## Organization of ribosomal RNA gene repeats of the mouse

R.Kominami, Y.Urano, Y.Mishima and M.Muramatsu

Department of Biochemistry, Cancer Institute, Japanese Foundation for Cancer Research, 1-37-1, Kami-Ikebukuro, Toshima-ku, Tokyo 170, Japan

Received 9 June 1981

### ABSTRACT

The organization of the ribosomal RNA (rRNA) genes of the mouse was determined by Southern blot hybridization using cloned rDNA fragments as probes, which could encompass the entire spacer region between two rRNA gene regions. The rRNA genes are organized into tandem repeats of nearly uniform length of about 44 kb. The heterogeneity detected in the nontranscribed spacer appears to be caused by its sequence rather than its length difference. At least three kinds of repetitive sequences are present in

the non-transcribed spacer region; two of them are located 13 kb upstream from the 5'-end of 18S RNA gene and the other located 1 to 4 kb downstream from the 3'-end of 28S RNA gene.

### INTRODUCTION

In eukaryotes, ribosomal RNA (rRNA) genes or rDNAs are organized as a family of tandemly repeated genes at the nucleolus organizer regions of chromosQmes. Each repeating unit contains one coding sequence for each of the rRNAs (18S, 5.8S and 28S RNA in this order) together with transcribed and non-transcribed spacer regions. The length of the tandem repeat varies widely among animals (1,2,3), and even within a single individual of some species (4,5). Although the variation results mainly from differences of spacer regions (6), one case of repeat unit heterogeneity has been described in Drosophila, where insertion sequences of variable lengths are present within some of the 28S RNA genes (4,7). The rDNAs of many eukaryotes show heterogeneity due to these structural features.

In higher vertebrates, however, the organization of rDNA is not fully characterized at present. Cory and Adams (2) studied the organization of the rDNA of the mouse, by means of Southern blot hybridization (8) using 18S, 5.8S and 28S RNA as probes, demonstrating that the size of the repeating unit for mouse rDNA was at least 44 kb. The upper limit, however, could not be defined clearly. Human rDNA was also partially characterized by blot hybridization and electron microscopy (9).

Recently, we have cloned three kinds of rDNA fragments from the mouse, which covers all of the rRNA gene regions (10, unpublished data). These DNA fragments provide good probes for studying the organization of rRNA gene repeats in more detail.

In this paper, we show that rDNA of the mouse consists largely of 44 kb repeats and the heterogeneity is due to its sequence rather than its length difference in the non-transcribed spacer region. This is in contrast with the length heterogeneity seen in Xenopus species, but rather agree with the findings with human rDNA (9).

In addition, we will show the presence of at least three kinds of repetitive sequences in the non-transcribed spacer region of rDNA; two of them located 13 kb upstream from the 5'-end of 18S RNA and the other located 1 to <sup>4</sup> kb downstream from the 3'-end of 28S RNA gene.

## MATERIALS AND METHODS

### Enzymes and Isotopes

Restriction enzymes were obtained from Takara Biochemicals, Kyoto. DNA polymerase I and T4-polynycleotide kinase were purchased from Boehringer Mannheim.  $[\alpha -^{32}P]$ dCTP and  $[\tilde{\Lambda} -^{32}P]$ ATP were obtained from Amersham Searle.

# Preparation of DNA and RNA

Livers were excised from mice and homogenized in a Potter homogenizer with 10 vol. of 0.02M Tris-HCl, pH7.6, 0.1M NaCl, 2mM MgCl<sub>2</sub> at  $0^{\circ}$ C. Nuclei were pelleted and resuspended in 10 vol. of 50mM Tris-HCl, pH7.6, 5mM EDTA, 0.1M NaCl, 0.5% SDS containing 200 pg/ml proteinase K. The viscous lysate was digested for 15 h at 37°C and extracted with phenol three times. The aqueous layer was dialyzed extensively against 0.05M Tris-HCl, pH7.6, 5mM EDTA, 0.01M NaCl at  $4^{\circ}$ C, and finally against 0.01M Tris-HCl, O.lmM EDTA (11).

Ribosomal 28S RNA was prepared from ddy mice after isolation of 60S ribosomal subunits with EDTA.

# Recombinant Plasmids Carrying rRNA Gene Fragments

EcoRI digestion of total mouse DNA yields three fragments containing rRNA gene region (See Fig. 1). The 15 kb fragment designated E-A contains a half of the 18S RNA gene and the upstream sequence (8). The 6.6 kb fragment (E-B) cloned by Tiemeier et al. (12) contains the 18S and 28S RNA sequences separated by the internal spacer.

Recently we have cloned a 6.6 kb DNA fragment carrying 3'-end of 28S RNA and the downstream sequence (E-C), which will be published elsewhere. The recombinant plasmids (pBR322) containing these fragments are propagated in E.coliXl776 under the Guidelines set by Japanese Government.

# Electrophoresis, Southern Blot Hybridization, and Filter Hybridization

Enzyme digestions were performed according to the conditions specified by the venders. Nick-translation was conducted as previously described (13). Specific activities ranged from  $10^7$  to  $10^8$  cpm/ug DNA. 28S RNA was labeled by  $\left[\right.\right.^{32}P$  | ATP and T4-polynuclotide kinase after partial alkali digestion. Specific activity was 1 to 5 x  $10^7$  cpm/uq RNA.

Agarose gel electrophoresis and Southern transfer were done by standard techniques (8) and filters were hybridized with RNA or DNA probe in 50% formamide, 0.72M NaCl, 0.04M Pipes-NaOH, pH6.8, lmM EDTA, 0.05% SDS at 37°C. When DNA probes were used, the solution contained 10 x Denhardt's solution (14).

After specified incubation times, filters were washed in 300 ml of 0.1 x SSC with occasional shaking at 50°C or 67°C.

 $[\tilde{\Lambda}^{-32}P]$  labeled HinfI fragments(~0.1µg) of EH2.0 kb DNA hybridized to lOpg of sonicated mouse liver DNA which had been bound to nitrocellulose filters, in the 50% formamide buffer described above. The hybridized fragments were identified by electrophoresis on an acrylamide-7M urea gel.

## RESULTS

# The Structure of the Repeat Unit of rRNA Gene

Fig. 1 illustrates three EcoRI fragments, E-A, E-B, and E-C, hybridizing to 18S and/or 28S rRNAs. These fragments are essentially equivalent to those detected in mouse DNA by Southern (8)



Fig. 1. A map of cleavage sites by various restriction enzymes within the mouse ribosomal RNA gene (rDNA).

ECORI, Hind II, and BamHI sites are shown by E, H, and B, respectively. Vertical lines indicating cleavage sites are Vertical lines indicating cleavage sites are divided into two groups. Shorter lines indicate partial digestion sites due probably to sequence heterogeneity or methylation and longer ones show complete cutting sites due to sequence homogeneity of repeating units.

and Cory and Adams (2), but the sizes of the fragments given in Fig. <sup>1</sup> were directly measured by an agarose gel electrophoresis of the cloned fragments.

Cory and Adams demonstrate that there is an additional spacer region of a mouse rDNA repeat between E-A and E-C fragments, which is 14 kb or longer.

We examined the length of this additional spacer segment. The 15 kb E-A DNA cloned in our laboratory contains a Hind $\overline{\mathbf{u}}$  $(H_2,$  see Fig 1.) site 2 kb downstream from the left end of the E-A fragment (10). When HindMI digests of mouse DNA are subjected to Southern blot hybridization, this EH 2.0 kb probe enables us to locate a HindII site in the upstream region. If the repeating unit is 44 kb, this upstream HindIII site should be the site on the  $18S$  RNA gene and the length between these two Hind $I\!I\!I$  sites should be 29 kb, which could be detected also by 28S RNA probe (See Fig. 1).

High molecular weight DNA prepared from C3H/He mouse liver was digested with EcoRI, HindII, and BamHI, and the digests were fractionated by gel electrophoresis. After transfer of the DNA fragments to nitrocellulose filters, they were hybridized with labeled 28S RNA, EH 2.0 kb fragments, and E-C (6.6 kb) fragments (Fig. 2-a).

In the HindmI digests, a 29 kb fragment  $(H_1/H_3)$  which hybridized to 28S RNA and E-C probes, hybridized also to EH 2.0 kb probe. These findings indicated that a HindIII site was present

E H B  $E$   $H$   $B$ EM <sup>B</sup> a  $24 Kb 24Kb$ to.A S  $\frac{9.5}{6.7}$  = 9.5<br>6.7  $43$ so  $43$  $2.3 -$ -2.3  $-2.0$ k 0 , . . . 0 ..l  $\blacksquare$ .... \_ \_ <sup>W</sup>  $\cdot$ ..] .FF  $\mathsf{C}$ I qwv  $\sim$  13 "a 28 S  $EH$  2.0

Fig. 2. DNAs from different mouse strains were digested with EcoRI, HindII, and BamHI, electrophoresed on agarose gels, and transferred to a nitrocellulose filter. a, C3H/He mouse DNA b, DDD mouse DNA c, BALB/c mouse DNA. Hybridization was carried out for 20 h with either labeled 28S RNA, EH 2.0 kb fragment, or E-C fragment as probe, washed with  $0.1 \times$  SSC at  $67^{\circ}$ C, and autoradiographed. E, H, and B denote digestions with EcoRI, Hind $\overline{\mathbf{u}}$ and BamHI, respectively. Size markers shown are derived from  $\lambda$ <br>phage DNA digested with HindII (23.7, 9.5, 6.7, 4.3, 2.3 and 2.0 kb).

at 41 kb (29 kb plus 12 kb) upstream from the 5'-end of the 18S RNA gene, the very Hind II site on the 18S RNA gene of the neighboring rRNA gene repeat. One rRNA gene repeat unit is thus inferred to be about 44 kb long. The other two bands of 14 kb and 1.8 kb hybridizing to EH 2.0 kb probe(Fig 2)were accounted for by the  $H_2/H_4$  and  $H_2/H_3$  fragments shown in Fig 1, respectively.

A similar hybridization experiment was carried out using the labeled E-C fragment as a probe, since this fragment contained two BamHI sites which could be utilized for determination of the repeat unit. BamHI digestion yielded a DNA fragment  $(B_5/B_6)$  of 32 kb hybridizing to E-C fragment, indicating that a BamHI site was located about 37 kb (4.7 kb plus 32 kb) downstream from the 3'-end of 28S RNA gene. This 32 kb fragment was also detected by the EH 2.0 kb probe. Thus, this 32 kb fragment begins at a site near the 5'-end of 18S RNA gene  $(B<sub>6</sub>)$  and extends through the so-called external transcribed spacer to the right side BamHI site  $(B_5)$  within E-C fragment. These results confirm those obtained by the Hind III digestion of DNA that the rDNA repeating unit of C3H/He mouse is about 44 kb. The other bands in the BamHI digests detected with 28S or E-C probe are explained by the map shown in Fig 1.

The data presented here agree with the map of the rDNA repeats as shown in Fig. 1.

In order to check this organization in mouse species, we next examined DNA from other strains of the mouse. Fig. 2-b and Fig. 2-c ahow the autoradiographs of DNA digests from DDD and BALB/c mice, respectively. HindJI digestion of each of mouse DNA generated a fragment of 29 kb which hybridized to all 28S RNA,E-C and EH 2.0 kb probes. Although the band of 29 kb detected in DDD mouse DNA was weak in intensity, it was probably due to the fact that most of the repeating units contained the second Hind $m$  site (H<sub>2</sub>) 14 kb upstream from the 5'-end of 18S RNA gene. In the BamHI digests of DDD and BALB/c mice, a band of 32 kb was detected by EH 2.0 kb probe as well as by E-C probe.

These findings obtained from DDD and BALB/c mice are consistent with the results obtained with C3H/He mice that one repeating unit of mouse rRNA gene is about 44 kb in length.

Heterogeneity in the non-transcribed spacer of rDNA

The DNA fragments detected in BamHI digests revealed that there was heterogeneity in the repeat units. 28S RNA probe hybridized to four fragments of 6.5kb, 5.Okb, 2.3kb and 1.3kb, 5.Okb and 1.3kb fragments of which were generated by the BamHI sites in the 18S and 28S RNA gene region (2,12). The intensities of the 6.5kb and 2,3kb bands varied in reciprocal manner in each mouse strain, suggesting that these two fragments were derived from separate repeating units differing in this region.

Fig. <sup>3</sup> shows two alternative models to explain this heterogeneity; i.e. length or sequence heterogeneity. In model A, some of the repeating units contain an inserted DNA sequence, located somewhere in the spacer region, or a DNA segment is deleted in

A. length heterogeneity



B. Sequence heterogenenity



Fig. 3. Order of BamHI fragments of mouse ribosomal gene repeats, showing two models to account for the heterogeneity within the repeating units.

some of the repeating units. In model B, some repeating units have two Bam sites,  $B_4$  and  $B_5$ , in the spacer region, but the remainder have only site  $B_5$  but not  $B_4$ .

In BamHI digests of cell DNA, E-C probe hybridized to  $B_A/B_E$ fragment of 4.2kb in addition to 2.3kb  $B_3/B_4$  and 6.5kb  $B_3/B_5$ fragments and the large 32kb DNA. The detection of  $B_A/B_F$  fragment hybridizing to E-C probe supports model B. Besides, E-C fragment (6.6kb) cloned by us contained two BamHI sites corresponding to  $B_A$  and  $B_E$  sites, which could not be accounted for by model A. Therefore, it seems that the heterogeneity in the spacer DNA is due to its sequence rather than its length differrence.

Two closely located HindIII sites  $(H_2$  and  $H_3)$  at about 13kb upstream from the 5'-end of 18S RNA gene are also interpreted in the same way (Cory and Adams, 2). In fact, the  $H_2/H_3$  fragment of 1.8kb was detected by hybridizing to EH2.Okb probe, demonstrating that sequence heterogeneity existed in this region. The Presence of the Three Different Repetitive Sequences in rDNA

In the course of this blot hybridization, we discovered the presence of repetitive sequences within the E-A and E-C fragments.

When 28S RNA and E-B DNA probes were used for hybridization, autoradiograms showed always patterns with clear bands (Data for E-B DNA probe are not shown). EH 2.0 kb and E-C DNA probes, however, yielded smears which seemed to be a high background but was in fact a continuum of hybridizing bands. Besides, in EcoRI digests, some extra-bands were seen in addition to the expected bands of the rDNA, which could not be explained from the structure of rDNA. The patterns of the extra-bands were different among these three mouse strains (Fig. 2). These observations suggested that some sequences present in EH 2.0 kb and E-C DNA were not unique to the rRNA gene family but were present in other parts of the genome.

In order to substantiate this interpretation, reciplocal hybridization experiments were carried out (15). Total mouse DNA was sheared, labeled with  $32P$ , and hybridized to each of the restriction enzyme digests of the E-A and E-C DNAs. E-A fragment was digested with either SalI or HindII and subjected to blot hybridization under a low Cot condition (Fig. 4-a). Five frag-



Fig. 4. Mouse liver DNA was sheared to 500-700 bp by sonication, labeled with 32p by nick-translation, and hybridized to filters carrying E-A and E-C fragments.

The E-A fragment was digested with either SalI or Hind $I\!I\!I$ , electrophoresed in a 1% agarose gel, stained with ethidium bromide (A-a) and transferred to a nitrocellulose filter.

Hybridization was carried out in 50% formamide et 37°C for 4 h as described in Materials and Methods. The filter was washed with 0.2 x SSC at 50°C, and autoradiographed (A-b).

pBR322 carrying E-C fragment at the EcoRI site was digested with SalI plus PstI, electrophoresed with HindII digests fo SV40 DNA, and subjected to Southern blot hybridization described above. B-a and B-b show the ethidium bromide staining pattern and

the autoradiogram, respectively.

S, H and P denote digestions with SalI, HindIII and PstI, respectively.

ments (8.0, 3.2, 1.9, 1.2, and 0.6 kb) and two fragments (13 and 2 kb) were generated by SalI and Hind $m$ , respectively. The 8.0 kb and the 2.0 kb fragments were detected in SalI and Hind $I\!I\!I$ digest, respectively. This indicates that repetitive sequence(s) are located within EcoRI/HindII 2.0 kb region.

Next, pBR322 carrying E-C DNA was digested with SalI plus PstI to yield four fragments (3.1, 2.0, 1.2, and 1.1 kb). Upon blotting, SalI/PstI 3.1 kb DNA was hybridized with labeled total cell DNA (Fig. 4-b), Under the hybridization conditions used, the rRNA gene sequences with a few hundred repetitions were not detected. Incidentally, there was no cross-hybridization between this 3.1 kb fragment and the EH 2.0 kb fragment (data not shown).

To further confirm the presence of repetitive sequences in the rRNA gene repeats, liquid hybridization studies were done with nick-translated EH 2.0 kb fragment and BamHI-digested fragment of rRNA gene region  $(B_2/B_3)$  as tracers and sheared total cell DNA as driver. As shown in Fig. 5 whereas 28S RNA sequence



Fig. 5. Analysis of the repetition frequency of the EH 2.0 kb DNA. Sonicated mouse DNA was incubated at a concentration of 16  $\mu$ g/ml with nick-translated EH 2.0 kb DNA ( $\circ$ -----  $\circ$ ). When BamHIdigested fragment of rRNA gene region  $(B_2/B_2)$  was used as tracer, <sup>200</sup> pg/ml or cell DNA was subjected to hybrldization () Duplex formation was assayed by hydroxyapatite chromatography.

with a repetition frequency of a few hundreds reannealed at  $\cot^1$ of about 5, EH 2.0 kb fragments reassociated at a Cot<sub>2</sub> of 0.1. We estimate from this analysis that there are about 10,000 to 30,000 copies of the repetitive sequence of EH 2.0 kb DNA in mouse genome. These results demonstrate clearly that some sequence within EH 2.0 kb DNA is repeated many times elsewhere in the genome.

In order to locate the repetitive sequence within EH 2.0 kb DNA. We constructed a finer map and determined which fragments could hybridize with sonicated cellular DNA bound to nitrocellulose filters more rapidly than the other fragments. EH 2.0 kb DNA was sub-cloned in a pBR322 vector and the locations of the cleavage sites were determined by double digestion with two enzymes and by partial digestion of the terminally-labeled fragment with an enzyme(16). The results are shown in Fig. 6-a. Hinf <sup>I</sup> fragments of EH 2.0 kb DNA were 5'-terminally labeled with  $[\delta -^{32}P]$  ATP and T4-polynucleotide kinase. The labeled Hinf I fragments hybridized with sonicated cellular DNA fixed to nitrocellulose filters and the hybridized fragments were then eluted with a buffer at  $100^{\circ}$ C. The fragments were identified by electrophoresis on an acrylamide-7M urea gel after alkali denaturation. As shown in Fig. 6-b, the recoveries of the two bands of 460 bp and 200 bp were much higher than the other bands. Under these electrophoretic conditions, 490 bp and 460 bp fragments exibited rather broad bands probably because of strand separation. 280 bp fragment was recovered more than 740 bp and 490 bp fragments. It remains unclear whether this 280 bp fragment contains a repetitive sequence. These results indicate that repetitive sequences are located within two segments of 460 bp and 200 bp defined by Hinf I cutting sites. These two fragments did not hybridize with each other (Data not shown), demonstrating that two kinds of repetitive sequences were present in EH 2.0 kb DNA.

## DISCUSSION

The present work demonstrates that the rRNA genes of the mouse are organized into tandem arrays of repeats of about 44 kb long. These results are obtained by the blot hybridization using



Fig. 6  $\,$  EH 2.0 kb, DNA was digested with Hinf I and 5'-terminally labeled with [ $\delta^{-3/2}$ P]ATP and polynucleotide kinase after bacterial alkaline phosphatase treatment. The labeled fragments were hybridized with sonicated cellular DNA fixed to nitrocellulose filters and the hybridized fragments were then eluted with 50mM Tris HC1, pH 7.6, lmM EDTA and 0.1% sodium dodecyl sulfate at 100°C. The recovered fragments were electrophoresed on a 5% acrylamide-7M urea gel after alkali denaturation.

- A. Cleayage sites of EH 2.0 kb DNA<br>B. a) <sup>32</sup>P-HinfI fragments before hybridization
	- b) Fragments hybridized to cell DNA on nitrocellulose filter.

cloned fragments as probes which could encompass the entire nontranscribed spacer region between two gene regions.

This long repeat unit of the mouse is <sup>3</sup> to <sup>4</sup> times larger than well-known repeats of Xenopus rDNA. DNAs from three different strains of mice were analyzed and the equivalent results were obtained.

Heterogeneity in rDNA repeat length is observed in most of the higher eukaryotes examined so far (3,4,5,6,17,18), although uniform repeating units are reported in yeast, tetrahymena and silkworm (19,20,21). Human rDNA appears largely homogeneous in

length (9).

The heterogeneity of Xenopus leavis rDNA apparently results from variability in the number of copies of a tandemly repeated short sequence within the non-transcribed spacer region (3,5). unequal crossing-over within these regions which may account for homogeneity of rRNA gene sequences within species probably gives rise to the length heterogeneity (3). Arnheim and Kuehn (22) have reported that this type of length variation also occurs in the spacer region of the mouse rDNA. They found that HindII digests of mouse DNA (including BALB/c mice) revealed heterogeneity in the size of fragments homologous to a 1.7 kb SalI fragment of mouse ribosomal spacer region located at 5 kb upstream from the 5'-end of 18S RNA gene. This length heterogeneity was supposed to be caused by different numbers of copies of PVu II subrepeating units. On the other hand, the presence of the major EcoRI fragment of approximately 15 kb covering the 1.7 kb SalI fragment and hybridizing to 18S RNA is confirmed by different authors especially for BALB/c mouse (2,10,23). Therefore, the length variation, if it exists due to the different numbers of Pvu II repeats, must be small.

The heterogeneity in mouse rDNA revealed by HindII and Bam HI digestions cannot be accounted for by the presence of insertions of different lengths. It is better explained by a sequence change which has altered the number of HindII or Bam HI sites in a set of rDNA repeats.

Most, if not all, rDNA repeating units have rather uniform lengh of 44 kb, although small size variation, within a few kilobases maybe present which would not be detectable in these fragments when examined by Southern blotting experiments. The length of mouse rDNA repeating units is thus rather uniform, as suggested for human rDNA (9) and in contrast with the heterogeneous repeats of Xenopus rDNA. This may have some implication in the evolution and conservation of the tandemly repeated rRNA genes in mammals or higher vertebrates.

In the course of this study, we found repetitive sequences at three loci in the non-transcribed spacer region. They appear to be scattered in other parts of the mouse genome. One of the sequences located 13 kb upstream from the 5'-end of the 18S RNA

gene probably correspond to the sequence recently found by Arnheim et sl. (24). Whatever the role of these repetitive sequences may be, it is of interest to determine whether every repeating unit of rDNA contains these repetitive sequences at the same loci.

Since Hind III digestion of mouse DNAs yielded 1.8 kb fragments  $(H_2/H_3)$  hybridizing to EH 2.0 kb probe, and the repetitive sequences were located within this 1,8 kb sequence, we infer that most, if not all, of rDNA repeating units contain this repetitive sequence. Namely, some of the repetitive sequence families found in the EH 2.0 kb region are present in a certain region of another tandemly repeated sequence family, i.e. rDNA, although these repetitive sequence families may be generated and conserved in different ways.

Two posibilities may explain this; one is that there is a specific site in rDNA repeats to accept the repetitive DNA sequence and the other is that an rDNA repeat which had carried the repetitive sequence was amplified in mouse genome.

EcoRI digestion of mouse DNA generated, in addition to smears, fragments of several discrete sizes hybridizing to EH 2.0 kb probe, suggesting that some sequences present in other parts of the mouse genome are embedded in certain flanking sequences which are also repeated (further data to be published). It is noteworthy that the bands vary among different strains of the mouse, indicating that the emergence of these discrete structures containing the repetitive sequence is a relatively new event and the mode and units of amplification of the sequence are different among the strains.

### AKNOWLEDGEMENTS

We thank Dr. P. Leder, National Institutes of Health for the generous gift of a phage carrying the 6.6 kb rDNA (E-B). This work was supported in part by grants from the Ministry of Education, Science and Culture of Japan. Y.M. was a recipient of a fellowship from Japan Society for the Promotion of Science.

### REFERENCES

Tartof, K.D. and David, I.B. (1976) Nature 263, 27-30

