Nucleotide sequence of the genetically labile repeated elements 5' to the origin of mouse rRNA transcription

Michael Kuehn and Norman Arnheim

Biochemistry Department and Molecular Biology Program, SUNY at Stony Brook, Stony Brook, NY 11794, USA

Received 20 August 1982; Revised and Accepted 1 December 1982

ABSTRACT

We have determined the complete nucleotide sequence of a cloned Balb/c mouse rDNA NTS fragment containing 13 tandem copies of a 135 bp subrepeating segment. This repetitious region (VrDNA) lies close to the origin of ribosomal RNA transcription. Analyses of these VrDNA subrepeats from Balb/c and a related species, Mus pahari, reveal regions of inverted repeat DNA as well as large poly T tracts, either of which may be significant to the generation of the high levels of VrDNA copy number variation found in wild and inbred mice and/or the modulation of rRNA synthesis. Unlike the highly homogeneous subrepeats in the Xenopus laevis NTS repetitious regions, the VrDNA subrepeats differ from one another on the average by about 13%. Sequence analysis and Southern hybridization studies have also shown that, unlike the Xenopus and Drosophila NTS, extensive duplications of sequences found surrounding the mouse rRNA initiation site are found neither in the VrDNA region nor 6 kb further upstream in the NTS.

INTRODUCTION

The eukaryotic ribosomal gene (rDNA) repeating unit consists of both transcribed and non-transcribed segments. The transcribed portions of the tandemly arranged repeating units therefore are separated from one another by non-transcribed spacer (NTS) DNA. The function of the NTS is unknown. One feature which is characteristic of many species is the variation in absolute length of the NTS found among rDNA repeating units both within and between individuals. This length variation has been localized in some species to specific regions of subrepeated DNA within the NTS (for reviews see 1-3).

In the inbred mouse, a variably sized region of repeated DNA [VrDNA (4)] lies only 210 bp upstream from the origin of transcription of the rRNA precursor (5-7). The mouse VrDNA regions are composed of variable numbers of tandemly repeated copies of an approximately 135 base pair (bp) subrepeat. Each tandem array of subrepeats is bounded by Hind II sites and can encompass an area of between 500 bp and 3000 bp within the

0305-1048/83/1101-0211$2.00/0

total 40 kb of a single mouse rDNA repeating unit. Possible mechanisms to generate this length variation would involve recombination between or within rDNA repeats, occurring at VrDNA regions in which the subrepeats are in unequal alignment. With respect to the ability of VrDNA sequences to promote or engage in recombination events, it was quite interesting to find that cloned mouse rDNA fragments carrying VrDNA regions are unstable during propagation in E. coli (4). Deletions and additions of integral numbers of subrepeats which mimic precisely the nature of the length variation detected for VrDNA regions in the mouse genome were found to occur at a significant rate.

The close proximity of subrepeated sequences to the rRNA transciption initiation site in the mouse and several other organisms (see 2,3) raises the possibility that these sequences might serve a function in the modulation of rRNA synthesis. Although there is no compelling evidence yet available to support this hypothesis for RNA polymerase I transcription units, there is a growing body of data demonstrating the in vivo modulating effect of specific sequences lying at considerable distances upstream from the initiation sites for RNA polymerase II transcription (8-12).

As a first step in understanding the relationship between VrDNA structure and function, we have determined the nucleotide sequence of one tandem cluster of subrepeats and directly flanking sequences cloned from Balb/c mice. Upon examination, a number of structural features are apparent in this sequence which could be involved in the modulation of rRNA synthesis as well as in recombination events capable of generating genomic length variation. We have also examined the VrDNA subrepeats of a distantly related mouse species, Mus pahari, and have found these same features to be present.

MATERIALS AND METHODS

DNA Samples

Subcloning of the VrDNA region from a cloned Eco RI fragment of Balb/c rDNA and purification of plasmid DNA has been described (4). DNA purification from frozen livers of the inbred strains Balb/cJ, DBA/2N, as well as the wild species Mus musculus musculus (Czech.I), Mus musculus domesticus (Centerville light), Mus cervicolor, Mus pahari, Mus spraetus hispanicus, Mus musculus molossinus, Mus caroli, Mus musculus musculus (Danish), Mus musculus praetexus, Mus musculus brevirostris, Mus musculus

castaneus and Mus cervicolor popaeus was performed by a modification of the procedure of Walker and McClaren (13). Liver samples from these wild species were kindly provided by Dr. Michael Potter. Filter lifts from platings of a Charon 28 library of Mbo I partially digested Mus pahari DNA (kindly provided by Larry Stanton) were performed according to Benton and Davis (14) and screened with the Balb/c VrDNA probe to select Mus pahari VrDNA containing ribosomal gene clones. The VrDNA regions from these clones were then subcloned into pBR-322.

Methods for restriction enzyme digestion, agarose gel electrophoresis, Southern transfer, nick translation, hybridization and autoradiography have also been described previously (4).

Nucleotide Sequence Analysis

5'-labeling of restriction fragments, recovery from polyacrylamide gels and DNA sequencing were performed by the procedures of Maxam and Gilbert (15). Figure 2B outlines the strategy employed to determine the sequence of the 1767 bp Balb/cJ VrDNA clone. The Mus pahari VrDNA subclone was 5'-labeled at its terminal Hind II (Sal I) sites, and sequenced after strand separation performed according to Maxam and Gilbert (15). Sequences were analyzed by computer and by inspection to determine regions of homology and potential secondary structure.

RESULTS

Further Studies on Length Heterogeneity in Mouse VrDNA

Our earlier studies revealed extensive length variation in the VrDNA region of several inbred and a few closly related wild mice (4). In this report we extend our study of the genomic organization of VrDNA sequences to include a large sample of more distantly related wild mice. DNA samples digested with Hind II and electrophoresed on agarose gels were hybridized to a cloned VrDNA fragment after Southern transfer. As shown in Figure 1, all the individuals examined in this analysis have multiple Hind II fragments that are detected by the VrDNA probe. For instance, Hind II digestion of Balb/cJ DNA produces one minor and 4 major VrDNA fragments with sizes of about 0.73, 1.20, 1.35, 1.75 and 2.70 kb. Some wild species present virtually uninterrupted stepladders of VrDNA fragments. This is most striking in the DNA from Mus pahari and Mus musculus molossinus. Our analysis of the pattern of Hind II fragments from both these species clearly shows that size differences are in units of 135 bp even though the absolute sizes of the fragments vary among species.

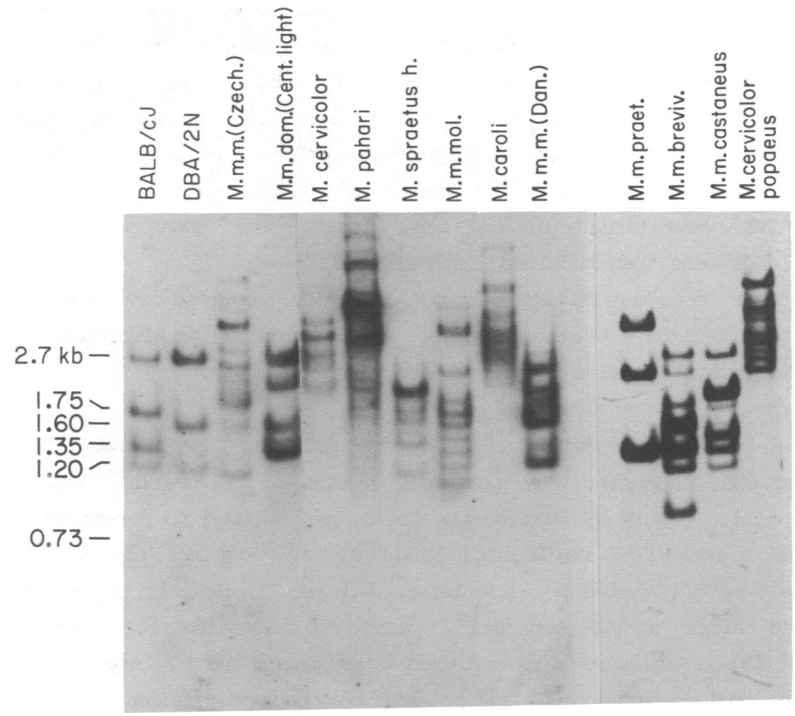

Figure 1. Southern transfer and hybridization analysis of Hind II digests of purified mouse DNA using the Balb/c VrDNA clone as probe. Sources for the DNA samples are denoted above each lane. Complete names for the different species can be found in the materials and methods section.

Individuals from some of the mouse species presented here as well as those analyzed previously tend to have a more limited number of VrDNA size classes with some having as few as three major fragments. Our statistical analysis of the previously reported VrDNA fragments (4) suggests however that the relatively few fragments found in a particular individual can be considered to be a random subset of a larger series of fragments that vary from each other by 135 bp.

Nucleotide Sequence Determination of a VrDNA Region

We have previously reported on the construction of a plasmid clone containing a 1767 bp VrDNA region derived from Balb/cJ mice (4). The exact strategy used for determining the sequence of the insert is shown in Figure 2 as are the cleavage sites for all restriction enzymes used in the sequencing. This cloned Sal I fragment contains 12 full copies and 1

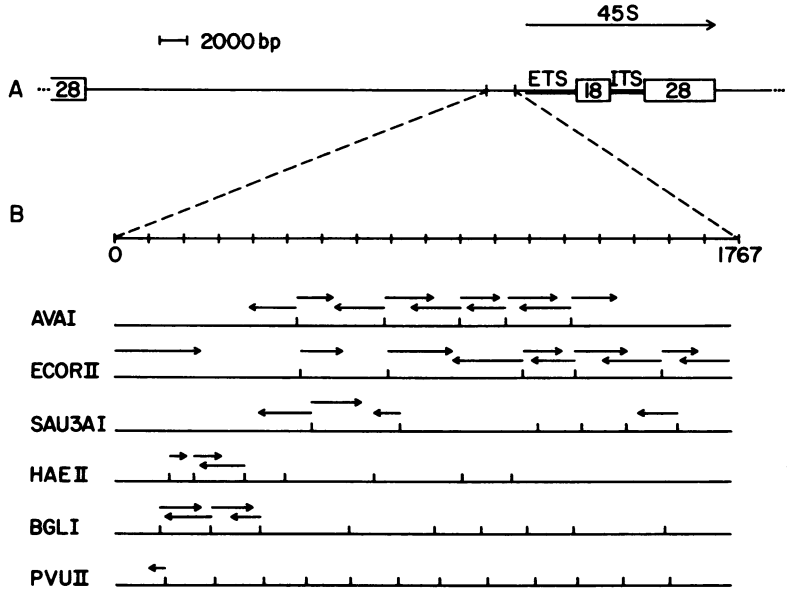

Figure 2. A. Diagram of the mouse rDNA repeating unit. The transcribed region is delineated by the arrow marked 45S. The hatch marked region to the left of the ETS contains the VrDNA region.
B. The VrDNA region of A. Each hatch-mark represents 100 bps. The sites for cleavage by restriction enzymes used in the determination of the nucleotide sequence are shown below. Arrows indicate regions that were sequenced after being end labeled at the cleavage site.

incomplete copy of the basic VrDNA subrepeating unit. Although very similar in sequence, a number of differences do exist among the subrepeats. This feature was indispensible in the correct ordering of the smaller sequenced fragments.

Most of the sequence has been determined from both strands or from the same strand more than once. We felt that extensive overlapping was essential for such a sequence where small variation between repeats could be overlooked or mistakenly assumed given the normal rate of human error in sequencing. Aspects of the sequence which were problematic included methylated bases, contiguous identical residues and regions capable of forming snapback DNA. When methylated, the second C in the Eco RII recognition sequence, CCAGG, appears as a blank in the sequence ladder. We have seen up to three contiguous C residues migrating as one band. Fragments produced by the sequencing reactions which have 3' ends capable

GTCGACCAGAAGGCTTAAGTCCTACCCCCCCCCCC $(T)_{14}$

REPEAT	
13	CCTCCAGAAGCCCTCTCTTGTCCCCGTCACCGGGGGCAC CGTACATCTGAGGCCGAGAGGACGCGATGGGCCC GGCTTCCAAGCCGGTGTGGCTCGGCCAGCTGGCGCTTCGGGTC $(T)_{25}$
12	CCTCCAGAAGCCTTGTCT GTCGCTGTCACCGGGGCGC TGTACTTCTGAGGCCGAGAGGACGCGATGGGCCCCGGCTTCCAAGCCGGTGTGGCTCGGCCAGCTGGAGCTTCGGGTC $(T)_{27}$
11	C TCCAGAAGCCTTGTCT GTCGCTGTCACCGGGGCGC TGTACTTCTGAGGCCGAGAGGACGCGATGGGTC GGCTTCCAAGCCGATGTGGCGGGGCCAGCTGGAGCTTCGGG $(T)_{11}$
10	CCTCCAGAAGCCCTCTCTTGTCCCCGTCACCGGGGCGC TGTACTTCTGAGGCCGAGAGGACGTGATGGGCCC GGGTTCCAGGCGGATGTCGCCCGGTCAGCTGGAGCTTTGGATC $(T)_{15}$
9	CCTCCAGAAGCCCTCTCTTGTCCCCGTCACCGGGGCAC CTTACATCTGAGGGCGAGAGGACGTGATGGGTCC GGCTTCCAAGCCGATGTGGCGGGGCCAGCTGGAGCTTCGGG $(T)_{12}$
8	CCTCCAGAAGCCCTCTCTTGTCCCCGTCACCGGGGCGC TGTACTTCTGAGGCCGAGAGGACGTGATGGGCCC GGGTTCCAGGCGGATGTCGCCCGGTCAGCTGGAGCTTTGGATC $(T)_{11}$
7	CCCTCCAGAAGCCCTCTCTTGTCCCCGTCACCGGGGCAC CGTACATCTGAGGCCGAGAGGACACGATGGGCCT GTCTTCCAAGCCGATGTGGCCCGGCCAGCTGGAGCTTCGGGTC $(T)_{15}$
6	CCTCCAGAAGCCTTGTCT GTCGCTGTCACCCGGGGCGC TGTACTTCTGAGGCCGAGAGGACGCGATGGGCCC GGCTTCCAAGCCGGTGTGGCTCGGCCAGCTGGAGCTTCGGGTC $(T)_{21}$
5	CCTCCAGAAACCTTGTCT GTCGCTGTCACCCGGGGCGCTTGTACTTCTGATGCCGAGAGGACGCGATGGGCCC GTCTTCCAGGCCGATGTGGCCCGGTCAGCTGGAGCTTTGGATC $(T)_{16}$
4	CCTCCAGAAGCCCTCTCTTGTCCCCGTCACCGGGGCAC CTTACATCTGAGGCCTAGAGGACACGATGGGCCC GGGTTCCAGGCCGATGTGGCCCGGTCAGCTGGAGCTTTGGATC $(T)_{14}$
3	CTTCCAGAAGCCCTCT T GTCCCCGTCACCGGTGGCAC TGTACATCTGAGGCCGAGAGGACATTATGGGCCC GGCTTCCAATCCGATGTGGCCCGGTCAGCTGGAGCTTTGGATC $(T)_{2}A(T)_{11}A(T)_{6}$
2	CTTCCAGAAGCCCTCT T GTCCCTGTCACCGGTGGCAC GGTACATCTGAGGCCGAGAGGACATTATGGGCCC GGCTTCCAGGCCGATGTGGCCCGGTCAGCTGGAGCTTTGGATC $(T)_{15}C(T)_{6}$
1	CCTCCAGAAGCCCTCTCT GTCCCTGTCACCGGGGGCCC TGTACGTCTGAGGCCGAGGGAAAGCTATGGGCGC GGTTTTC

TTTCATTGACCTGTCGGTCTTATCAGTTCTCCGGGTTGTCAGGTCGAC

Figure 3. Nucleotide sequence of the cloned Balb/c Hind II fragment containing VrDNA. The 5' and 3' boundries of the VrDNA region with respect to the non-coding strand are at the upper left and lower right portions of the figure respectively. The subrepeats have been aligned to maximize homology. Individual subrepeats are bounded by T clusters, the lengths of which are denoted numerically. In repeat 8 an A residue immediately 5' to the $(T)_{11}$ cluster was inadvertently omitted. Subrepeat 1 is incomplete being only 80 bp in length.

of folding back and forming stable base pairs, even under our standard denaturing conditions (8M urea gels run at elevated temperatures), have increased mobility in the sequencing gels. This is probably due to the more compact structure of these molecules. Their anomolous migration also results in gaps and compressions in the sequence ladder. These problems were resolved by sequencing of the other strand or electrophoresis to a different position in a sequencing gel of the same or different percent acrylamide.

Structural Features of the Sequence

In Figure 3, the complete nucleotide sequence of the VrDNA containing fragment is presented to illustrate the repeating structure. Within the 1767 bp sequenced are 12 full subrepeats which vary in size from 124 bp to 142 bp and one truncated copy of approximately 80 bp. The 5' boundary of the VrDNA sequence is a cluster of 12 C residues followed by 14 contiguous T residues, reading the non-coding strand. Each subrepeat contains similar long runs of T's which range from 11 to 27 bp in length. The length variation seen between subrepeats is almost entirely due to this variation in length of the T clusters. The T clusters are bounded by an additional 5 to 7 pyrimidine residues. In one subrepeat (no.13) 36 out of 39 contiguous residues are pyrimidines, of which 29 are T's. Long stretches of T residues have also been reported for downstream regions within the ETS of mouse rDNA (16,17). The 5' end of the ETS and the 3' end of the VrDNA region are separated by approximately 210 bp.

In Figure 3, the subrepeats have been aligned to maximize homology. Other than T cluster length differences, the subrepeats show great similarity in sequence. Of the 118 nucleotide positions between successive T clusters, all subrepeats are identical at 76 positions. Variation at the other 42 sites is to differing degrees. On the average, variation between any two subrepeats occurs at 16 of the 118 positions giving an overall average similarity of 87%. The two most dissimilar subrepeats vary at 25 of the 118 sites (a similarity of 79%) while two subrepeats are identical at every position. Interestingly, these two invariant copies are not adjacent, being separated by one repeat, which led us to examine the degree of nucleotide sequence similarity among all the VrDNA subrepeats with respect to their positions within the linear array. Figure 4 shows this relationship which was determined by analysis of a similarity matrix generated from pairwise comparisons of all the

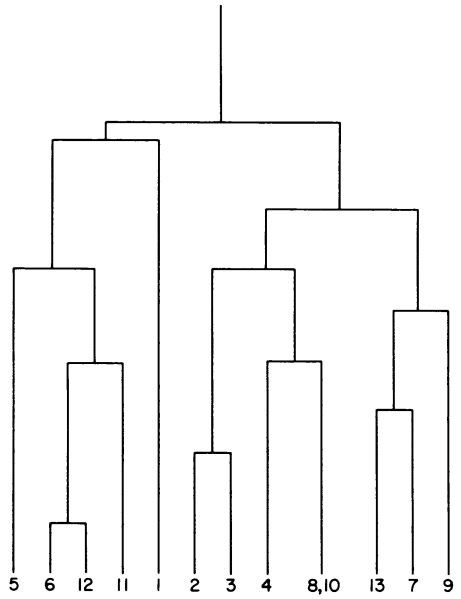

Figure 4. Tree representing the degree of nucleotide sequence relatedness among the subrepeats in the cloned VrDNA array. The numbers refer to the position of the subrepeat in the array (see Fig. 3). This tree was generated by the NT-SYS program developed by F. James Rohlf, John Kishpaugh and David Kirk. Our procedure for estimating the degree of divergence of the truncated subrepeat (No. 1) from the remainder of the subrepeats may have led to a considerable underestimate of the true level.

subrepeats at each of the variable sites. The numbers at the bottom of each tree branch indicate the position of that subrepeat within the linear array. It is clear that there is no simple relationship between the degree of homology among the subrepeats and their position within the linear array.

Given the proximity of VrDNA to the initiation site for rRNA synthesis (210 bp) and its extensive length variation in mice, as well as the high rate of recombination for this sequence during propagation as a cloned fragment (4), we examined the sequence for structural features that may be relevant to these processes. In addition to the long poly T tracts, we have found regions of inverted repeat DNA within the subrepeats. Under the correct conditions inverted repeats can form hair pins (stem and loop structures) (18-20). Such structures have been postulated as recognition elements mediating various genetic events

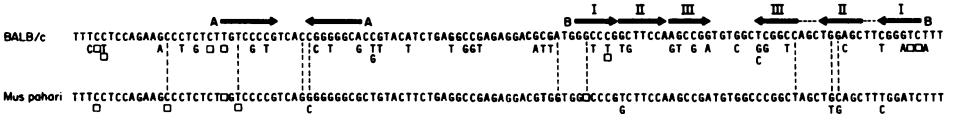

Figure 5. Comparison of composite subrepeat sequences for Balb/c VrDNA and Mus pahari VrDNA. Nucleotides below the line show positions of heterogeneity within each species. Vertical dashed lines indicate positions in the Mus pahari sequence where a nucleotide is found that is not seen at the homologous position in Balb/c VrDNA. The open squares denote the fact that some repeats are deleted at this site. Regions of inverted repeat sequences are overlined by the arrows marked A and B. The horizontal lines indicate marked interruptions in the inverted repeats.

including transcription, replication and recombination (see 18,20). Figure 5 shows the composite sequence for the Balb/c VrDNA subrepeats we examined. Consideration of unpublished sequence data (M. Kuehn and D. Treco) on 12 additional subrepeats obtained from 4 independently derived rDNA clones confirm the generality of this composite sequence for Balb/c mice. Arrows above the sequence illustrate the inverted repeats found in the VrDNA subrepeats. As described above, the ability of inverted repeat DNA to assume a secondary structure was quite apparent during sequencing. In fact, the regions which did show anomalous migration on sequence gels are exactly those inverted repeats delineated in Figure 5, labeled A and B. Of these, the longest perfect inverted repeat is A which can consist of an eight base pair stem and a four base loop. Inverted repeat B is an imperfect structure having a stem extending for approximately 20 base pairs with two mispaired regions and a loop of six bases. Other perfect inverted repeats are present, all having stems of seven or less basepairs. Due to sequence variation certain subrepeats lack part or all of the illustrated inverted repeats. A recent study using a mammalian 'site specific' endonuclease, which appears to recognize stem and loop structures rather than specific nucleotide sequences, demonstrated the presence of distinct cleavage sites in the mouse VrDNA subrepeats (35).

Primary sequence data for a VrDNA region cloned from a different mouse species, Mus pahari, was obtained. Approximately 500 bp of Mus pahari VrDNA was sequenced covering about 3 full and 2 incomplete subrepeats. A composite sequence for the Mus pahari VrDNA subrepeating elements is presented below the Balb/c composite sequence in Figure 5. Their VrDNA sequences are virtually identical. At 9 positions in the Mus

pahari sequence a nucleotide is present which has never been found in the homologous position in the total of 25 Balb/c VrDNA subrepeats sequenced. The Mus pahari subrepeats also have long runs of T residues and contain regions of inverted repeats in positions homologous to those in the Balb/c subrepeats.

DISCUSSION

Our analyses of the rDNA from a number of mouse species and inbred strains has revealed the presence of a repetitious region terminating 210 bp upstream from the origin of rRNA synthesis. In addition to mice, a series of repetitive elements adjacent to the origin of rRNA transcription is a feature characteristic of rDNA from Xenopus (21-25), Drosophila (26-29), Tetrahymena (30) and human (31). Sequences in close proximity to the origin of transcription may play some role in the control of rRNA synthesis.

In addition to the several classes of short tandemly subrepeated elements found among these organisms, extensive duplications of the sequences at the rRNA initiation site have also been detected in the NTS in some of these species. For example, sequence analysis of the NTS of Xenopus laevis (23,24) has revealed the presence of two non-tandem duplications, the so-called Bam islands, which are highly homologous to sequences at and around the rRNA initiation site (119 out of 132 bp or 90% are identical). More recently, analysis of the Drosophila melanogaster NTS 240 bp subrepeated elements has shown them to contain a 68 bp stretch of nucleotides with extensive homology (92%) to the initiation site (33 and E. Coen and G.A. Dover, personal communication). Although this situation is similar to Xenopus, the organization of the sequences homologous to the Drosophila initiation site is quite distinct: being in the subrepeats themselves they are tandemly arranged, present in many more copies and reside closer to the actual initiation site. As proposed in a model for NTS participation in the control of rDNA expression, duplications of the initiation regions in these two organisms may function to provide multiple sites on rDNA genes for the binding and sequestering of Pol I molecules in preparation for transcription (23). In fact, transcription has occasionally been observed to initiate at these upstream sites (see 2,32) but whether this transcription is significant to normal rDNA function in unknown.

Doubts on the generality of this model are raised by the absence of

extensive duplications of sequences found at the respective initiation sites of mouse and human rDNA in the NTS regions adjacent to the origins of transcription of these two organisms. The mouse VrDNA elements show only short stretches of perfect homology (less than 9 bp) to sequences found surrounding the initiation site. These perfectly homologous stretches can be part of longer but highly interrupted homologies. Southern transfer and hybridization studies (4) under moderately stringent conditions (3XSSC, 65 C; hybridization and wash conditions) also extend for a further 6 kb upstream the absence of significantly long sequences homologous to the initiation site. Sequence analysis of 480 bp upstream from the human rDNA initiation site and Southern hybridization studies covering an additional 12 kb of contiguous NTS (31) point to similar conclusions. Thus if some aspect of rRNA transcription commonly involves duplications of sequences surrounding the initiation site, the requirement in mouse and human must be satisfied by duplications which are much less extensive and/or which reside very much further upstream than is the case in Drosophila and Xenopus.

Though the VrDNA subrepeats lack extensive homology to the mouse rRNA initiation site these and the short tandemly subrepeated elements found in other organisms may have other functional roles in rDNA transcription. The expression of certain genes transcribed by RNA polymerase II have been shown to be influenced by the presence of sequences upstream from recognized promoter regions (8-12). Perhaps the best characterized examples of these so called enhancer or activator sequences are the 72 bp tandem repeats of SV40 which are essential for the expression of viral early genes (10). Recently, a positive cis-acting effect of the 72 bp elements on transcription of heterologous genes has also been seen and has been shown to occur even over relatively large distances (11).

Mouse rDNA NTS length polymorphisms, which can be found even within the same individual, result from differences in the number of copies of the tandemly repeated sequence which comprises the VrDNA region. The observed VrDNA length (subrepeat copy-number) variation is most likely the byproduct of continual recombination events involving rDNA genes. By one mechanism, crossing over between different rDNA genes occurring at the tandemly repetitious VrDNA region may be unequal due to the large number of homologous but out-of-register subrepeat pairing combinations possible between exchanging VrDNA arrays. The presence of inverted

repeat elements in VrDNA may be relevant to this process since several molecular models of recombination postulate an essential role for inverted repeats (see 34). The presence of long runs of T residues which are heterogeneous in length in different subrepeats suggests that these may be the sites of intermolecular recombination. Unequal alignment at the T clusters of the subrepeats involved in an inter-VrDNA recombination event would result in new T cluster lengths by analogy to the production of new VrDNA lengths (copy-number). However T cluster length variation might also arise by polymerase slippage during DNA replication (37) as has been seen for Polymerase I of E coli (43). Variation in subrepeat copy number may be entirely due to intramolecular deletions and/or additions of whole subrepeats occurring within the VrDNA segment of a single rDNA gene. The presence of the inverted repeat sequence elements within the subrepeats as well as long clusters of T residues may have significance to these processes. Single strand DNA loops containing one or more whole subrepeats, perhaps stabilized at their base by inverted repeats (self complementary sequences contributed by different subrepeats), could be degraded resulting in deletions (36). Such a model is particulary attractive for tandemly repeated sequences, where the production of single stranded loops equal in length to one or more whole subrepeats could result from slipped mispairing within the array (37). A number of observations in other systems fit such a model of the mispairing of directly repeated DNA leading to deletions (37-41). Alternatively, the loop could serve as the template for DNA synthesis with the resultant duplex incorporated into the molecule in a subsequent step (36,42) thereby leading to duplications. Taken together, these models can explain how contraction and expansion of a tandem array is possible without requiring recombination with other homologous arrays.

The Xenopus laevis non-transcribed spacer has been completely sequenced from a single clone. Of the four repetitious regions which have been identified (24) there is extensive sequence data available for two, repetitious regions 1 and 2. Within a particular region, the subrepeats are virtually identical. Our analysis shows that the subrepeats in repetitious region 1 are on the average 96% identical, Those in repetitious region 2 show an even higher (98%) similarity although these similarities approach 100% when the terminal subrepeats are removed from the analysis. This stands in contrast to the average difference of 13% among all the VrDNA subrepeats. This average

difference is unaffected by exluding the terminal subrepeats from the analysis. Assuming that the rates of mutation are identical in Xenopus and mice, the differences in levels of subrepeat variation between the cloned Xenopus and mouse arrays may reflect species-specific differences in the rates of fixation.

The role of specific sequence elements in the genetic behavior of the VrDNA region is open to further experimental analysis. Because of certain specific differences in sequence among subrepeats, recombination products occurring in either procaryotic or eucaryotic vector systems can be analyzed at the nucleotide sequence level to determine precisely the sites at which new joints are formed.

REFERENCES

1. Reeder, R.H., (1974) In Ribosomes, Nomura, M., Tissieres, A., and Lengyel, P., Eds., Cold Spring Harbor Press, New York, pp. 489-518.
2. Long. E.O. and Dawid, I.B. (1980) Ann. Rev. Biochem. 49, 727-764.
3. Treco, D., Brownell, E. and Arnheim, N. (1982) In The Cell Nucleus, H. Busch, Ed., Volume XII, Academic Press, N.Y. In Press.
4. Arnheim, N. and Kuehn, M. (1979) J. Mol. Biol. 134, 743-765.
5. Miller, K.G. and Sollner-Webb, B. (1981) Cell 27, 165-174.
6. Mishima, Y., Yamamoto, O. Kominami, R. and Muramatsu, M. (1981) Nucl. Acids Res. 9, 6773-6785.
7. Grummt, I. (1981) Nucl. Acids Res. 9, 6093-6102.
8. Grosschedl, R. and Birnstiel, M. (1980) Proc. Natl. Acad. Sci. USA 77, 1432-1436.
9. Benoist, C. and Chambon, P. (1981) Nature, 290, 304-310.
10. Gruss, P., Dhar, R. and Khoury, G. (1981) Proc. Natl. Acad. Sci. USA 78, 943-947.
11. Banerji, J., Rusconi, S. and Schaffner, W. (1981) Cell 27, 299-308.
12. McKnight, S.L. and Kingsbury, R. (1982) Science 217, 316-324.
13. Walker, P.M.B. and A. McLaren (1965) J. Mol. Biol. 12, 394.
14. Benton, W.D., and R.W. Davis. (1977) Science 196, 180.
15. Maxam, A.M. and Gilbert, W. (1980) In Method in Enzymology, Vol. 65, Grossman, L. and Moldave, K., Eds., Academic Press, New York, pp. 499-560.
16. Urano, Y., Kominami, R., Mishima, T. and Muramutsu, M. (1980) Nucl. Acids Res. 8, 6043-6058.
17. Bach, R., Grummt, I. and Allet, B. (1981) Nucl. Acids Res. 9, 1559-1569.
18. Panayotatos, N. and Wells, R.D. (1981) Nature 289, 466-470.
19. Lilley, J.M.J. (1980) Proc. Natl. Acad. Sci. USA 77, 6468-6472.
20. Lilley, D.M.J. (1981) Nucl. Acids Res. 9, 1271-1289.
21. Wellauer, P., Dawid, I.B., Brown, D.D. and Reeder, R.H. (1976) J. Mol. Biol. 105, 461-486.
22. Botchen, P., Reeder, R.H. and Dawid, I.B. (1977) Cell 11, 599-607.
23. Boseley, P., Moss, T., Machler, M., Portmann, R. and Birnstiel, M.D. (1979) Cell 17, 19-31.
24. Moss, T., Boseley, P.G., and Birnstiel, M.L. (1980) Nucl. Acids Res. 8, 467-485.
25. Bach, R., Allet, B. and Crippa, M. (1981) Nucl. Acids Res. 9, 5311-5330.

26. Wellauer, P. and Dawid, I.B. (1977) Cell 10, 193-212.
27. Long, E.O., and Dawid, I.B. (1979) Nucl. Acids Res. 7, 205-215.
28. Rae, P.M.M.; Barnett, T. and Murtif, V.L. (1981) Chromosome (BerL.) 82, 637-655.
29. Renkawitz-Pohl, R., Glatzer, K.H. and Kuntz, W. (1980) Nucl. Acids Res. 8, 4593-4611.
30. Niles, E.G., Sutiphony, J., and Haque, S. (1981) J. Biol. Chem. 256, 12849-12856.
31. Miesfeld, R. and Arnheim, N. (1982) Nucl. Acids Res. 10, 3933-3949.
32. Kohorn, B.D. and Rae, P.M.M. (1982) Proc. Natl. Acad. Sci. USA 79, 1501-1504.
33. Kohorn, B.D., Roberts, M.P. and Rae, P.M.M. (1982) J. Cell Biol. 95, 474A.
34. Radding, C.M. (1978) Annu. Rev. Biochem. 47, 847-880.
35. McKenna, W.G., Brown, F.L., Musich, P.R. and Maio, J.J. (1982) J. Mol. Biol. 154, 379-384.
36. Radding C.M. (1979) Cold Spring Harbor Symp. Quant. Biol. 43, 1315-1316.
37. Streisinger, G., Okada, Y., Emrich, J., Newton, J., Tsugita, A., Terzaghi, E. and Inouye, M. (1966) Cold Spring Harbor Symp. Quant. Biol. 31, 77-84.
38. Farabaugh, P.J. Schmissner, V., Hofer, M. and Miller, J.H. (1978) J. Mol. Biol. 126, 847-863.
39. Albertini, A.M., Hofer, M., Calos, M.P. and Miller, J.H. (1982) Cell 29, 319-328.
40. Foster, T.J., Lundblad, V., Halling, S.M., Hanley-Way, S. and Kleckner, N. (1981) Cell 23, 215-227.
41. Efstratiadis, A., Posakony, J.W., Maniatis, T., Lawn, R.M., O'Connell, C., Spritz, R.A., DeRiel, J.K., Forget, B.G., Weissman, S.M., Slightom, J.L., Blechl, A.E., Smithies, O., Barelle, F.E., Shoulders, C.C. and Proudfoot, N.J. (1980) Cell 21, 653-668.
42. Maio, J.J., Brown, F.L. and Musich, P.R. (1977) J. Mol. Biol. 117, 637-655.
43. Kornberg, A. (1980) DNA Replication, W.H. Freeman, San Francisco.