Eur. J. Biochem. *115*, 227–233 (1981)
© FEBS 1981

The Major Components of the Mouse and Human Genomes

1. Preparation, Basic Properties and Compositional Heterogeneity

Gérard CUNY, Philippe SORIANO, Gabriel MACAYA, and Giorgio BERNARDI

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

(Received October 27, 1980)

Main-band DNA from mammals and birds can be resolved by density gradient centrifugation techniques into three or four families of fragments of different dG + dC contents. 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.

In the present work, the four major components of mouse and human DNAs were prepared and characterized in several basic properties: relative amounts, dG + dC contents, buoyant densities and compositional heterogeneity. The results obtained lead to the following conclusions: (a) the major DNA components of mouse and man form at least 85% and possibly the totality of the main bands of these DNAs; (b) they have very low compositional heterogeneities over a wide molecular weight range; (c) they derive from very large chromosomal DNA segments of fairly homogeneous base composition, for which the name 'isochores' is proposed.

A comparison of the compositional heterogeneity of main-band DNAs from warm-blooded and cold-blooded vertebrates confirms our previous conclusion that these DNAs are characterized by a different sequence organization.

At the present time, genome organization and regulation of gene expression are reasonably well understood in the case of prokaryotes, but problems are still open in the case of eukaryotes. This very wide gap in our knowledge is a most serious one, since these issues are crucial for our understanding of differentiation and evolution. Most of the current research activity on eukaryotic genomes is devoted to the study of the organization of individual genes. While this approach will hopefully shed some light on the regulation of gene expression, it may be of rather limited help as far as genome organization is concerned, since it deals with an extremely small fraction of the genome and neglects, by definition, the vast majority of the non-coding sequences which are so characteristic and abundant in eukaryotic genomes. For these reasons, the problem of sequence organization of such genomes is best studied by global approaches at the molecular level.

So far, the most widely used global approach has involved the study of the reassociation kinetics exhibited by denatured fragments of eukaryotic DNA. The main contributions of this approach have been: (a) the demonstration [1] that eukaryotic (in contrast to prokaryotic) genomes contain, beside unique sequences (namely sequences present only once or a very small number of times in the haploid genome), interspersed repeated sequences (namely sequences sharing a certain degree of homology and present a large number of times in the haploid genome); and (b) information on the interspersion patterns of these sequences in the genomes of a number of species. Two main interspersion patterns have been recognized; the first one, as found in *Xenopus* [2], is characterized by the fact that slightly more than half of the unique sequences, about 700 residues long, are interspersed with short repeated sequences 300–400 residues in length (short-period interspersion), the remaining unique sequences being at least 4000 residues long and being interspersed in a long-period interspersion pattern. The second one, the *Drosophila* pattern [3], consists of only a long-period interspersion.

Another global approach has been developed in our laboratory [4–11] and is based on the fractionation of native eukaryotic DNAs by density gradient centrifugation in the presence of DNA ligands, like Ag^+ or bis(acetato-mercurimethyl)dioxane (BAMD). In general, these techniques separate native DNA fragments containing repeated oligonucleotides according to sequence, and other DNA fragments containing short sequences closer to random according to base composition. They have not only allowed the separation of satellite DNAs [1] and of minor components, sometimes formed by repeated genes containing short repeated sequences in their spacers (see, for example, [12]) but also, more interestingly, the fractionation of main-band DNAs from higher vertebrates into a small number of discrete components.

In a study of DNAs from eukaryotes covering a very wide phylogenetic range [6], we observed that (neglecting satellite and minor components) DNAs from unicellular organisms and invertebrates exhibit symmetrical CsCl band profiles, as

Nomenclature. The nomenclature used here will follow conventions used in previous papers from our laboratory (e.g. [9]). We call DNA components (and indicate them by their buoyant densities in CsCl) the families of native genome fragments which can be separated from each other by density gradient centrifugation techniques. We distinguish three groups of DNA components according to their nucleotide sequence patterns and relative amounts in the genome as follows. (a) Satellite components are formed by short tandem repeated sequences; each satellite component usually represents a small percentage of the genome. (b) Minor components each account for less than 3% of the genome; they are not formed by short repeated nucleotide sequences, yet they may contain a certain amount of them in spacer sequences; several minor components are formed by repeated genes. (c) Major components, making up the main-band DNA, which corresponds to the bulk of mammalian and avian genome, contain interspersed unique and repetitive sequences.

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

do prokaryotic DNAs; DNAs from cold-blooded vertebrates show either no asymmetry or a slight asymmetry on the heavy side of their CsCl bands; DNAs from warm-blooded vertebrates are characterized by a strong asymmetry on the heavy side of their CsCl bands. This was shown to be due to the fact that mammalian and avian genomes contain two heavy DNA components, exhibiting buoyant densities of 1.704 and 1.708 g/cm^3 and representing about 25% and 10% of the genome, respectively, in addition to one or two (according to the species) light components. The latter have buoyant density(ies) in the range 1.697–1.703 g/cm^3 (which is the same range as for main-band DNAs from most cold-blooded vertebrates [6,13]) and represent about 60% of the genome. The fractionation of DNA fragments from mouse into three or four discrete major components was shown to be possible over a very wide molecular weight range, 2×10^6 to over 2×10^8 [7], indicating that the major components of these genomes (a) do not exhibit large intermolecular compositional heterogeneities at molecular weights as low as 2×10^6; (b) correspond to very long stretches of DNA having homogeneous average base compositions.

In the present work, we have isolated the four major components of two mammalian genomes, those of mouse and man, and determined their buoyant densities, relative amounts and dG+dC contents [2]. In addition, we have investigated the intermolecular compositional heterogeneity, or buoyant density heterogeneity, of these major components at different molecular weight levels and compared them to those of the major components of the chicken genome and of DNAs from fishes [13]. An investigation on the reassociation kinetics of the major components of mouse and human DNAs is reported in the following paper [14].

A brief, preliminary report on some aspects of the present work has been presented elsewhere [15].

MATERIALS AND METHODS

Preparation of the Major Components of the Mouse Genome

DNA was prepared according to the method of Kay et al. [16], as modified by Bernardi and Sadron [17], from thymus glands of 5–7-week-old Balb-c mice, and purified by chromatography on hydroxyapatite in the presence of 3 M KCl [5,18]; the sedimentation coefficient, $s_{20,w}$, of this preparation was 25 S. A different DNA preparation, $s_{20,w}$ = 39 S, was used to obtain an additional sample of the least abundant component, having a buoyant density of 1.708 g/cm^3.

The major components of mouse DNA were prepared by centrifugation in Cs_2SO_4/Ag^+ density gradient at pH 9.2, a choice suggested by previous work [5,6] as outlined below. In this gradient, dA+dT-rich fragments are heavier than dG+dC-rich fragments, but the dA+dT-rich satellite DNA of mouse remains light in the range of r_f (the Ag^+/nucleotide molar ratio) used (Fig. 1).

The first preparative centrifugation (Fig. 2) was performed on 627 A_{260} units at an r_f = 0.25, a value providing, according to analytical experiments (Fig. 1), the best separation of the satellite peak from the main band. Incidentally, under the experimental conditions used (DNA load and gradient slope) the satellite peak exhibited a fine structure with at least three different components (data not shown), similar to those reported by Zardi et al. [19]. All the DNA fractions from the first preparative centrifugation were examined by analytical centrifugation in CsCl, using the conditions of Thiery et al. [6]. The results so obtained were essentially identical to those presented in detail by Macaya et al. [7]. In the present work the gaussian analysis of the CsCl profiles of the fractions was not performed since these data were available from the previous work of Macaya et al. [7]. Instead, all profiles were simply examined as to their modal buoyant densities, band widths, asymmetries, presence of absence of shoulders, multimodality. Some fractions were found to correspond to almost pure components, as in the previous work of Macaya et al. [7]; most of them were, however, mixtures of neighboring components, as shown by their modal buoyant densities, which were intermediate between those of known components, and by the shape (band width and asymmetry) of their CsCl profiles. The criteria mentioned above were used to pool all the fractions (except for the heavy fractions containing the satellite DNA and several dG+dC-rich minor components and for the lightest fractions) into four pools, each one of which predominantly contained a given DNA component (Fig. 2).

Pools from the first centrifugation corresponding to the four major components of main-band DNA were recentrifuged in preparative Cs_2SO_4/Ag^+ density gradients at an r_f value of 0.28, a value providing a maximal spreading and skewness of the main band (Fig. 1), except in the case of the 1.708-g/cm^3 component where an r_f value of 0.25 was used to remove contaminating traces of satellite DNA. Again, all DNA fractions were examined by analytical centrifugation in CsCl. In all cases, the central fractions from each preparative gradient, which represented about 75% of main-band DNA, exhibited narrow, symmetrical bands, with practically identical modal buoyant densities, whereas side fractions were noticeably heterogeneous.

The central fractions from the second centrifugation were pooled and recentrifuged once more in Cs_2SO_4/Ag^+ density gradient at an r_f of 0.28. Analytical centrifugation in CsCl showed that this time not only the central fractions of the DNA peak, but also most of the side fractions were practically homogeneous. In each case, pools were made after the elimination of the outermost 3–5% of the DNA and shown by analytical centrifugation in CsCl to correspond to the major DNA components previously identified in the mouse genome (see Results, Fig. 3). These pools contained 70% of main band DNA (itself representing 89% of total DNA; the rest is formed by satellite and minor components, see [6]) after correction for mechanical losses at the preparative centrifugation steps, at dialysis and at analytical CsCl centrifugations.

The side fractions from the second centrifugation, which contained most of the remaining 30% of main-band DNA, were examined by preparative Cs_2SO_4/Ag^+ centrifugation followed by analytical CsCl centrifugation and found to consist, for at least 50%, of the expected major components. Further processing of this material was not performed because its enrichment in lower-molecular-weight DNA, present in the side fractions of preparative centrifugation bands, made its fractionation difficult. In estimating the relative amounts of the major components, the material from the side fractions was apportioned according to the amounts of the contiguous major components.

Preparation of the Major Components of the Human Genome

DNA was prepared and purified from one placenta recovered at a girl's birth, as described above. Its $s_{20,w}$ was 32 S. The major components were prepared as outlined below.

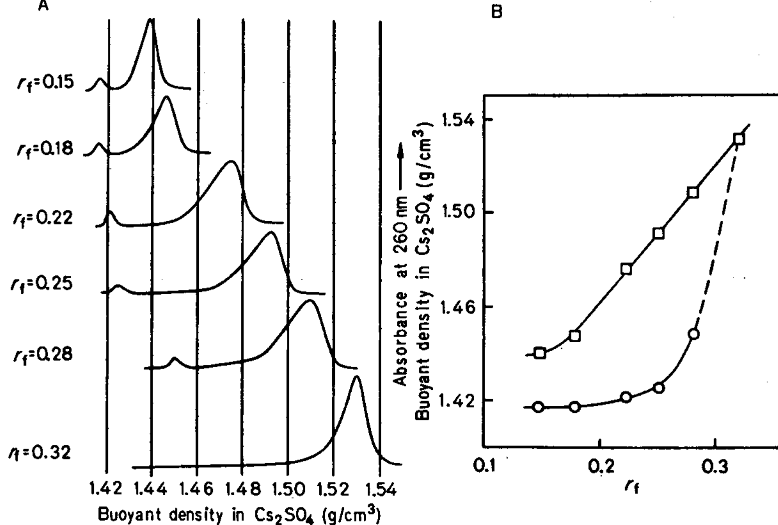

Fig. 1. *Centrifugation of total mouse DNA in analytical Cs_2SO_4/Ag^+ density gradients, at pH 9.2.* (A) Buoyant density profiles obtained at different values of r_f(Ag^+/nucleotide molar ratio) are shown. (B) The buoyant densities in Cs_2SO_4 of main-band (□) and satellite (○) DNAs are plotted against r_f values

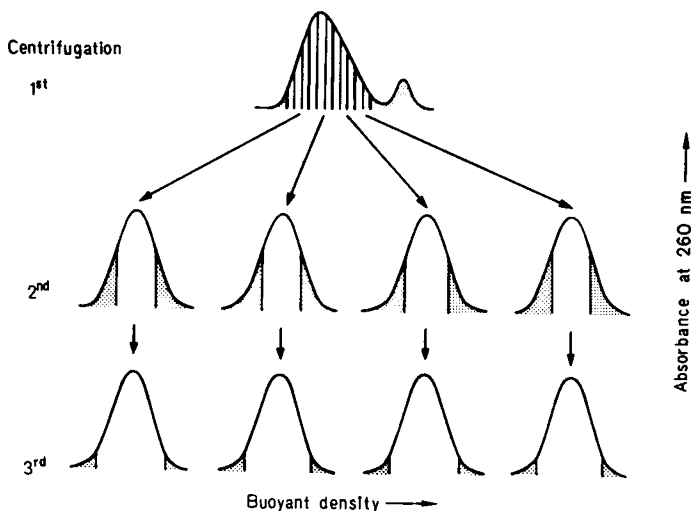

Fig. 2. *Scheme of the procedure used to prepare the major components of the mouse genome.* Details are given in Materials and Methods. Shadowed fractions were not processed further

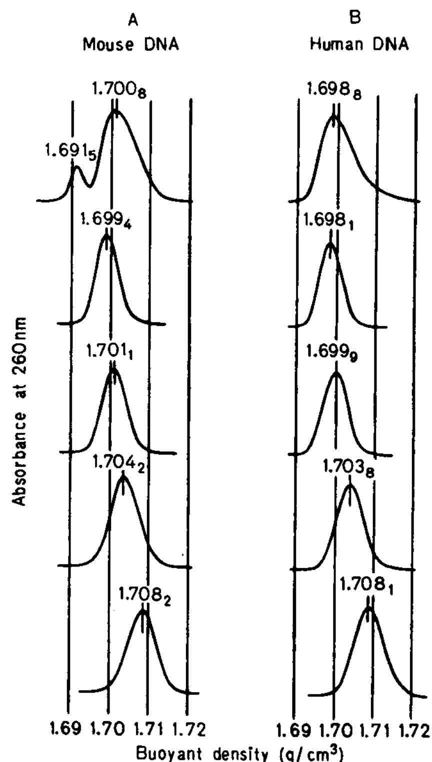

Fig. 3. *Analytical CsCl density profiles of (A) total mouse DNA and its major components with buoyant densities of 1.699, 1.701, 1.704 and 1.708 g/cm³, and (B) total human DNA and its major components with buoyant densities of 1.698, 1.700, 1.704 and 1.708 g/cm³*

The first preparative Cs_2SO_4/Ag^+ density gradient centrifugation was performed on 716 A_{260} units at pH 9.2 and at an r_f value of 0.32, a choice determined by an experiment, identical to that of Fig. 1, reported by Macaya et al. [8]. 25–30 fractions were collected and analyzed by analytical CsCl density gradient centrifugation. Six pools were made from these fractions using the criteria presented in the preceding section.

The two extreme pools, representing about 7% of total DNA, contained satellite and minor components and were not investigated further; the other four pools, corresponding to the four major components, were centrifuged once more in preparative Cs_2SO_4/Ag^+ density gradients, using r_f values providing an optimal resolution of the major components ($r_f = 0.32$, except for the 1.704-g/cm³ component, where the r_f value was 0.41). Once again, fractions were collected and analyzed by analytical CsCl density gradient centrifugation. Central fractions, corresponding to essentially pure components, were pooled; the side fractions displayed considerable heterogeneity.

The four central pools corresponding to the four major components were centrifuged a third time in Cs_2SO_4/Ag^+ density gradients; the outermost, slightly heterogeneous side fractions, which amounted to less than 10% of each component, were eliminated and central fractions, containing purified components, were pooled. The final yield of the major components was about 70% of main-band DNA.

Other Methods

Sedimentation velocity experiments were done as described [20]. Base compositions were determined by chromatography on Aminex A6 (Bio-Rad, Richmond, CA) of nucleosides obtained by enzymatic digestion [21]. Intermolecular compositional heterogeneity or buoyant density heterogeneity was estimated according to Schmid and Hearst [22], using an independent measurement by sedimentation velocity [23] and/or gel electrophoresis [24] of the molecular weights of DNA samples.

Table 1. *Basic properties of mouse and human DNAs and their components* ϱ_0 is the modal buoyant density in CsCl; the error is ± 0.0005 g/cm³. H is the intermolecular, or buoyant density, heterogeneity calculated as a percentage of dG + dC content. dG + dC content was calculated from ϱ_0 values according to the relationship of Schildkraut et al. [25] or determined by nucleoside analysis; the error is $\pm 0.5\%$

Species	DNA	Amount	ϱ_0	H	dG+dC content	
					calculated	determined
		%	g/cm³	%		
Mouse	Main-band	93	1.700_8	4.0	41.6	40.6
	1.699-g/cm³	22	1.699_4	2.5	40.2	38.0
	1.701-g/cm³	34	1.701_1	2.9	41.9	40.2
	1.704-g/cm³	26	1.704_2	3.1	45.1	43.2
	1.708-g/cm³	7	1.708_2	3.5	49.2	48.0
	Satellite	7	1.691_5	1.0	32.1	36.0
Human	Total	100	1.698_8	4.8	39.6	40.3
	1.698-g/cm³	29	1.698_1	2.6	38.7	36.7
	1.700-g/cm³	33	1.699_9	2.7	40.7	38.5
	1.704-g/cm³	22	1.703_8	3.2	44.7	42.9
	1.708-g/cm³	9	1.708_1	3.8	49.1	49.2

RESULTS

Preparation of the Major Components by Mouse and Human DNAs

The preparative procedure used provided three main results. First, a comparison of the analytical CsCl profiles (not shown) of fractions obtained from the preparative density gradients with those obtained in previous work [6, 7] indicates that a better separation of the major components was achieved here. This appears to be due to the higher molecular weight and the lower diffusion coefficient of the DNA preparations used in the present work, as shown by the better separation of major components from 39-S DNA compared to 25-S DNA preparations from mouse (results not shown).

Secondly, as indicated in the preceding section (see also Fig. 2), the side fractions of the second centrifugation (a) had modal buoyant densities intermediate between those of the neighboring (practically pure) major components, (b) were broad or multimodal in CsCl profiles since they were formed by mixtures of those components and (c) were present in small amounts compared to neighboring major components, thus marking discontinuities in a plot of DNA amount versus buoyant density.

Thirdly, the preparative approach used here has proved that the major components form at least 85% and very possibly all of the main-band DNAs from mouse and man, as hinted by previous gaussian analysis [6]. In other words, the existence in the main band of other major components having buoyant densities different from those of the components studied here is highly unlikely. Finally, it should be mentioned that major components indistinguishable in their properties from those studied here have been prepared from mouse DNA using the Cs_2SO_4/BAMD centrifugation approach, also used for preparing the major components of the chicken genome [11].

Basic Properties of Major Components

The relative amounts of the major components in the mouse and human genomes as determined from their yields are given in Table 1, along with the corresponding modal buoyant densities, ϱ_0, the intermolecular, or buoyant density, heterogeneities, H, and dG + dC contents derived from ϱ_0 or from nucleoside analysis.

The CsCl band profiles (Fig. 3) are characterized by a remarkable symmetry, the mean buoyant density, $\langle \varrho \rangle$, (see [6] for a definition) being identical with ϱ_0; this is in sharp contrast with the asymmetry of the profiles of unfractionated DNAs. As expected on the basis of previous work [6], both buoyant densities and relative amounts of corresponding major components are very similar in the two genomes.

The modal buoyant densities and the relative amounts are in good agreement with those of Thiery et al. [6], except for the relative amounts of the light components. The overall agreement is remarkable in view of the fact that in the present work ϱ_0 values were determined on the components as isolated by preparative centrifugation, and relative amounts were based on the yields of the components, whereas in previous work both sets of results were obtained from the gaussian analysis of CsCl bands of total DNA or Cs_2SO_4/Ag^+ fractions. The discrepancy concerning the light components is probably due in part to their incomplete preparative separation and in part to the difficulty of obtained precise determinations of relative amounts by gaussian analysis of components so close in both quantity and buoyant density. In any case, it should be recalled that the two light components were previously considered to correspond to subcomponents, because they form a single component in some mammalian genomes [6]; the results of the present work confirm a great similarity in the properties of the two light components of both mouse and human genomes. This is also supported by reassociation studies [14].

The estimates of dG+dC contents of the major components as obtained from ϱ_0 (using a relationship established for bacterial DNAs [25]) were systematically higher compared to the results of chemical analysis (Table 1). A first possible explanation for such discrepancy is methylation, since a comparison (not shown) of digest patterns from the major components by restriction endonuclease *Hpa*II, an enzyme which only splits the sequence d(C-C-G-G), and by *Msp*I, an enzyme which splits both d(C-mC-G-G) and d(C-C-G-G) [26], indicated that all major components are methylated, roughly to similar extents. This possibility is, however, ruled out by the fact that methylation causes a decrease in buoyant density [27] and should therefore lead to dG + dC estimates deviating from those of chemical analysis in a direction opposite to that observed. Another possible explanation is a different frequency of short oligonucleotides in mouse and human DNAs and in their components compared to bacterial DNAs [28, 29]. This

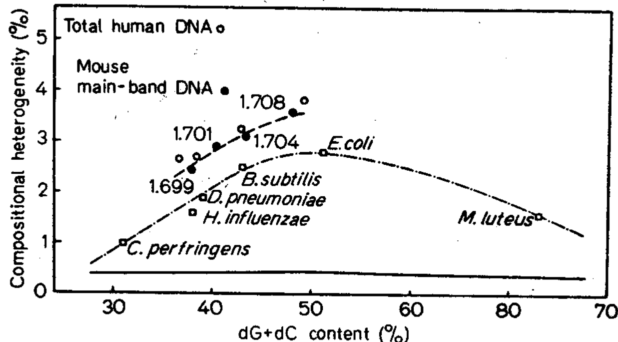

Fig. 4. *The intermolecular compositional heterogeneity, or buoyant density heterogeneity, plotted against composition for statistical DNAs (——), bacterial DNAs (□), mouse major components and main-band DNA (●) and human major components and total DNA (○).* Values for statistical DNAs were calculated according to Sueoka [33]; those for bacterial DNAs were estimated by Sueoka [33] for *D. pneumonia* (39% dG + dC), by Yamagishi [38] for *C. perfringens* (31% dG + dC), *B. subtilis* (43% dG + dC) and *M. luteus* (73% dG + dC), or were calculated according to Schmid and Hearst [22] for *H. influenzae* (38% dG + dC) and *E. coli* (51% dG + dC). All compositions are from Devillers-Thiéry [29] except for that of *D. pneumoniae*, which is from Belozersky and Spirin [50]. The different mouse DNAs are indicated by their buoyant density values (g/cm³)

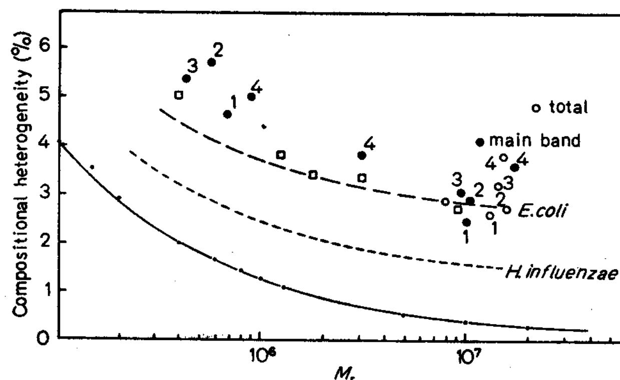

Fig. 5. *Buoyant density heterogeneity plotted against molecular weight for a statistical DNA,* H. influenzae *DNA,* E. coli *DNA, and mouse and human DNAs.* The values for statistical DNA (points) were obtained by simulating a cleavage by restriction enzymes of a DNA having a random sequence built using the Monte Carlo method, with a Control Data 6600 computer; the solid line is the relationship predicted by Sueoka [33]. The values for *E. coli* DNA (□) are from Miyazawa and Thomas [35] for the lowest-molecular-weight sample, from Yamagishi [38], or calculated according to Schmid and Hearst [22] for the other ones. The values for *H. influenzae* DNA were also calculated according to Schmid and Hearst [22]. The values for mouse DNA components (●) are noted 1–4 for components with ϱ_0 of 1.699–1.708 g/cm³. Values for components with ϱ_0 = 1.699, 1.701, and 1.704 g/cm³ are for undegraded and *Hae*III-degraded samples. Values for the 1.708-g/cm³ component are from undegraded, *Hpa*II-degraded, and *Ava*II-degraded samples, respectively. The values for human DNA components (○) are likewise noted 1–4 for components with ϱ_0 of 1.698–1.708 g/cm³

appears to be the correct explanation since it is well established that such frequencies affect the buoyant density of DNA [4,30,31].

Along the same line, data on main-band and satellite DNA from mouse (Table 1) are in agreement with previous ones [4]. Interestingly, while the dG + dC estimate from ϱ_0 for main-band DNA expectedly shows the same derivation from the analytical value as for major components, that for satellite DNA shows a stronger deviation in the opposite direction; such a deviation is likely to be largely due to the high methylation level [32] of this DNA component, but the particular frequency of short nucleotide sequences in satellite DNA may also play a role.

Compositional Heterogeneity of Major Components

Two points should be mentioned before describing the results obtained on the intermolecular compositional heterogeneity, or buoyant density heterogeneity, of mouse and human DNAs and their major components. (a) For DNAs with a random sequence, it was predicted [33] that the compositional heterogeneity increases when dG + dC approaches 50% and when the molecular weight of the fragments decreases; we have verified both effects (Fig. 4 and 5), the first one however is extremely small. (b) For bacterial DNAs, the effect of dG + dC contents on heterogeneity was found to be strong (Fig. 4) and the effect of molecular weight decrease was found to parallel that shown by DNA with a random sequence (Fig. 5).

At high molecular weight ($M_r > 10^7$), the heterogeneities of the major components of mouse and human DNAs varied with dG + dC contents (Fig. 4 and Table 1) like bacterial DNAs, but were systematically slightly higher than the latter; expectedly, they were lower than those of mouse main-band DNA and human total DNA (the latter being higher than the former for reasons given in the Discussion).

In plots against fragment molecular weight (Fig. 5), the heterogeneities of mammalian major components were slightly higher than those of bacterial DNAs of the same base composition; in the case of the light components of mouse they also showed a higher increase with molecular weight decrease. It should be noted that even at the lowest molecular weight examined, no increase in the asymmetry of CsCl bands, nor any fractionation in Cs_2SO_4/Ag^+ preparative density gradients were observed (data not shown).

Fig. 6 presents the heterogeneity ratio of main-band mouse DNA, total human DNA, their major components, and mouse satellite DNA. The heterogeneity ratio is the result of dividing the heterogeneity of a eukaryotic DNA by that of a bacterial DNA having the same molecular weight and the same dG + dC content. This normalization facilitates comparison which is extended, in Fig. 6, to chicken total DNA and its major components [11] and to fish main-band DNA (the data presented is, in this case, the average of DNAs from 33 different species [13]).

DISCUSSION

Most of the results on the preparation and the basic properties of the major components of mouse and human DNAs have already been commented upon in Results above and do not need further discussion. This section will, therefore, be devoted to the compositional heterogeneity of the major components and to the origin of the latter.

While some information on compositional heterogeneity of phage and bacterial DNA is available [22,33–38], very little exists in the literature on eukaryotic DNAs, apart from the pioneering work of Sueoka [33,34]. The results of the

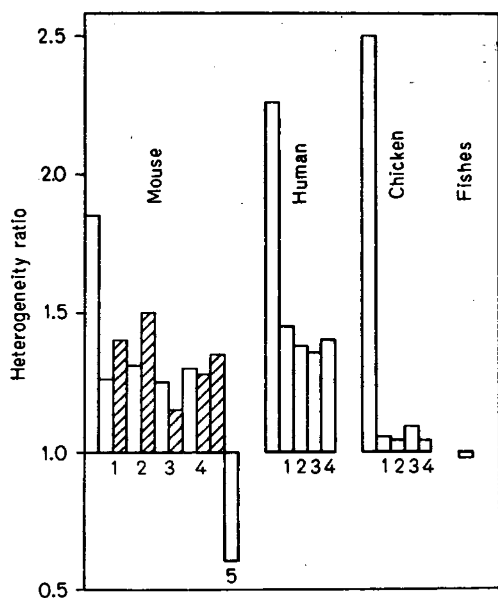

Fig. 6. *Histogram of heterogeneity ratios for main-band mouse DNA, total human and chicken DNAs and their major components (marked 1–4 in order of increasing density, 5 is for satellite DNA)*. In the case of mouse DNA, hatched bars refer to degraded samples (see legend of Fig. 5). Also shown is the average ratio for 33 fish DNAs [13]

present work provide, therefore, the first quantitative estimates of compositional heterogeneity of two mammalian DNAs and their major components.

The compositional heterogeneities of the major components of mouse and human DNAs at high molecular weight ($> 10^7$) are not much larger (only by 20–40%) than those of comparable bacterial DNAs (namely bacterial DNAs having the same molecular weight and the same dG + dC content); expectedly, such heterogeneities are lower than those of mouse main-band DNA and human total DNA and of degraded light components but higher than that of mouse satellite DNA. These results are remarkable in two respects: (a) because of the much higher and different complexities of major components compared to bacterial DNA; (b) because of the very different ratios of unique to repeated sequences in different major components [14]. In other words, neither 50–200-fold increases in complexity relative to bacterial DNAs nor different relative amounts of repeated sequences appear to affect significantly the intermolecular heterogeneity of the components. It should be added here, however, that upon degradation to molecular weights in the 5×10^5 range the heterogeneity of the light components (which contain most of the repeated sequences in mouse DNA [14]) increases up to 50% relative to comparable bacterial DNAs. This increase is not accompanied by any alteration of band symmetry, nor by any fractionation upon Cs_2SO_4/Ag^+ preparative centrifugation.

The data of Fig. 6 indicate that the chicken and human total DNAs are more heterogeneous than main-band mouse DNA. This is due to the presence in the former two DNAs of a number of satellite DNAs, which have been removed from the latter; in the case of human DNA two other factors are to be taken into account: the wider spread in modal buoyant densities of major components (the lightest component, at 1.698 g/cm^3 is 1 mg/cm^3 lighter than those of the other two genomes) and the greater heterogeneity of isolated major components. The extremely low heterogeneity of major components from chicken DNA and of the average of main-band DNAs from fishes is very striking; in the last case, many DNAs do exhibit a heterogeneity which is lower than that of comparable bacterial DNAs [13], as is the case of mouse satellite DNA.

A comparison of the compositional heterogeneities of main-band DNAs from warm-blooded and cold-blooded vertebrates (Fig. 6; see also [13]) confirms our previous conclusion [6] that a major discontinuity exists in the sequence organization of the genomes of vertebrates. Cold-blooded vertebrates are characterized by DNAs exhibiting a low degree of compositional heterogeneity; DNA from warm-blooded vertebrates, in contrast, is remarkably heterogeneous, essentially because it contains dG + dC-rich components which are either totally absent or very scarce in cold-blooded vertebrates.

Several lines of evidence indicate that the families of DNA fragments, called here major components, originate from very large chromosomal DNA segments of fairly homogeneous dG + dC contents, for which we propose the name isochores (meaning similar regions).

a) A comparison of the present results with previous ones has shown that the higher the molecular weight of DNA preparations used, the better the resolution of major components in preparative density gradients; obviously an opposite result would be expected if the DNA fragments forming different major components were interspersed at molecular weights of $10^7 - 10^8$.

b) The results of Cortadas et al. [11] on the chicken genome have likewise shown excellent separations for DNA fragments having molecular weights of about 50×10^6. In this case, the distribution of DNA fragments carrying the ovalbumin gene in density gradient fractions exactly follows the distribution of a given component ($\varrho_0 = 1.7025$ g/cm^3), proving that the gene is carried by DNA fragments belonging to this component. The chromosomal segment represented by the population of overlapping fragments containing the gene has an M_r as large as 10^8, namely 25 times larger than the smallest fragment containing all the coding sequence (4×10^6 [11]). This indicates homogeneity of base composition over chromosomal fragments as large as 10^8 (about 1.5×10^5 base pairs).

c) These conclusions are in agreement with the findings of Macaya et al. [7] that CsCl profiles of mouse DNA of molecular weight higher than 2×10^8 show discontinuities corresponding to the buoyant densities of the major components.

The points just made indicate that we can only estimate the lower limits of the isochore size, the actual values are not known. Likewise, we do not know whether there is a relationship between isochores and chromosome organization. We would like to suggest, as a working hypothesis which can be put to experimental test, that isochores may correspond to the Giemsa and reverse bands of the chromosomes of higher vertebrates. The reasons for such a suggestion are as follows. (a) There are strong indications [39] that Giemsa bands correspond to dA + dT-rich, late-replicating DNA and reverse bands to dG + dC-rich, early-replicating DNA; similarly, light and heavy isochores are, in all likelihood, interspersed on chromosomes. (b) Giemsa banding is very evident in warm-blooded, but not in cold-blooded vertebrates [40–42] (and personal communications from W. Beçak and P. Leon), a feature paralleled by the remarkably low compositional heterogeneity of cold-blooded vertebrates. (c) Giemsa-banding patterns appear to be highly conserved in birds [41] and throughout orders of mammals [43–48] as are

the relative amounts and dG + dC contents of major components. (d) The amount of DNA per chromosome band (account being taken of the high resolution achieved by Yunis [49] is compatible with that present in isochores.

We wish to thank Dr Hugues Blanc for his participation in several experiments involving restriction enzyme degradations, as well as for writing some of the computer programs, Ms Martine Brient for typing this manuscript, Mr Philippe Breton for the drawings and Mr François Strauss for his help with some of the computer programs.

REFERENCES

1. Britten, R. J. & Kohne, D. E. (1968) *Science (Wash. DC) 161*, 529–540.
2. Davidson, E. H., Hough, B. R., Amenson, C. S. & Britten, R. J. (1973) *J. Mol. Biol. 77*, 1–23.
3. Manning, J. E., Schmid, C. W. & Davidson, N. (1975) *Cell, 4*, 141–155.
4. Corneo, G., Ginelli, E., Soave, C. & Bernardi, G. (1968) *Biochemistry, 7*, 4373–4379.
5. Filipski, J., Thiery, J. P. & Bernardi, G. (1973) *J. Mol. Biol. 80*, 177–197.
6. Thiery, J. P., Macaya, G. & Bernardi, G. (1976) *J. Mol. Biol. 108*, 219–235.
7. Macaya, G., Thiéry, J. P. & Bernardi, G. (1976) *J. Mol. Biol. 108*, 237–254.
8. Macaya, G., Thiéry, J. P. & Bernardi, G. (1977) in *Molecular Structure of Human Chromosomes* (Yunis, J. J., ed.) pp. 1–33, Academic Press, New York.
9. Macaya, G., Cortadas, J. & Bernardi, G. (1978) *Eur. J. Biochem. 84*, 179–188.
10. Cortadas, J., Macaya, G. & Bernardi, G. (1977) *Eur. J. Biochem. 76*, 13–19.
11. Cortadas, J., Olofsson, B., Meunier-Rotival, M., Macaya, G. & Bernardi, G. (1979) *Eur. J. Biochem. 99*, 179–186.
12. Meunier-Rotival, M., Cortadas, J., Macaya, G. & Bernardi, G. (1979) *Nucleic Acids Res. 6*, 2109–2123.
13. Hudson, A., Cuny, G., Cortadas, J., Haschemeyer, A. E. V. & Bernardi. G. (1980) *Eur. J. Biochem. 112*, 203–210.
14. Soriano, P., Macaya, G. & Bernardi, G. (1981) *Eur. J. Biochem. 115*, 235–239.
15. Bernardi, G. (1979) in *Recombinant DNA and Genetic Experimentation* (Morgan, J. & Whelan, W. J., eds) pp. 15–20, Pergamon Press, New York.
16. Kay, E. R. M., Simmons, N. S. & Dounce, A. L. (1952) *J. Am. Chem. Soc. 74*, 1724–1726.
17. Bernardi, G. & Sadron, C. (1964) *Biochemistry, 3*, 1411–1418.
18. Bernardi, G. (1971) *Methods Enzymol. 21D*, 95–139.
19. Zardi, L., Siri, A., Santi, L. & Rovera, G. (1977) *FEBS Lett. 79*, 188–190.
20. Prunell, A. & Bernardi, G. (1973) *J. Biol. Chem. 248*, 3433–3440.
21. Thiery, J. P., Ehrlich, S. D., Devillers-Thiery, A. & Bernardi, G. (1973) *Eur. J. Biochem. 38*, 434–442.
22. Schmid, C. W. & Hearst, J. E. (1972) *Biopolymers, 11*, 1913–1918.
23. Eigner, J. & Doty, P. (1965) *J. Mol. Biol. 12*, 549.
24. Prunell, A., Kopecka, H., Strauss, F. & Bernardi, G. (1977) *J. Mol. Biol. 110*, 17–52.
25. Schildkraut, C. L., Marmur, J. & Doty, P. (1962) *J. Mol. Biol. 4*, 430–443.
26. Waalwijk, C. & Flavell, R. A. (1978) *Nucleic Acids Res. 5*, 4631–4641.
27. Kirk, J. T. O. (1967) *J. Mol. Biol. 28*, 171–172.
28. Bernardi, G., Ehrlich, S. D. & Thiery, J. P. (1973) *Nat. New Biol. 246*, 36–40.
29. Devillers-Thiéry, A. (1974) Thesis, Université Paris VII.
30. Wells, R. D. & Blair, J. E. (1967) *J. Mol. Biol. 27*, 273–288.
31. Wells, R. D., Larson, J. E., Grant, R. C., Shortle, B. E. & Cantor, C. R. (1970) *J. Mol. Biol. 54*, 465–497.
32. Salomon, R., Kaye, A. M. & Herzberg, M. (1968) *J. Mol. Biol. 43*, 581–592.
33. Sueoka, N. (1962) *Proc. Natl Acad. Sci. USA, 48*, 582–592.
34. Sueoka, N. (1959) *Proc. Natl Acad. Sci. USA, 45*, 1480–1490.
35. Miyazawa, Y. & Thomas, C. A. (1965) *J. Mol. Biol. 11*, 223–237.
36. Yamagishi, H. (1970) *J. Mol. Biol. 49*, 603–608.
37. Yamagishi, H. (1971) *J. Mol. Biol. 57*, 369–371.
38. Yamagishi, H. (1974) *J. Mol. Biol. 3*, 239–242.
39. Comings, D. E. (1978) *Annu. Rev. Genet. 12*, 25–46.
40. Bailly, S., Guillemin, C. & Labrousse, M. (1973) *C.R. Acad. Sci. Paris, 276*, 1867–1869.
41. Stock, A. D. & Mengden, G. A. (1975) *Chromosoma, 50*, 69–77.
42. Schmid, M. (1978) *Chromosoma, 68*, 131–148.
43. Dutrillaux, B., Rethoré, M. O. & Lejeune, J. (1975) *Ann. Genet. 18*, 153–161.
44. Wurster-Hill, D. H. & Gray, C. W. (1975) *Cytogenet. Cell. Genet. 15*, 306–331.
45. Dutrillaux, B. (1977) in *Molecular Structure of Human Chromosomes* (Yunis, J. J., ed.) pp. 233–261, Academic Press, New York.
46. Pearson, P. L. (1977) in *Molecular Structure of Human Chromosomes* (Yunis, J. J., ed.) pp. 267–293, Academic Press, New York.
47. Grouchy, J. de, Turleau, C. & Finaz, C. (1978) *Annu. Rev. Genet. 12*, 289–328.
48. Seuanez, H. N. (1979) *The Phylogeny of Human Chromosomes*, Springer, Berlin.
49. Yunis, J. J. (1976) *Science (Wash. DC) 191*, 1268–1270.
50. Belozersky, A. N. & Spirin, A. S. (1960) in *The Nucleic Acids* (Chargaff, E. & Davidson, J. N., eds) vol. 3, p. 147, Academic Press, New York.

G. Cuny, 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