *Chromosoma* **66**, 23–32.
34. Darling, S. M., Crampton, J. M. & Williamson, R. (1982) *J. Mol. Biol.* **154**, 51–63.
35. Kramerov, D. A., Grigoryan, A. A., Ryskov, A. P. & Georgiev, G. P. (1979) *Nucleic Acids Res.* **6**, 699–713.
36. Pan, J., Elder, J. T., Duncan, C. & Weissman, S. M. (1981) *Nucleic Acids Res.* **9**, 1151–1170.