
5S Ribosomal RNA genes of the newt *Notophthalmus viridescens*

Brian K.Kay* and Joseph G.Gall

Department of Biology, Yale University, New Haven, CT 06511, USA

Received 1 September 1981

ABSTRACT

The genes which code for the 5S ribosomal RNA in the newt, *Notophthalmus viridescens* have been cloned and analyzed. Two types of repeating unit were detected: a major type consisting of a 120 bp coding region with a 111 bp spacer, and a minor type composed of a coding region, a pseudogene, and a 113 bp spacer. The pseudogene is a 36 bp segment which corresponds to the 3' terminal third of the 5S RNA gene, and is situated immediately 3' to the gene, being separated from it by 2 bp. Two recombinant plasmids were obtained in which the major and minor units were arranged in an interspersed pattern.

INTRODUCTION

The DNA which codes for the 5S ribosomal RNA of eukaryotes consists of tandemly repeating units containing a gene and a spacer sequence. Recently, the ability to construct recombinant DNA molecules has facilitated the fine structure analysis of individual and multiple 5S DNA repeats from the genomes of several eukaryotes (15). We were interested in analyzing the structure of the 5S DNA in the newt, *Notophthalmus (Triturus) viridescens*, for several reasons. We wondered what kind of heterogeneity might exist in newt 5S DNA, as there are four chromosomal loci which contain the genes (9,22). In addition, recombinant DNA molecules containing 5S RNA genes could be used for *in situ* hybridization and *in vitro* transcription experiments, so as to determine the structure of 5S RNA transcription units on lampbrush chromosomes and to identify the sequences in 5S DNA which regulate transcription.

In this paper we describe the structure and sequence organization of single and multiple 5S DNA repeats from the newt. We have found that the 5S DNA consists of long tandem arrays of repeats containing a short spacer sequence. In some repeats we have also detected a pseudogene, which consists of the 3' terminal 36 bp of the 5S RNA coding region.

EXPERIMENTAL PROCEDURES

Hind III, Mbo I, Msp I, Bst NI, Alu I, Tha I, and T4 DNA ligase were obtained from New England Biolabs. Bam HI, Eco RI, polynucleotide kinase, *E. coli* DNA polymerase I, Klenow DNA polymerase, and calf intestine alkaline phosphatase were purchased from Boehringer-Mannheim. Taq I and Sau 3A were the gifts of H. Erba and C. Yen, respectively.

To isolate DNA enriched in 5S RNA genes, Notophthalmus DNA was centrifuged in Hg^{++}/Cs_2SO_4 gradients, with Hg^{++} :DNA = 0.10 (weight:weight). The 5S DNA sequences were located by denaturing gradient fractions, affixing them to nitrocellulose filters (0.45 μ m, Millipore), and hybridizing the filters with 3H -5S RNA. 3H -5S RNA was isolated by preparative gel electrophoresis of RNA from 3H -uridine labelled Xenopus kidney cell cultures (4). The 5S DNA was centrifuged to the light side of the bulk of the DNA and was further enriched by successive centrifugations. This DNA was used for all cloning experiments. Twenty micrograms of this DNA were digested with Alu I and electrophoresed in a 2% low gel temperature agarose gel (Biorad) according to published procedures (1). A region of the gel which contained 231 bp fragments was excised, and the DNA was recovered by phenol extraction of the melted gel slice. To the blunt ends of the Alu I fragments, synthetic Eco RI and Bam HI linkers (Collaborative Research) were attached with T4 DNA ligase (17). After digestion with Eco RI and Bam HI, the 231 bp fragments were recovered by gel electrophoresis and ligated to pBR322 DNA (2) which had previously been cleaved with Eco RI plus Bam HI. *E. coli* strain X1776 (gift of R. Curtiss, III) was transformed according to a published protocol (10) and transformants were selected on agar plates containing ampicillin (100 μ g/ml; Sigma). Colonies which harbored plasmids with 5S DNA inserts were identified by colony hybridization (8) with Xenopus 5S RNA, end-labelled with ^{32}P (16).

All initial work involving recombinant DNA was performed under P2 and EK2 containment conditions in accordance with NIH guidelines published in June 1976. Later, with the issuance of subsequent guidelines, work was done under P2 and EK1 conditions.

Gradient enriched 5S DNA was partially digested with Mbo I so as to cleave at about 10% of the Mbo I sites. Mbo I was heat inactivated (5 min at 65°C) and the DNA ligated to a Bam HI-digested pBR322 DNA. Prior to transformation of *E. coli* strain HB101, the ligated DNAs were digested with Bam HI to select for recombinant DNA molecules (see Results). Colonies

which contained recombinant plasmids with 5S DNA inserts were identified as above.

Transfer of DNA fragments from gels to nitrocellulose filters, preparation of ^{32}P -labelled DNA by nick-translation, and filter hybridizations were as described (1). Labelled DNA segments were sequenced by the chemical cleavage method (18). DNA restriction fragments for sequencing were labelled at their 5' ends by polynucleotide kinase after the terminal phosphate had been removed by calf intestine alkaline phosphatase. Fragments were labelled at their 3' ends with Klenow DNA polymerase and (α - ^{32}P) dNTPs (19). Sequencing gels were prepared and electrophoresed as described elsewhere (23). Autoradiography was done with preflashed Kodak SB-5 film and Lightning Plus intensifying screens (Cronex) at -75°C . Restriction sites were mapped in the inserts of plasmids p5S101 and p5S102 DNA by the partial digest method (25). Size markers for these experiments were prepared by end-labelling lambda Hind III and pBR322 Taq I fragments with Klenow polymerase and (α - ^{32}P) dATP and (α - ^{32}P) dCTP, respectively. ^{32}P -labelled X174 DNA was prepared by nick-translation and was the gift of N. Kan.

RESULTS

To determine the best method for cloning Notophthalmus 5S RNA genes, genomic DNA was digested with a number of restriction enzymes, electrophoresed on agarose gels, transferred to nitrocellulose filters, and then hybridized with ^{32}P -5S RNA. The majority of enzymes used did not cut the 5S DNA as evidenced by hybridization at the limit of mobility in the gel. Several enzymes (Alu I, Eco RI*, Taq I) however, produced a series of hybridizing fragments that were 231 bp or integral multiples thereof in length. Using DNA enriched for 5S DNA as a starting material, we have cloned 231 bp Alu I fragments into the vector pBR322 and identified several bacterial transformants which contained plasmids with 5S DNA sequences. The insert in one recombinant plasmid, p5S10, has been sequenced and the noncoding strand is displayed in Fig. 1. By comparison with the published 5S RNA sequence of Xenopus (7,21), a putative coding region has been found; it is 55% G+C (guanosine + cytosine) and has 91% homology with the major X. laevis 5S RNA gene. Preliminary sequencing results on newt 5S RNA (not shown) and in vitro transcription experiments (13) confirm the newt 5S coding region. If the newt 5S DNA is assumed to be organized as tandemly repeating units, as data will show, then the spacer is 111 bp long, without

```

5' (AG)CTAATTCTC CAAACGTCAA TGCTGGGTTT GCTTACGGCC ATACCACCCT GAAAGCACCT GATCTCGTCT GATCTCAGAA
                                     1      10      20      30      40      50
                                     60      70      80      90      100     110     120
GTTAAGCAGG GTCGGGCTTG GTTAGTACCT GGATGGGAGA CCGCCTTGGG ATACCAGGTG CTGTAGGCTT TTGAACTTTT
TATTTTTTCC TTTAAGTGAA GTGGATTGCA TGTCCTCTTG CTATGGACAG CACAGAAGTC GAGTGAAAAA AG(CT) 3'

```

Figure 1. The nucleotide sequence of *N. viridescens* 5S DNA in p5S10. Both strands of the Eco RI-Bam HI insert of p5S10 were sequenced by the chemical cleavage method. The coding region is underlined, with positions 1 and 120 of the 5' and 3' ends noted, respectively. Only the noncoding strand is shown. The original Alu I termini are shown in parentheses, whereas the linker sequences are not shown.

internally repeating sequences, and 65% A+T (adenosine + thymine). We point out two features of the spacer: first, there is a large number of T residues downstream from the gene; second, there are some oligonucleotides which match sequences upstream from the 5' end of other 5S genes (14).

We have hybridized ³²P-p5S10 DNA to restriction enzyme digests of newt DNA. The 'ladder' pattern of hybridization is not altered by repeated digestion and is identical to that obtained when ³²P-5S RNA is used as a probe. In the Taq I and Alu I digests the smallest band of hybridization is 231 bp. These data agree with the presence of single Alu I and Taq I sites in the 5S DNA segment in p5S10. The enzyme Hae III produces some smaller genomic 5S DNA fragments and appears to cut three times per 231 bp repeat. There are only two Hae III sites in the 5S DNA of p5S10, but a third site has been found in the spacer of another single repeat plasmid, where the A in GGAC of p5S10 5S DNA is replaced by a C (12). We conclude that the insert of p5S10 is representative of the major class of newt 5S DNA.

To investigate in more detail the arrangement of 5S DNA repeat units in the *Notophthalmus* genome, we have cloned longer segments of 5S DNA by inserting Mbo I restriction fragments from a partial digest of enriched 5S DNA into the Bam HI site of pBR322 DNA. Since both of the Mbo I sites present in the 5S DNA sequence (Fig. 1) are flanked on the 3' and 5' sides by T's, insertion of these fragments into pBR322 destroys the Bam HI site. We utilized this fact as a means of selecting recombinant plasmids by cleaving the ligated DNAs with Bam HI. In Fig. 2, the Taq I digestion patterns for five such 5S DNA plasmids are shown. All five plasmid DNAs share a 231 bp fragment, although in varying amounts. Two plasmids (lanes c and d) also contain other fragments, 255 and 269 bp, which are from the

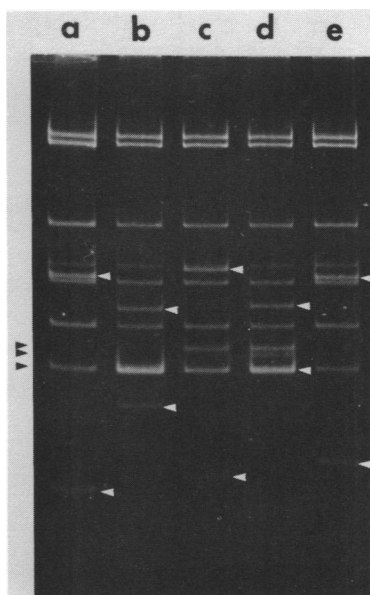


Figure 2. DNA of five recombinant plasmids that carry multiple 5S DNA repeat units cleaved with Taq I and electrophoresed in a 5% polyacrylamide gel. Fragments derived entirely from the inserts are noted on the left by black arrows. The white arrows designate fragments which carry both vector and 5S DNA sequences.

insert. Fragments which extend from Taq I sites within pBR322 to sites within the insert are designated by white arrows in the Figure. The insert sizes of the five recombinant plasmids have been determined from Msp I digests, since this enzyme does not cleave 5S DNA and the size of the longest fragments reflects the size of the insert. The plasmids contain 2, 18, 5, 15, and 2 repeat units (Fig. 2, lanes a-e).

In order to analyze the overall sequence organization within the inserts of the two largest recombinant plasmids, p5S102 (Fig. 2, lane d) and p5S 101 (Fig. 2, lane b), we used the partial digest method (25). A restriction fragment of p5S102, which contained 5S DNA sequences and was labeled with ^{32}P at one end, was partially digested by Taq I and Alu I and produced the array of fragments shown in Fig. 3. The data show that Taq I and Alu I sites are orderly spaced across the fragment with each Alu I site 10 bp closer to the fragment's labelled end than each Taq I site. This finding agrees with the fact that the Taq I and Alu I sites in the 5S DNA of p5S10 are only 10 bp apart, and suggests that the repeats are organized

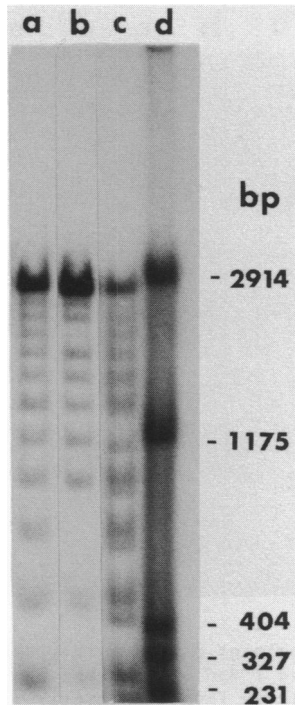


Figure 3. Gel electrophoretic analysis of partial restriction digests of the insert from p5S102 DNA. Restriction sites were mapped in a 2.8 kb Eco RI-Sal I fragment, labelled at the Eco RI end. The products of several digests were electrophoresed in a 1.4% agarose gel: (a) partial Taq I digest; (b) partial Alu I digest; (c) partial Hae III digest; (d) complete Taq I digest of ^{32}P - ϕX174 DNA. The sizes of the ϕX174 Taq I fragments are marked on the right side of the autoradiogram.

in a head-to-tail fashion. A map of the data is shown in Fig. 4. Taq I and Alu I cleave regularly every 231 bp, except that Alu I cleaves occasionally at 462 bp intervals. These 'dimers' are probably due to sequence divergence and are similar to the dimers detected in restriction digests of genomic 5S DNA. For the 18 repeat plasmid, p5S101, the Taq I and Alu I sites are likewise spaced in an orderly fashion across the insert (Fig. 5). When we did partial digests of an insert fragment which had both ends labelled, two fragment ladders were visible in the autoradiogram. The two ladders are out of register because of the different lengths of the attached vector DNA segments. At three places (the 4th, 9th, and 14th doublets) the doublet bands are unexpectedly close to one another. We have

p5S102 DNA

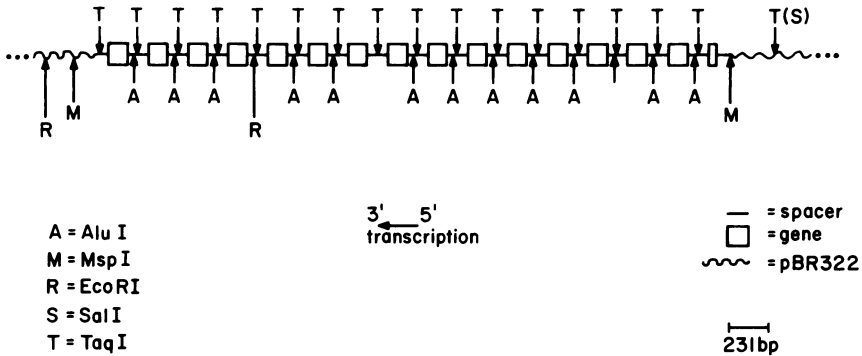


Figure 4. Map of restriction sites in the p5S102 insert. The 3.7 kb Msp I fragment of p5S102 DNA is drawn with the wavy lines, straight lines, and boxes representing vector, spacer, and gene sequences, respectively. The location of the restriction sites are noted with letters (Alu I = A, Msp I = M, Eco RI = R, Sal I = S, Taq I = T). This map represents data from Figure 4 and data not shown about the 1.3 kb Eco RI fragment. The direction of transcription is shown for the genes.

concluded that these differences in the spacing are caused by length heterogeneity of repeat units in the p5S101. This conclusion is supported by the observation that a Taq I digest of p5S101 DNA produces 231, 255, and 269 bp fragments (Fig. 2, lane d). From the above results, we have constructed a map of the arrangement of the repeat units in the insert of p5S101 (Fig. 6). Four of the 18 repeat units in p5S101 are larger than 231 bp.

We have sequenced the 231 bp Taq I fragments of p5S101, using the chemical cleavage method (18). An unambiguous sequence was established, except for the internal 10 bp Sau 3A fragment which was not analyzed. The sequence differs by only four base changes from the noncoding strand of the p5S10 insert, with three changes in the spacer and one in the gene (Fig. 7). Some variability among repeats has probably been overlooked; for instance, there are two Hha I restriction sites in the 5S DNA of p5S101 which were not observed in the DNA sequence from p5S10.

DNA sequencing was also used to determine the nature of the 269 bp repeat units of p5S101. The 269 bp Taq I fragments yielded an unambiguous sequence identical to that in the 231 bp units, except for an additional 38

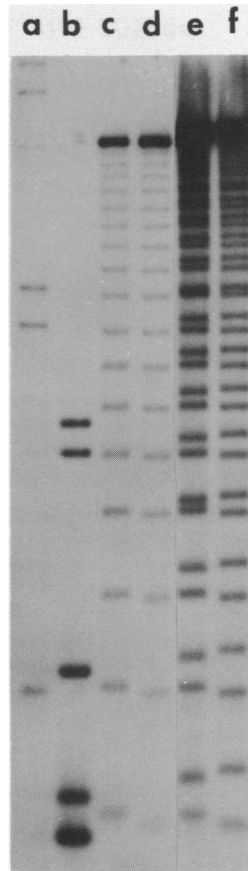


Figure 5. Gel electrophoretic analysis of partial restriction digests of the insert from p5S101 DNA. The 4.4 kb Msp I fragment was labelled at both ends with Klenow polymerase, and one half of the DNA was cleaved with Tha I to prepare a 4.2 kb Msp I-Tha I fragment that was labelled only at the Msp I end. In a 1.4% agarose gel the following digests were electrophoresed and autoradiographed: (a) lambda Hind III fragments; (b) pBR322 Taq I fragments; (c) partial Taq I digests of the 4.2 kb fragment; (d) partial Alu I digests of the 4.2 kb fragment; (e) partial Taq I digests of the 4.4 kb fragment (f) partial Alu I digests of the 4.4 kb fragment.

bp located 3' to the coding region (Fig. 7). The sequence of the additional nucleotides corresponds to a duplication of the 3' terminal third of the gene and two nucleotides of the adjacent spacer sequence. We have designated this unusual structure a pseudogene. Such a structure has also been found in another newt 5S DNA recombinant plasmid (data not

Map of Taq I sites in Msp I fragment of p5S101

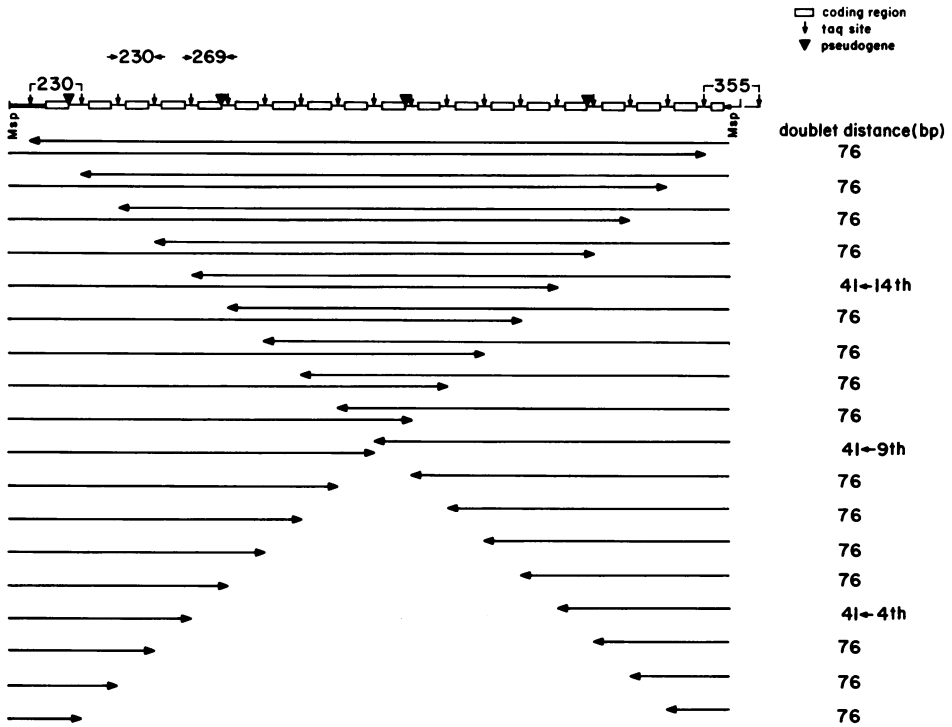


Figure 6. Schematic representation of the data shown in Figure 5. The 4.4 kb Msp I fragment of p5S101 is drawn with thick lines, thin lines, and boxes representing the vector, spacer, and gene sequences respectively. The pairs of arrows correspond to the doublet bands in the autoradiogram in Figure 6. The arrowheads represent the pseudogene and the bracketed fragments represent the terminal Taq I fragments.

shown). The structure of the 255 bp unit in p5S101 is not yet known.

To determine whether 269 bp repeat units are present in the genome, we looked for a 5S DNA unit of this size in restriction digests of total DNA. Fig. 8 shows hybridization of 5S DNA from p5S101 to restriction fragments of total DNA transferred from a polyacrylamide gel to a nitrocellulose filter. There is an intense band of hybridization at 231 bp and a weaker band which corresponds to 269 bp (lanes f,g). We conclude from this result that the 269 bp repeat units are present in the genome, but at a low frequency. We have termed these 231 and 269 bp units major and minor repeat units, respectively.

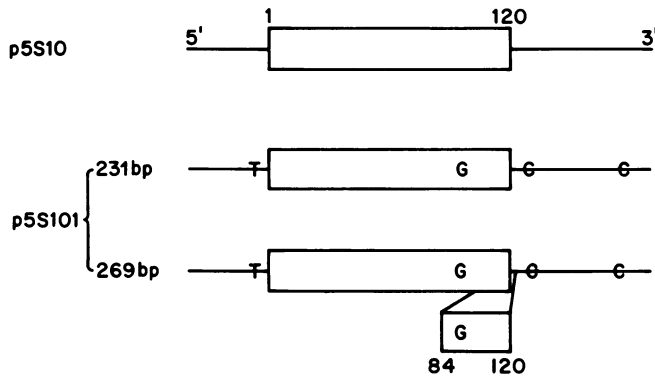


Figure 7. Comparison of p5S10 with the 231 and 269 bp units of p5S101. The boxes and thin lines represent gene and spacer, respectively. The 5' and 3' termini of the coding region are numbered 1 and 120. Nucleotides in p5S101 which differ from those in p5S10 are noted. The pseudogene in the 269 bp unit is a direct repeat of nucleotides 84 through 120.

DISCUSSION

We report here that 5S DNA in Notophthalmus is organized as tandemly repeating units of gene and spacer, without much length or sequence heterogeneity. We have described two size classes of 5S DNA repeats, one with and one without a pseudogene. While the sequence organization of 5S DNA in the newt is similar to that described for Xenopus and Drosophila 5S DNAs, the major 5S DNAs of these organisms have a great deal of length heterogeneity (6,14,26). This is due to the presence of a simple sequence which exists in variable numbers in the spacer (7,26). Such a sequence is missing in newt 5S DNA. The two classes of newt 5S DNA are also similar in organization to one of the Drosophila melanogaster satellite DNAs, where two classes of repeats are identical except for a 98 bp deletion (or insertion) in one class (5). However, both types of satellite units exist in separate tandem arrays, unlike the interspersed arrangement of the two newt 5S DNAs.

The spacer is about 240 bp shorter in newt 5S DNA than in the major 5S DNA of Xenopus laevis (7). It is of interest to note, that even though there is a 14-fold difference in genome size of Notophthalmus and Xenopus (3), this difference is not correlated with the size of the 5S spacer. This lack of correlation between genome size and spacer length is also evident in the comparison of the rDNA nontranscribed spacers of the ribosomal genes of Notophthalmus and Xenopus because the spacers are

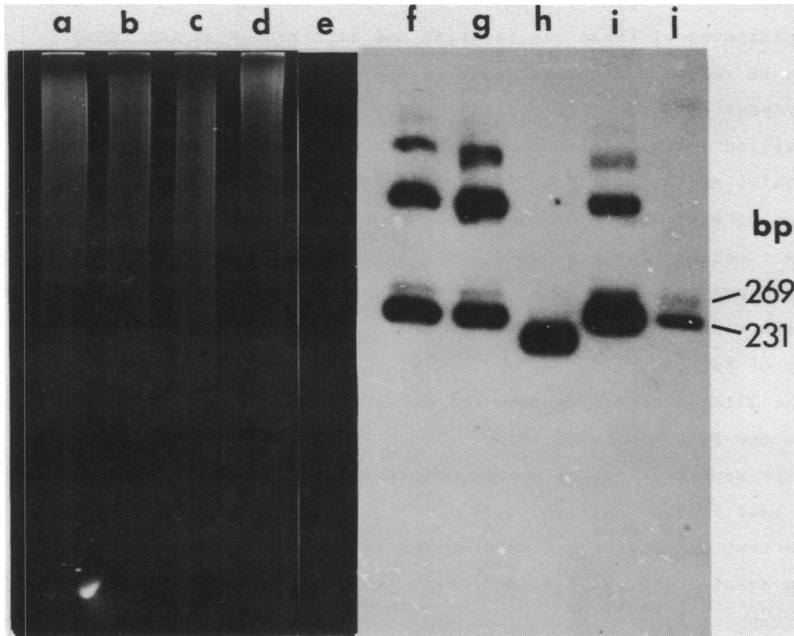


Figure 8. Detection of 269 bp units in restriction digests of genomic DNA. Staining patterns are shown for several restriction digests electrophoresed in an 8% polyacrylamide gel. In each case 3 μ g of newt DNA was digested except lane (e) which contains 10 ng of p5S101 DNA. (a) Taq I; (b) Alu I; (c) Bst NI; (d) Sau 3A. The nitrocellulose bound DNAs were hybridized with 32 P-p5S101 insert DNA (4.4 kb Msp I fragment) and autoradiographed, with lanes (f-j) corresponding to lanes (a-e), respectively. The sizes of the labelled fragments are recorded on the right. Note that the hybridizing fragments in lanes f, g and j are 231 and 269 bp, whereas the fragments in lane i are 221 and 259 bp. The band of hybridization in lane h corresponds to fragments 205 bp long, since Bst NI cleaves 5S DNA several times per repeat.

approximately the same length (20,27,12).

A 5S pseudogene has been described for both Xenopus laevis (11) and Xenopus borealis (14). The pseudogenes of Notophthalmus and Xenopus have some similar and some different features. The major similarity is their location with respect to the 5S genes: the pseudogenes are downstream from the 3' end of the genes, and have the same transcriptional polarity. One major difference between the pseudogenes is their length: they are 36, 101, and 120 bp in Notophthalmus, X. laevis, and X. borealis, respectively. One other difference is their varying degree of homology to the 5S genes: the two Xenopus pseudogenes differ by 20% from the 5S genes, whereas the

Notophthalmus pseudogene is exactly homologous in sequence to the gene. The significance of these similarities and differences is not known.

The 18 repeat recombinant plasmid p5S101 contains 4 pseudogenes, while the 15 repeat plasmid p5S102 carries none. Among the five plasmids characterized only two contain pseudogenes. In these two plasmids the 269 bp units are separated by four or five 231 bp units. The relative frequency of the pseudogenes among cloned 5S repeats is 14% (i.e. 6/42). This low frequency roughly correlates with the low frequency of the 269 bp units detected in genomic DNA (Fig. 9). From these findings we infer two arrangements of 5S DNA repeat units in the newt genome: one composed entirely of 231 bp units, and another with the 269 bp units interspersed among the 231 bp units. We have not detected any tandem arrays of 269 bp units in the recombinant plasmids. These two arrangements could be physically separated in the genome, since there are four chromosomal loci for the newt 5S RNA genes (9,22).

The newt pseudogene may have arisen either by an insertion or a deletion event. The exact site of the event cannot be defined because of the overlapping homologies between the gene, spacer, and pseudogene. The first nucleotide of the pseudogene corresponds to either nucleotide 84, 85, or 86 of the coding region. If the insertion/deletion event is the outcome of unequal crossing over, then only one of the products may be stable, as a pseudogene with only the first 84 nucleotides of the gene has not been detected in newts. The conservation of the pseudogene sequence may be a consequence of the mechanisms which maintain sequence and length homogeneity in repeated gene families.

The presence of the pseudogene in a 5S DNA repeat does not prevent some correct transcription from the adjacent gene. We have reported experiments which show that the 269 bp units can produce 5S RNA transcripts in an in vitro transcription system (13). However, the same DNA also directs the synthesis of some longer transcripts, the result of continuation of transcription through the end of the gene with termination at the end of the pseudogene. It is not known whether this sort of 'piggy-back' transcription occurs in vivo.

ACKNOWLEDGEMENTS

This paper is based on a dissertation submitted by B.K.K. in partial fulfillment of the requirements for the Ph.D. degree at Yale University. Supported by NIH Research Grant GM 12427. We thank N. Kan and R. Steele

for helpful discussions.

*Current address: Laboratory of Biochemistry, National Cancer Institute, Bethesda, MD 20205, USA

REFERENCES

- 1 Barnett, T. and Rae, P.M.M. (1979). *Cell* 16, 763-775.
- 2 Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heyneker, H.L. and Boyer, H.W. (1977). *Gene* 2, 95-113.
- 3 Brown, D.D. and Dawid, I.B. (1968). *Science* 160, 272-280.
- 4 Brown, D.D. and Weber, C.S. (1968). *J. Mol. Biol.* 34, 661-680.
- 5 Carlson, M. and Brutlag, D. (1977). *Cell* 11, 371-382.
- 6 Carroll, D. and Brown, D.D. (1976). *Cell* 7, 467-475.
- 7 Fedoroff, N.V. and Brown, D.D. (1978). *Cell* 13, 701-716.
- 8 Gergen, J.P., Stern, R.H. and Wensink, P. (1979). *Nucl. Acid Res.* 7, 2115-2136.
- 9 Hutchison, N. and Pardue, M.L. (1975). *Chromosoma* 53, 51-69.
- 10 Inoue, M. and Curtiss, R. III (1977). In: *Molecular Cloning of Recombinant DNA*, Miami Winter Symposia (W.A. Scott and R. Werner, eds), Vol. 13, p. 248.
- 11 Jacq, C., Miller, J.R. and Brownlee, G.C. (1977). *Cell* 12, 109-120.
- 12 Kay, B. (1980). Ph.D. Thesis, Yale University.
- 13 Kay, B.K., Schmidt, O. and Gall, J.G. (1981). *J. Cell Biol.* 90, 323-331.
- 14 Korn, L.J. and Brown, D.D. (1978). *Cell* 15, 1145-1156.
- 15 Long, E.O. and Dawid, I.B. (1980). *Ann. Rev. Biochemistry* 49, 727-764.
- 16 Maizel, N. (1976). *Cell* 9, 431-438.
- 17 Maniatis, T., Hardison, R.C., Lauer, J., O'Connell, C., Sim, G.K. and Efstratiadis, A. (1978). *Cell* 15, 687-701.
- 18 Maxam, A. and Gilbert, W. (1977). *Proc. Nat. Acad. Sci. USA* 74, 560-564.
- 19 Maxam, A. and Gilbert, W. (1979). In: *Methods in Enzymology* (eds. Grossman, L. and Moldave, K.) vol. 65, 499-560.
- 20 Miller Jr. O.L. and Hamkalo, B.A. (1972). *Int. Rev. Cytology.* 33, 1-25.
- 21 Miller, J.R., Cartwright, E.M., Brownlee, G.G., Fedoroff, N.V. and Brown, D.D. (1978). *Cell* 13, 717-725.
- 22 Pukkila, P.J. (1975). *Chromosoma* 53, 71-89.
- 23 Sanger, F. and Coulson, R. (1978). *FEBS Lett.* 87, 107-110.
- 24 Schultz, L.D., Kay, B.K. and Gall, J.G. (1981). *Chromosoma* 82, 171-187.
- 25 Smith, H.O. and Birnstiel, M.L. (1976). *Nucl. Acids Res.* 3, 2387-2398.
- 26 Tschudi, C. and Pirrotta, V. (1980). *Nucl. Acids Res.* 8, 441-451.
- 27 Wellauer, P.K., Dawid, I.B., Brown, D.D. and Reeder, R.G. (1976). *J. Mol. Biol.* 105, 461-486.