Volume 11 Number 18 1983 Nucleic Acids Research

Genomic organization of rat rDNA

Dona M.Chikaraishi*, Lisa Buchanan*, Kathleen J.Danna+ and Christina A.Harrington+*

*Division of Molecular and Cellular Biology, National Jewish Hospital/National Asthma Center, 3800 East Colfax Avenue, Denver, CO 80206, and +Department of Molecular, Cellular and Developmental Biology, University of Colorado, Campus Box 347, Boulder, CO 80309, USA

Received 14 March 1983; Revised and Accepted 12 August 1983

ABSTRACT

A detailed restriction map was determined for a 10.9 KB region that contains the initiation site for 45S pre-rRNA and the first 1.7 KB of the 18S rRNA coding region. When the restriction pattern of the cloned rDNA was compared with that of total rat DNA, the rDNA regions of both Sprague-Dawley and BD-9 rats were identical to each other and to that of the cloned rDNA. However, both strains exhibit a major polymorphism consisting of an insertion of 0.9 KB of DNA in the nontranscribed spacer between 0.29 KB and 1.8 KB upstream from the 45S RNA initiation site. This region consists of tandem repeats approximately 130 base pairs in length. These repeats contain large poly T tracts and are similar in sequence to analogous elements 5′ to the origin of mouse rRNA transcription. Regions containing highly repetitious DNA sequences were located at sites 2.8 KB and 4.3 KB upstream from the initiation site. The repetitive sequence at 2.8 KB from the initiation site anneal to a known Alu-equivalent type 2 sequence derived from the second intron of the rat growth hormone gene.

INTRODUCTION

There are about 160-200 copies of rat ribosomal DNA genes (rDNA) per haploid genome (1). They are arranged in tandem clusters (2) which can be identified cytologically by their ability to organize a nucleolus in interphase nuclei and by their localization to sites of secondary constriction in metaphase chromosomes. Three chromosomes (3,11 and 12) each carry about 60 gene copies (3). By visualizing R-loops consisting of 18S and 28S rRNA annealed to long fragments of genomic DNA, Stumph, Wu and Bonner (2) estimated the rDNA repeat unit to be 37.2 ∓1.3 KB in size. All the repeat units were approximately the same size, but a certain degree of length heterogeneity, possibly up to 4 to 5 KB, could not be ruled out by their data.

As in other eukaryotes, each rat rDNA repeat is composed of: 1) a nontranscribed spacer, which in mammals comprises two-thirds the length of the total repeat; 2) an external transcribed spacer, about 4.4 KB in mammals, which extends from the start site of pre-rRNA synthesis to the beginning of the 18S rRNA gene; 3) one coding region for each mature rRNA (18S, 5.8S and

28S); and 4) internal transcribed spacers separating each of the mature rRNAs. In higher eukaryotes, 5S rRNA is encoded by another region of the genome.

In order to examine the genomic organization of the rDNA repeat units, we selected a λCharon 4A recombinant phage which contained transcribed rDNA sequences, determined a detailed restriction enzyme map for a 10.9 KB fragment that contains the 45S RNA transcriptional initiation site, and compared the restriction pattern of the cloned sequence to that of total rat DNA by Southern blot analysis. In overall structure, the rat rDNA repeat is very similar to that reported for other mammals (4-15), having a conserved transcribed region as well as polymorphic and highly repetitious sequences in the non-transcribed spacer.

MATERIALS AND METHODS

Materials

All restriction enzymes were purchased from Bethesda Research Laboratories and used according to the vendor's recommendations, except that two to four fold more enzyme was used for digestion of total rat DNA. Bacterial alkaline phosphatase and nuclease S1 were purchased from P-L Biochemical or Bethesda Research Laboratories. Micrococcal nuclease was purchased from Sigma, and DNA polymerase was supplied by Boehringer-Mannheim.

Phage and plasmid stocks

The rat genomic DNA library in λCharon 4A was a gift of Drs. R.B. Wallace, T. Sargent and J. Bonner. It was constructed from a partial Eco RI digest of liver DNA from a single male Sprague-Dawley rat. λB3 was a recombinant selected from that library by Mr. James Beeson and Dr. Noboru Sueoka and contains transcribed highly repetitious DNA sequences. λ63, a gift of Dr. Warren Jelinek, contains a Chinese hamster ovary DNA insert homologous to the rodent Alu-equivalent type 1 sequence (16,17). A pBR322 plasmid containing a rat Alu-equivalent type 2 sequence from the second intron of the rat growth hormone (18) was a kind gift of Dr. Barbara Cordell.

λDNA was prepared as described by Kao et al. (20).

Single-stranded M13 DNA was purified from phage lysates as described by Harrington and Chikaraishi (21).

Cellular DNA was prepared from the liver of one adult male Sprague-Dawley rat and one adult male BD-9 rat by phenol extraction of isolated nuclei.

Cloning and Selection of Recombinants

λCharon 4A recombinants containing rDNA inserts were selected by plaque hybridization (22) with ^{32}P labeled 18S and 28S rRNAs and subcloned as described (21) and as shown in Figure 1.

Labeling of nucleic acids

RNA. RNA was labeled with [γ-^{32}P]ATP [9000 Ci/mmole] prepared by a modification of Johnson and Walseth (23,24)], and polynucleotide kinase according to Derman et al. (24).

Micrococcal nuclease digestion and labeling of M13 single-strand DNA with polynucleotide kinase. We have developed a simple and rapid method for preparing ^{32}P-labeled, single-strand M13 DNA probes by phosphorylation at the 5'-hydroxyl group generated by micrococcal nuclease. Purified single-strand M13 DNA (0.1 mg/ml) was digested for 3 minutes on ice with micrococcal nuclease at a final concentration of 6 µg/ml in 5 mM sodium acetate pH 5.0, 1 mM $CaCl_2$. Stock solutions of micrococcal nuclease at 1 mg/ml in H_2O should be stored frozen, thawed, and refrozen only three or four times, because the activity of the enzyme decreases with repeated freeze-thaw cycles. Since the optimal pH for nuclease activity is between pH 9.2 and 9.6, the low pH of the incubation limits the enzyme activity such that DNA fragments between 50 and 100 nucleotides in length are generated. The nuclease was stopped by addition of EGTA (ethylene glycol-bis-(β-aminoethyl ether)N, N'-tetraacetic acid) to 8 mM to chelate Ca^{+2} and the reaction was boiled for 5 minutes. Any insoluble, denatured enzyme was removed by centrifugation in a microfuge for 5 minutes at 4°C. Generally, 20 µl of the supernatant containing the DNA fragments (2 µg) were then added to a microfuge tube containing 1 mCi of lyophilized [γ-^{32}P]ATP (9000 Ci/mmole). Buffer was adjusted to 50 mM Tris-HCl, pH 7.5, 18 mM $MgCl_2$, 5 mM dithiothreitol, 0.1 mM spermidine, 0.1 mM EDTA and 2 to 10 units of polynucleotide kinase added. The reaction was incubated for 45 minutes at 37°C and stopped with the addition of 100 µl of 2.5 M NH_4 acetate, 50 mM EDTA. The labeled DNA fragments were separated from free [γ-^{32}P]ATP by gel filtration on Sephadex G-50 equilibrated in 0.4 M NaCl, 10 mM Tris-HCl pH 7.5, 1 mM EDTA, 0.1% SDS. Carrier E. coli tRNA (20 µg) was added to the labeled DNA from the excluded peak and the nucleic acid was precipitated with ethanol. The ethanol precipitate was dissolved in 200 µl of 0.1 M sodium acetate pH 6.0 and extracted with Tris-saturated phenol and the phenol removed by ether extraction or by ethanol precipitation of the nucleic acids. The specific activity of the DNA was between 0.5 and 2 X 10^8 cpm/ µg.

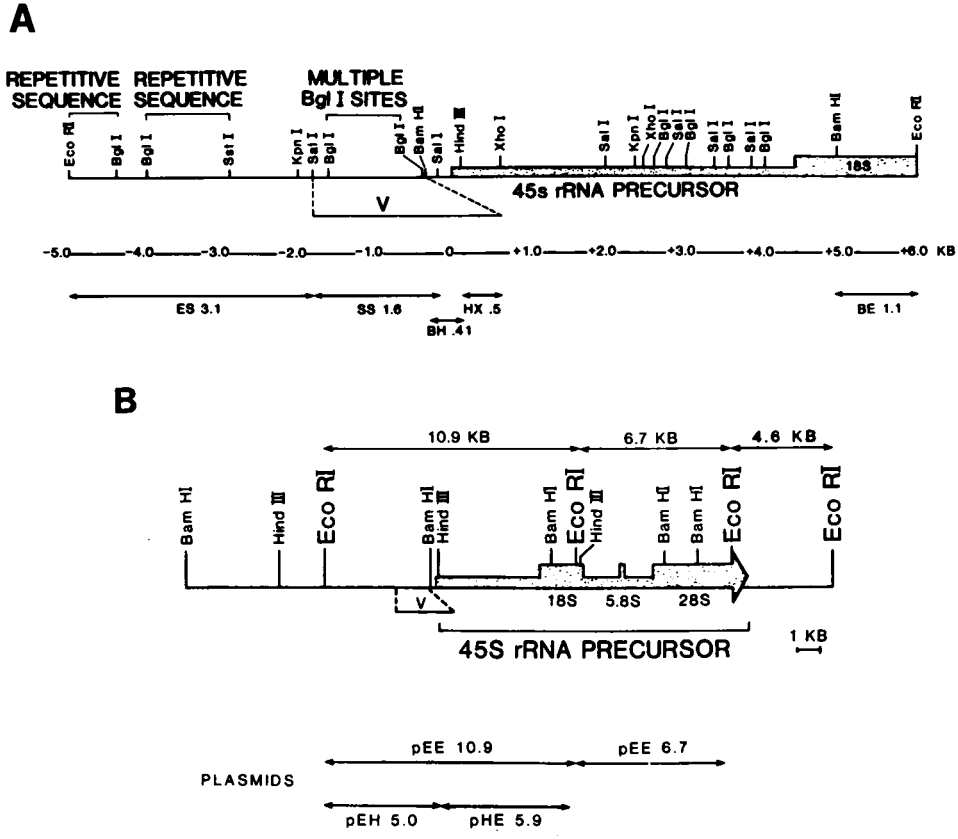


Figure 1.

A) Restriction map of the 10.9 KB EcoRI fragment containing the 45S RNA initiation site. The initiation site is designated as 0 on the map; sites in the transcribed region are shown as positive distances from the initiation site; those upstream from the initiation site are assigned negative distances. The position of the rDNA sequences contained in M13 recombinant clones is shown below the map.

B) Restriction map of the transcribed region of rat rDNA. The location of restriction sites was determined by standard mapping techniques on the cloned rDNA and on total genomic DNA as described in Materials and Methods. The position of EcoRI fragments (10.9 KB, 6.7 KB and 4.6 KB) is shown above the map and the location of sequences contained in recombinant plasmids pEE10.9, pEE6.7, pEH5.0 and pHE5.9 is shown below the map. The stippled bar represents the transcribed region, with the arrow indicating the direction of transcription. The position of the 45S and 18S RNAs were determined by transcription mapping as described in the text; the location of 5.8S and 28S RNAs should only be considered approximate and are based on the R-looping data of Erickson et al. (13) and Fuke et al. (37). "V" signifies an 0.9 KB insertion present in about half of the Sprague-Dawley rDNA repeats and in about 60-70% of the BD-9 rDNA repeats.

Restriction Mapping

Restriction mapping was performed by single and double digestions with various enzymes. Mapping of the SS1.6 fragment to delineate its repeat structure was accomplished by partial digestion of singly-end labeled purified fragment DNA. The right end of the fragment was labeled at the Bam Hl site (-.28) and recut at the Sal I site (-1.76). Mapping from the left end was accomplished using a 1.4 KB fragment labeled at the Sal I site (-1.76) and recut at the Alu I site (-.23). The partial digestion products were separated on 6% polyacrylamide-8M urea gels.

Southern blots and hybridization

10 µg of digested liver DNA or 0.02 µg of digested plasmid DNA were electrophoresed on 1.1% agarose gels (22 cm x 14 cm x 3 mm) in 40 mM Tris, 20 mM sodium acetate, pH 8.3, 2 mM EDTA for 5 hours at 125 V. The DNA was transferred to nitrocellulose filters by the method of Southern (25) and hybridized according to the procedure of Jeffreys and Flavell (26).

DNA sequencing

M13 recombinants were sequenced as described previously (21).

RESULTS

Cloning of rDNA Sequences

To identify the size of rDNA containing fragments, Sprague-Dawley and BD-9 rat liver DNAs were digested with Eco RI, electrophoresed on 1.1% agarose gels, and then blotted onto nitrocellulose by the method of Southern (25). When hybridized with labeled liver RNA (90% of which is 18S and 28S rRNA), three major bands at 11.4, 6.7 and 4.6 KB were observed (Figure 2). (The largest band is actually a doublet of 10.9 and 11.8 kb which did not resolve well). The same results were obtained by Tantravahi et. al. (27) with DNA derived from a rat cell line, K22, and by Braga et al. (4) with rat liver DNA. In experiments not shown, we have determined that both bands in the 10.9-11.8 KB doublet contain sequences homologous to 18S rRNA whereas the 6.7 KB fragment contains both 18S and 28S rRNA homologous sequences, and the 4.6 KB band bears only 28S sequences.

We selected λCharon 4A recombinant phage containing rDNA sequences from the genomic library of Sargent et al. (28). One clone contained the 10.9 and 6.7 KB Eco RI fragments which were cloned separately into pBR322. The 10.9 KB fragment contains a single Hind III site which was used to clone the two halves of the fragment via the Eco RI and Hind III sites in the rDNA and in

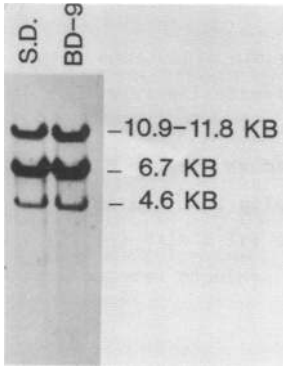


Figure 2. Southern blot of Eco RI digested rat DNA probed with [^{32}P]-labelled rat liver RNA. 10µg of Sprague-Dawley or BD-9 liver DNA were digested with Eco RI, electrophoresed, and transferred to nitrocellulose as described in Materials and Methods. The filter was annealed with ^{32}P-labelled rat liver RNA. The sizes of the autoradiographic bands were estimated from markers run on the same gel.

the pBR322. In addition, six smaller regions of rDNA were recloned into the single-strand vector M13 as shown in Figure 1.

A restriction map of the 10.9 KB fragment is shown in Figure 1. The initiation site for 45S pre-rRNA synthesis (21,29,30) is designated as 0 on the map. Numbering from a physiologically important site may be preferable to numbering from a given restriction site, which may vary among different repeat units, and will certainly vary between species.

The location of the 18S rRNA sequences was determined by S1 transcription mapping which indicated that the 18S coding region begins at +4.4 KB on the map in Figure 1, and defines the length of the external transcribed spacer to be 4.4 KB in rat (data not shown).

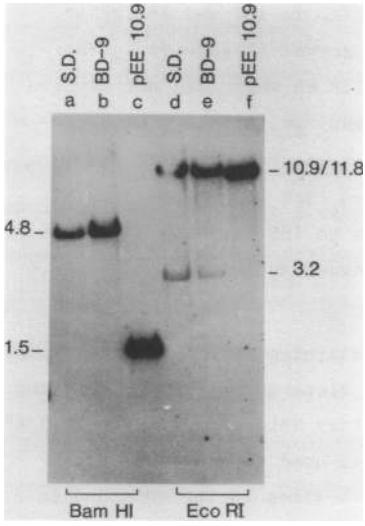


Figure 3. Southern blot of restricted rat DNA probed with BE 1.1 (+4.9 to +6.0). Sprague-Dawley genomic DNA (lanes a and d), BD-9 genomic DNA (lanes b and e), or clone pEE10.9 DNA (lanes c and f) was digested with Bam HI (lanes a, b, and c) or Eco RI (lanes d, e, and f), blotted, and probed as described in Materials and Methods. The 1.5 KB band (lane c) from Bam HI digestion of pEE10.9 results from rDNA sequences from +4.9 to +6.0 KB which are attached to pBR322 sequences from the Eco RI site to the Bam HI site.

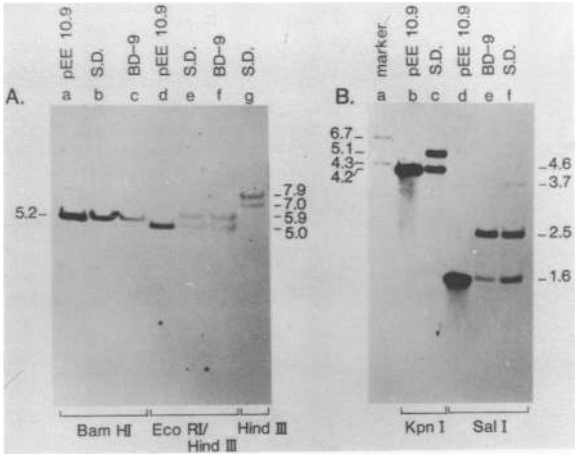


Figure 4.

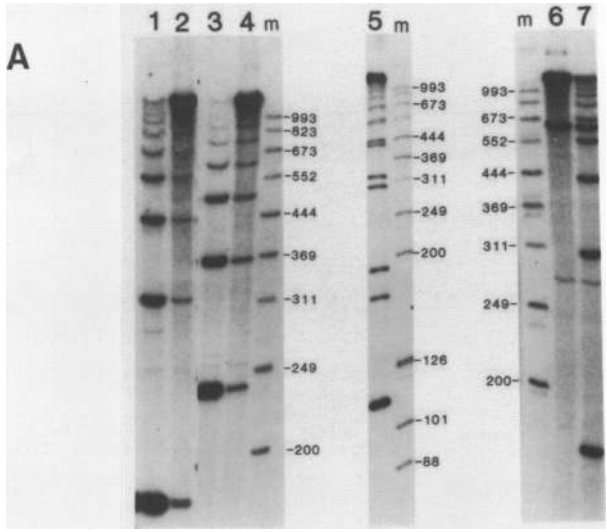
A) Southern blot of restricted DNA probed with BH.41 (-.28 to + .12). Sprague-Dawley genomic DNA (lanes b, e, and g), BD-9 genomic DNA (lanes c and f) or plasmid pEE10.9 DNA (lanes a and d) was digested with Bam HI (lanes a, b, and c), with Eco RI and Hind III (lanes d, e, and f), or with Hind III alone (lane g).

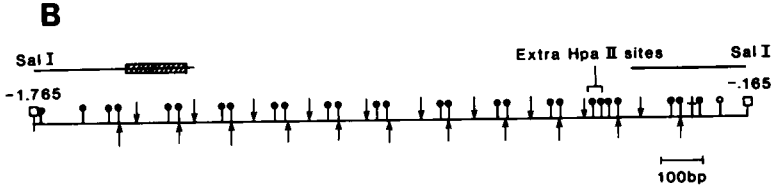
B) Southern blot of restricted DNA probed with SS1.6 (-1.8 to -.16). Sprague-Dawley genomic DNA (lanes c and f), BD-9 genomic DNA (lane e) or plasmid pEE10.9 (lanes b and d) was digested with Kpn I (lanes b and c) or Sal I (lanes d, e, and f). The minor 3.7 KB and 4.6 KB bands in 4B (lane f) are observed in other preparations of Sal I digested Sprague-Dawley DNA.

Genomic Organization of rDNA

To assess how similar the 160-200 different rDNA repeats are to each other and to the cloned rDNA sequence, Southern blot analyses were performed on liver DNA cleaved with different restriction enzymes and probed with various labeled rDNA cloned fragments.

The results showed that the transcribed region of the rDNA repeats was conserved with respect to the mapped restriction sites. Figure 3, lanes a and b, shows a typical result in which DNA from both strains exhibit one band homologous to the probe, in this case BE1.1 (+4.9 to +6.0), which is within the 18S rRNA coding region. The 4.8 KB band represents DNA sequences from +4.9 to +9.7 on the map in Figure 1 and demonstrates that the majority of rDNA repeats contain Bam HI sites at +4.9 and +9.7. The only instance in which minor bands generated from sequences within the transcribed region were reproducibly observed was when Eco RI cleaved DNA was probed with BE1.1 (+4.9 to +6.0), as shown in Figure 3, lanes d and e. In addition to the doublet band





C

```
-1765
GTCGACCAGA AGGCTTGGGC CCCGTCCCGG ATTTTTTTTT TTTTTTCCA ATTTTTTTTT TTTAAATTTT
TTATTTTATT TTATTTTATT TTTTTTATTT CTTCCCTCCT CTCCGGTCCC AGCCAGCGGG ACCATAGACA
CTTTTGAGGC CGAGAGGGTG CTGTGTCCCC GGCCTCCGAG CCGAGGTGGC CCGGCTAGGG GCGCCACGGA
GTTTTTTTTT TTTTTTTTCT TTTCCAGGAG TCCCTTCGGT CCCAGCCAGC GGGAGCCATA GACACTTTTG
AGGCCGAGAG GGTGCTGTGT CCCCGGCCTC CGAGCCGAGG TGGCCCGGCT AGGGGCGCCA cGGATTTTTT
//....../ CTgTGTCCCC GGCCTCCGAG CCGAgGTGGC CCGGCTAGTG GCGCCACGGA TTTTTTTTTT
TTCTTTTCCA GGAGTCCTTC GGTCCCTGAT AGCGGGAGCC TGGACCTCTG AGGCCGAGAG GGTGCTGTGT
CCCCGGCCTC CGAGCCGAGG TGGCCCGGCT AGGGGGCGCC TCCGAGACTT TATTTTTTCC AGGATCCTCC
CCGGTCCCTG CCAGCGGGAG CATGGACTTC TGAGGCCGAG GGGAAGCTGT GTTCCAGGCT ATCTACCATG
                                  -165
GCCTCCTCGG TCTGTGAGCA CTCAGGGTTC TAAGGTCGAC
```

at 10.9 - 11.8 KB, there is a faint doublet at 3.2 KB. By densitometry, these minor bands were estimated to be about 4% of the intensity of the major band.

Length Polymorphism in the Non-transcribed Spacer, Region -1.8 to -0.28

When rDNA sequences outside the transcribed region were investigated, it was found that about half of the Sprague-Dawley and 60-70% of the BD-9 rDNA repeats contain an insertion of 0.9 KB just upstream from the initiation site of rDNA transcription (designated as "V" for variable in Figure 1). Figure 4A is an autoradiogram of a Southern blot probed with BH.41, a region from -0.28 to +0.125. When rat DNA was digested with Hind III or a mixture of Hind III and Eco RI and probed with BH.41, two bands result (Figure 4A, lanes e,f,g), suggesting that the rDNA repeat in the +0.125 to -4.9 region occurs in two discrete lengths. Figure 4A (lane d) shows that the cloned rDNA is representative of the smaller repeat length. To localize the insertion, we probed Bam HI digested DNA with the BH.41 probe and detected no length polymorphism (Figure 4A, lane b and c) suggesting that the insertion is upstream of the Bam HI site located at -0.28. However, when DNA was cleaved with Sal I or Kpn I and probed with SS1.6 (-1.8 to -0.16) two bands were detected (Figure 4B), with the cloned rDNA again representative of the small repeat (lane d). This positions the insertion between -1.8 and 0.28. In all cases, the difference in length between the two rDNA repeat units is 0.9 KB.

This polymorphic region (-1.8 to -0.16) (V region) is composed of small repeating units 130 nucleotides in length. When the SS1.6 fragment is labeled at only one end (either at -1.76 at the Sal site or at -.28 at the Bam HI site) and partially digested with either Bgl I, Hinf I or Hpa II, a ladder of

Figure 5.

A) Mapping of the subunit repeat in the SS1.6 fragment by partial digestion with Bgl I, Hinf I and Hpa II. The SS1.6 fragment was singly end labeled at the Sal I site (-1.76), gel purified and digested with Bgl I at 0.15 units/ml (lane 1), and 0.015 units/ml (lane 2); with Hinf I at 0.3 units/ml (lane 3) and at 0.03 units/ml (lane 4); with Hpa II at 1 unit/ml (lane 5). A 1.4 KB BamHI-Sal I fragment was singly end labeled at the BamHI site (-0.28) gel purified and digested with .015 units/ml of Bgl I (lane 7). Lane 6 shows the undigested fragment; the bands at 650 and 275 bp are contaminating fragments from pBR322 and are not digested with Bgl I. Lanes "m" are SV40 Eco RII fragments; sizes are given in base pairs.

B) Restriction map of the SS1.6 fragment. The lines indicate sequenced regions shown below. The box indicates the subunit repeat whose sequence is underlined in C. Hpa II, Hinf I, Bgl I, Bam HI, Alu I, Sal I.

C) Nucleotide sequence of the regions indicated in B. Underlined sequence corresponds to a representative single repeat unit. The identity of three of the bases reported here is questionable and they are indicated in lower case letters.

fragments is generated consistent with an internal nucleotide subunit repeat (Figure 5A).

The cloned SS1.6 fragment contains eleven tandem repeats (Figure 5B) plus an incomplete repeat extending to about -.22. Figure 5C shows the sequence of approximately 300 bases at each end of the SS1.6 fragment with the region of one repeat underlined. (The underlined repeat corresponds to a consensus sequence for the 3.5 repeats that have been sequenced). Based on the position of the restriction sites and the available sequence data the repeats are quite similar but not identical; 79% of the bases in the sequenced repeats are the same. As in the mouse (31), the repeat units are separated by variable length tracts of T's.

Two Different, Highly Repetitious DNA Sequences Reside in the Non-transcribed Spacer Regions -4.9 to -4.3 and -3.9 to -2.8

When ES3.1 DNA (-4.9 to -1.8) was used to probe Southern blots, strong bands were observed corresponding to the fragment lengths predicted by the restriction pattern of the cloned rDNA (Figure 6). In addition, ES3.1 probes gave an intense smear of radioactivity that followed the ethidium bromide stained DNA (compare Figures 6A and 6B). This smear was detected when the filters were washed under the standard condition of 0.5 X SSC at 60°C but was more pronounced if the stringency of washing was reduced by eliminating the 0.5 X SSC salt wash. These results suggest that ES3.1 may contain repetitive sequences found elsewhere in the rat genome.

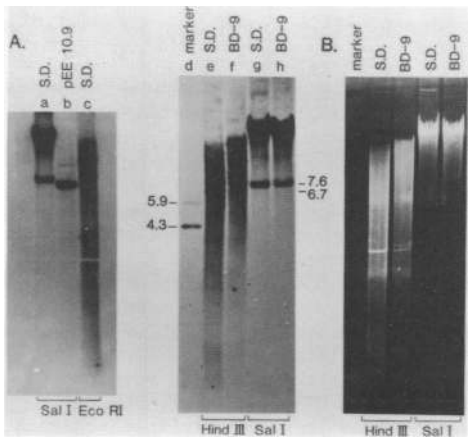


Figure 6.
A) Southern blot of restricted DNAs probed with ES3.1 (-4.9 to -1.8). Sprague-Dawley genomic DNA (lanes a, c, e, and g), BD-9 genomic DNA (lands f and h), or plasmid pEE10.9 DNA (lane b) was digested with Sal I (lanes a, b, g and h), with Eco RI (lane c) or with Hind III (lanes e and f). The digested DNAs were processed as described in Materials and Methods, except that the filter bearing lanes a, b, c and h was not washed in 0.5 x SSC, 0.1% SDS.
B) The agarose gel of digested DNAs used for Southern blot of lanes d, e, f, g, and h was stained with ethidium bromide and photographed under u.v. illumination.

These highly repetitious sequences were compared to those found in two other rat DNA clones. One clone contains the repeat sequence from the second intron of the rat growth hormone gene which by sequence analysis belongs to the rodent Alu-equivalent type 2 family described by Haynes and Jelinek (17). The other clone, λB3 (J. Beeson and N. Sueoka, personal communication), contains two different repeat elements, one Alu type 2 and one non-Alu related sequence.

When DNA from -4.9 to + 0.125, was digested with Bgl I (see Figure 7A for restriction map) and probed with nick-translated λB3 DNA, two cross-hybridizing rDNA bands at 1.7 and 2.4 KB resulted (Figure 7B, lane c). The 1.7 KB band contained 0.6 KB of rDNA from -4.9 to -4.3, the rest of the length being pBR322 DNA. The 2.4 KB band contained rDNA between -3.9 and -1.5 KB. When the Bgl I digest was further cut with Sst I, the 2.4 KB band disappeared and a radioactive 1.1 KB band resulted which contained rDNA from -3.9 to -2.8 (Figure 7B, lane b). In reciprocal hybridizations, when restricted λB3 DNA was probed with ES3.1, all fragments known to contain repeat elements hybridized with ES3.1 (data not shown). These results suggest that the cloned rDNA repeat has two (-4.9 to -4.3 and -3.9 to -2.8) non-contiguous regions contain-

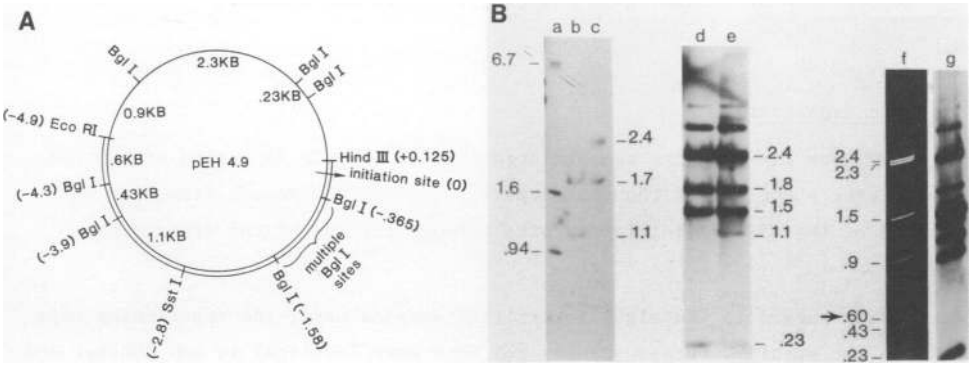


Figure 7.
A) Restriction map of pEH5.0. The double line represents rDNA sequences.
B) pEH5.0 DNA digested with Bgl I and Sst I (lanes b and e), Bgl I and EcoRI (lanes f and g) or Bgl I alone (lanes c and d), probed with nick-translated DNA from λB3 (lanes b and c) or nick-translated DNA from pBR322 bearing the rat growth hormone intron sequence (lanes d, e, and g). Lane f is the ethidium bromide stained gel of the DNA used for the Southern blot shown in lane g. The strongly hybridizing bands in lanes d, e, and g are due to the cross-annealing of pBR322 sequences. The rDNA containing band at .6 KB is indicated with an arrow. Lane a contains molecular weight markers.

ing highly repetitious DNA, at least one of which contains an Alu-equivalent type 2 sequence.

To determine which region contained the Alu related sequence, pEH5.0 DNA was digested with Bgl I and Sst I and probed with a clone containing sequences from the intron of the rat growth hormone. A 1.1 KB fragment (-3.9 to -2.8) annealed (Figure 7B, lane e), suggesting that the -3.9 to -2.8 region contains the Alu-equivalent type 2 sequence. The strong hybridization bands in lanes d, e and g are due to the cross-annealing between pBR322 sequences in pEH5.0 and the nick-translated probe.

If pEH5.0 DNA was double digested with Bgl I and EcoRI, the resulting -4.9 to -4.3 fragment could be seen as an ethidium-bromide staining band at 0.6 KB lane f), which failed to anneal to growth hormone intron sequences, suggesting that its repetitive element is not homologous to an Alu type 2 sequence (Figure 7B, lane g).

These data suggest that the rat rDNA repeat contains two non-contiguous regions (-4.9 to -4.3 and -3.9 to -2.8) that contain highly repetitious sequences. The two regions differ in that only the -3.9 to -2.8 fragment is homologous to the rodent Alu type 2 sequence found in the rat growth hormone intron while neither is homologous to the Chinese hamster Alu type 1 sequence, judged by lack of annealing to λ63 (16,17) (data not shown).

DISCUSSION

Genomic Organization

We have examined the genomic organization of a 22 KB region of rat DNA comprising about 60% of the rDNA repeat. The region extends from 9 KB upstream of the 45S RNA initiation site through the end of the transcribed region.

With regard to the eight restriction enzymes used, the restriction maps of two rat strains, Sprague-Dawley and BD-9 were identical to one another and to the cloned rDNA sequence derived from a recombinant library of Sprague-Dawley DNA. While all major bands were common between the DNAs using probes from the external and non-transcribed spacer regions some minor bands were detected suggesting some heterogeneity. It is unknown whether these differences reflect minor heterogeneity in various rDNA repeats, junctional rDNA regions abutting non-rDNA sequences, or cross-hybridizing non-nucleolar sequences, unrelated to rDNA. Although it was unclear as to what strain of rat was used, Braga *et al.* (4) demonstrated substantially greater hetero-

geneity in Southern blots of rDNA using cDNAs prepared from 18S and 28S rRNAs.

Restriction maps for other rat rDNA clones (also from Sprague-Dawley libraries) have been published (4,27,32,33) and, in general, agree with our proposed map (Figure 1). One difference among the cloned rDNAs from different laboratories is the estimated size of the Eco RI fragment containing the initiation site. Our clone contains a 10.9 KB fragment that corresponds in size to an analogous fragment in about half the Sprague-Dawley genomic rDNA repeats; the other repeats being 11.8 KB. Fuke et al. (32) isolated two clones, whose corresponding Eco RI fragments were smaller, 10.1 and 9.4 Kb; likewise, Rothblum et al. (33) described two clones whose Eco RI fragments were 9.5 and 10.5 KB while Braga et al. (4) described two clones, 0.3 KB different in size, which were both about 13 KB in length. Whether these differences arise as errors in estimating DNA lengths or actually reflect the sizes of different rDNA repeats is unclear. However, since our clone is the same size as that in the genome, its size is unlikely to have been generated artificially during cloning. Arnheim and Kuehn (10) have shown that the analogous fragment in mouse rDNA can undergo deletions and insertions during propagation in E. coli. Therefore, it may be important to insure that any clone from this region has not undergone rearrangement during cloning.

Since only eight restriction enzymes were employed in our analysis, it is possible that sequence variation not resulting in deletions or insertions would be undetected. Even with regard to the sites examined, minor rDNA repeats present at less than 2-3 copies per genome could not be detected. The presence of minor heterogeneity among repeats is, however, evident by comparing the sequences of rDNA regions. We (21), Financsek et al. (29) and Rothblum et al. (30) have reported the nucleotide sequence surrounding the rat 45S rDNA initiation site. Of the 293 nucleotides common to our sequence and that of Financsek et al. (29), there are 6 differences; 2 single base substitutions and 4 single base insertions. Comparison of the 900 bases of our sequence that overlap with that of Rothblum et al. (30) shows 12 single base differences. However, at this time, we have no way of assessing how minor sequence variation among repeats affects function.

Length Heterogeneity in the Non-transcribed Spacer

Rat rDNA has a length polymorphism 5′ to the transcribed region consisting of the insertion of 0.9 KB of DNA in the non-transcribed spacer, between -1.8 and -0.28 KB from the initiation site. Both Sprague-Dawley and the BD-9 strains exhibit this variability. Restriction mapping and partial sequencing

of the region indicate that it is composed of repeating units, 130 nucleotides in length. While the smaller polymorphic length consists of approximately 11.5 tandem repeats it is likely that the larger repeat contains correspondingly more repeat units. In this regard, Arnheim and Kuehn (10) have shown that the analogous position in mouse rDNA consists of a variable number of copies of a 135 nucleotide sequence. Mice exhibit greater polymorphism with regard to this region than rat, since any given inbred strain is likely to have 4 to 6 major sizes of DNA fragments homologous to this region. Almost a continuous variation in size exists in the spacer of Xenopus rDNA, again due to an internally repetitious sequence (34,35,36). In addition, Braga et al. (4) have also demonstrated an internally repetitious region 3' to the rat rDNA repeat. Both sequence and length polymorphisms have been used as genetic markers for determining the distribution of various mammalian rDNA repeats among different nucleolar organizer-bearing chromosomes (37,38).

By sequence analysis the rat repeat unit does not bear homology to the promoter region (21,29,30) and in this regard is not analogous to the "Bam islands" described for Xenopus rDNA (36) nor to the repetitious rDNA spacers of Drosophilia that contain regions homologous to the polymerase I initiation site (39,40). However, it is very similar in sequence to the 135 nucleotide repeat of mouse (10,31) including the extensive poly T tracts that both species share.

Highly Repetitious DNA in the Non-transcribed Spacer

We detected two regions upstream of the initiation site (-4.9 to -4.3 and -3.9 to -2.8) that contain highly repetitive DNA elements. The two fragments differ in that only the -3.9 to -2.8 region contains an Alu-type 2 sequence. Since we have not directly cross-hybridized the two rDNA repeat-containing fragments, it is possible that they share some repeat elements but have other elements (such as the Alu type 2 sequence) unique to one or the other.

Mouse rDNA contains at least two different repeat elements in their non-transcribed spacers. Kominami et al. (14) showed that mouse rDNA contains two different non-contiguous repeat elements at about -7.0 to -7.5 and -8.2 to -8.4 (relative to the mouse 45S RNA initiation site at 0). Arnheim et al. (12) demonstrated that these elements belonged to highly repetitive and interspersed families. In addition to the upstream repeat elements both mouse and rat rDNA contain repeat elements downstream from the 28S gene (4,14).

Higuchi et al. (15) have demonstrated that human rDNA also has at least 4 regions containing highly repeated elements, two of which contain Alu repeats,

the other two bearing non-Alu-like elements. The repeats in human rDNA are considerably closer to the initiation site than those in rat or mouse rDNA, beginning at about -0.8 KB, relative to the transcription initiation site, and extending to about -6.8 KB.

While the three mammalian species (mouse, human and rat) that have been investigated have multiple repeat elements in their non-transcribed spacers, albeit at varying distances from the transcribed region, it is unclear how or if they function in rDNA. It is not even known if they are transcribed. In other locations in the genome, interspersed repeat elements are transcribed by RNA polymerase II as part of larger heterogeneous nuclear RNA molecules (41,42,43) and by polymerase III as independent transcription units (17,44). In addition to being possible templates for transcription, others have proposed that Alu-like sequences function as origins of replication or as vestiges of transposable elements (45,46,47).

ACKNOWLEDGEMENTS

We thank Drs. R.B. Wallace, T. Sargent and J. Bonner for the rat library; Dr. Barbara Cordell for the rat growth hormone clone; Dr. Warren Jelinek for λ63; Mr. James Beeson and Dr. Noboru Sueoka for λB3. We thank Barbara Masters and James Beeson for identifying the initial λ clone and Nadia DeStackleberg and Carol Breibart for help in preparation of the manuscript. We acknowledge Dr. Tom Walker and Dr. Elaine Lewis for their helpful discussion and critical reading of the manuscript.

This work was supported by NIH Grants GM29302 and GM29986 to D.M.C. and NIH Grant CA18212 to K.J.D. C.A.H. was supported by NIH Predoctoral Training Grant 5T32 GM 07385-04 from 1/81 to 7/81.

REFERENCES

1. Melli, M., Whitfield, C., Rao, K.V., Richardson, M., Bishop, J.O. (1971) Nature New Biol. 231, 8-12.
2. Stumph, W.E., Wu, J.R., Bonner, J. (1979) Biochem. 18, 2864-2871.
3. Szabo, P., Lee, M.R., Elder, F.B., Prensky, W. (1978) Chromosoma 69, 161-172.
4. Braga, E.A., Yussifov, T.N., Nosikov, V.V. (1982) Gene 20, 145-156.
5. Arnheim, N. and Southern, E.M. (1977) Cell 11, 363-370.
6. Wilson, G.N., Hollar, B.A., Waterson, J.R., Schmickel, R.D. (1978) Proc. Natl. Acad. Sci. USA 75, 5376-5381.
7. Krystal, M. and Arnheim, N. (1978) J. Mol. Biol. 126, 91-104.
8. Arnheim, N. (1979) Gene 7, 83-96.
9. Grummt, I., Soellner, C. and Scholz, I. (1979) Nucleic Acids Res. 6, 1351-1369.
10. Arnheim, N. and Kuehn, N. (1979) J. Mol. Biol. 134, 743-765.
11. Wellauer, P.K. and Dawid, I.B. (1979) J. Mol. Biol. 128, 289-303.

12. Arnheim, N., Seperack, P., Banerji, J., Lang, R.B., Miesfeld, R., Marcu, K.B. (1980) Cell 22, 179-185.
13. Erickson, J.M., Rushford, C.L., Dorney, D.J., Wilson, G.N., Schmickel, R.D. (1981) Gene 16, 1-9.
14. Kominami, R., Urano, Y., Mishima, Y. and Muramatsu, M. (1981) Nucleic Acids Res. 9, 3219-3233.
15. Higuchi, R., Stang, H.D., Browne, J.K., Martin, M.O., Huot, M., Lipeles, J. and Salser, W. (1981) Gene 15, 177-186.
16. Jelinek, W.R. (1978) Proc. Natl. Acad. Sci. USA 75, 2679-2683.
17. Haynes, S.R. and Jelinek, W.R. (1981) Proc. Natl. Acad. Sci. USA 78, 6130-6134.
18. Page, G.S., Smith, S., Goodman, H.M. (1981) Nucleic Acids Res. 9, 2087-2104.
19. Messing, J., Crea, J. and Seeburg, P.H. (1981) Nucleic Acids Res. 9, 309-321.
20. Kao, F-T., Hartz, J.A., Law, M.L., Davidson, J.N. (1982) Proc. Natl. Acad. Sci. USA 79, 865-869.
21. Harrington, C.A. and Chikaraishi, D.M. (1983) Nucleic Acids Res., 11, 3317-3332.
22. Benton, W.D. and Davis, R.W. (1977) Science 196, 180-182.
23. Johnson, R.A. and Walseth, T.F. Adv. Cyclic Nucleotide Res. 10, 135-168.
24. Derman, E., Krauter, K., Walling, L., Weinburger, C., Ray, M. and Darnell, J.E. (1981) Cell 23, 731-739.
25. Southern, E.M. (1975) J. Mol. Biol. 98, 503-517.
26. Jeffreys, A.J. and Flavell, R.A. (1977) Cell 12, 1097-1108.
27. Tantravahi, U., Guntaka, R.V., Erlanger, B.F. and Miller, O.J. (1981) Proc. Natl. Acad. Sci. USA 78, 489-493.
28. Sargent, T.D., Wu, J., Sala-Trepat, J.M., Wallace, R.B., Reyes, A.A. and Bonner, J. (1979) Proc. Natl. Acad. Sci. USA 76, 3256-3260.
29. Financsek, I., Mizumoto, K., Mishima, Y. and Muramatsu, M. (1982) Proc. Natl. Acad. Sci. USA 79, 3092-3096.
30. Rothblum, L.I., Reddy, R. and Cassidy, B. (1982) Nucleic Acids Res. 10, 7345-7362.
31. Kuehn, M. and Arnheim, N. (1983) Nucleic Acids Res. 11, 211-224.
32. Fuke, M., Dennis, K.J., Busch, H. (1981) Mol. Gen. Genet. 182, 25-30.
33. Rothblum, L.I., Parker, D.L., and Cassidy, B. (1982) Gene 17, 75-77.
34. Wellauer, R.K., Dawid, I.B., Brown, D.D., and Reeder, R.H. (1976) J. Mol. Biol. 105, 461-486.
35. Buongiorno-Nardelli, M., Amaldi, F., Beccari, E. and Junakovic, N. (1977) J. Mol. Biol. 110, 105-117.
36. Boseley, P., Moss, T., Machler, M., Portmann, R. and Birnstiel, M. (1979) Cell 17 19-32.
37. Arnheim, N., Krystal, M., Schmickel, R., Wilson, G., Ryder, O., Zimmer, E. (1980) Proc. Natl. Acad. Sci. USA 77, 7323-7327.
38. Krystal, M., D'Eustachio, P., Ruddle, F.H. and Arnheim, N. (1981) Proc. Natl. Acad. Sci. USA 78 5744-5748.
39. Kohorn, B.D. and Rae, P.M.M. (1982) Nucleic Acids Res. 10, 6879-6886.
40. Coen, E.S. and Dover, G.A. (1982) Nucleic Acids Res. 10, 7017-7026.
41. Robertson, H.D., Dickson, E. and Jelinek, W.R. (1977) J. Mol. Biol. 155, 571-589.
42. Jelinek, W.R., Evans. R., Wilson, M., Salditt-Georgieff, M. and Darnell, J.E. (1978) Biochem. 17, 2776-2783.
43. Ullu, E., Murphy, S. and Melli M. (1982) Cell 29, 195-202.
44. Harada, F. and Kato, N. (1980) Nucleic Acids Res. 8, 1273-1285.
45. Haynes, S.R., Toomey, T.P., Leinward, L., Jelinek, W.R. (1981) Mol. Cell. Biol. 1, 573-583.
46. Van Arsdell, S.W., Denison, R.A., Bernstein, L.B., Weiner, A.M., Manser, T. and Gesteland, R.F. (1981) Cell 26, 11-18.
47. Jagadeeswaran, P., Forget, B.G. and Weissman, S.M. (1981) Cell 26 141-142.