

# The Two Embryonic U1 Small Nuclear RNAs of *Xenopus laevis* Are Encoded by a Major Family of Tandemly Repeated Genes

ELSEBET LUND,<sup>1</sup> JAMES E. DAHLBERG,<sup>1\*</sup> AND DOUGLASS J. FORBES<sup>2†</sup>

*Department of Physiological Chemistry, University of Wisconsin-Madison, Madison, Wisconsin 53706,<sup>1</sup> and Department of Biophysics and Biochemistry, University of California-San Francisco, San Francisco, California 94143*

Received 3 July 1984/Accepted 30 August 1984

**We have identified a large family of U1 RNA genes in *Xenopus laevis* that encodes two distinct species of U1 RNA. These genes are expressed primarily at the onset of transcription in the 4,000-cell embryo (D. J. Forbes, M. W. Kirschner, D. Caput, J. E. Dahlberg, and E. Lund, *Cell* 38:681-689, 1984). The two types of embryonic U1 RNA genes are interspersed and are organized in large tandem arrays. The basic 1.9-kilobase repeating unit contains a single copy of each of the embryonic genes and is reiterated ca. 500-fold per haploid genome. This repetitive U1 DNA accounts for more than 90% of all U1 DNA in *X. laevis*. In addition to this major family, there exist several minor families of dispersed U1 RNA genes, which presumably encode the oocyte and somatic species of *X. laevis* U1 RNA. Although the embryonic genes are normally inactive in stage VI oocytes, they are expressed when cloned copies are injected into oocyte nuclei.**

The U1 small nuclear RNA (U1 RNA), which is ubiquitous in higher eucaryotes, is encoded by multigene families. The estimates for gene copy number range from ca. 5 to 50 per haploid genome in *Drosophila* (S. M. Mount, Ph.D. thesis, Yale University, New Haven, Conn., 1983) sea urchin (5), chicken (24), rodent (16, 30), and human (9, 14, 15) DNAs. In spite of this multiplicity of genes, only one or two U1 RNA species have been detected in each of these organisms. In vertebrates, the majority of the U1 RNA genes appears to be dispersed in the genome but, at least in the case of humans, they are clustered in a small region of a single chromosome (13). In the sea urchin genome, all of the genes for N1 RNA (the equivalent of U1 RNA) are arranged in a single large tandem array (5). Mammalian genomes also contain several classes of abundant U1 RNA pseudogenes (8, 9, 15, 18, 21), whereas the sea urchin genome apparently does not (5).

It has been reported elsewhere that *Xenopus laevis* DNA encodes at least seven different U1 RNA species and that the transcription of several U1 RNA genes is under developmental control (10a). That is the first demonstration of controlled expression of U1 RNA genes. In particular, we find that two species of *Xenopus* U1 RNA which we call embryonic U1 RNAs, xU1a and xU1b, are synthesized at the onset of transcription in the 4,000-cell embryo but not in fully grown oocytes. At least four other U1 RNAs are synthesized during late oogenesis and in somatic cells.

As an initial step in the study of the mechanism of the developmental control of *X. laevis* U1 RNA synthesis, we have analyzed the structure and organization of the genes that are responsible for the synthesis of xU1a and xU1b RNAs early in embryogenesis. We report here that these genes are present in ca. 500 tandemly repeated units per haploid genome; the repeat unit contains one copy of each of the embryonic U1 RNA genes. Together, these two types of genes constitute the major U1 gene family, accounting for

more than 90% of all U1 RNA genes in *X. laevis*. The remaining 5 to 10% of the genes appear to be dispersed in the genome, analogous to the gene organization in other vertebrates, and may encode oocyte and somatic species of U1 RNAs (32; this study). Thus the U1 RNA genes expressed at various times of development differ both in number and in genomic arrangements. These different gene arrangements might be responsible, at least in part, for the developmental control of U1 RNA synthesis in *X. laevis*.

## MATERIALS AND METHODS

**Analysis of DNA.** *X. laevis* genomic DNA was prepared by standard procedures from liver or erythrocytes (the latter was a generous gift of D. D. Brown). DNA fragments were fractionated on 0.7 to 1.2% agarose gels or 6% (60:1) polyacrylamide gels and transferred to nitrocellulose filters as described by Southern (27) or by the bidirectional method of Smith and Summers (26). DNA blots were probed with <sup>32</sup>P-labeled human U1 DNA which was either 3'-end-labeled pU1.15b DNA (kindly provided by A. Weiner) or single-stranded cDNA made of M13mp7 U1-C DNA (14). Hybridization conditions were 5× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate; pH 7.0)-0.1% sodium dodecyl sulfate-1 mM EDTA-20 mM Tris-hydrochloride (pH 7.0)-5× Denhardt solution (7)-50% formamide at 37, 42, 50, or 55°C for 20 to 30 h (14).

**Cloning of *X. laevis* U1 RNA genes.** The bacteriophage λ Charon 4A library of *X. laevis* embryo DNA (28) was screened by the method of Benton and Davis (2) with the pU1.15b human U1 DNA probe (see above). All of the five U1 DNA clones (AX1U1-G-1 through -5) obtained by screening a total of 500,000 plaques were found to contain minor *X. laevis* U1 RNA genes (see Results).

To obtain *X. laevis* genomic DNA enriched in the major U1 RNA genes, total *X. laevis* (liver or erythrocyte) DNA was digested with a mixture of *Bam*HI, *Bgl*III, *Kpn*I, and *Xba*I (or *Bam*HI, *Bgl*III, and *Eco*RI) and size fractionated on 5 to 20% NaCl gradients (23). Fractions containing large (>25 kilobases [kb]) U1 DNA fragments (identified by

\* Corresponding author.

† Present address: Department of Biology, University of California-San Diego, San Diego, CA 92096.

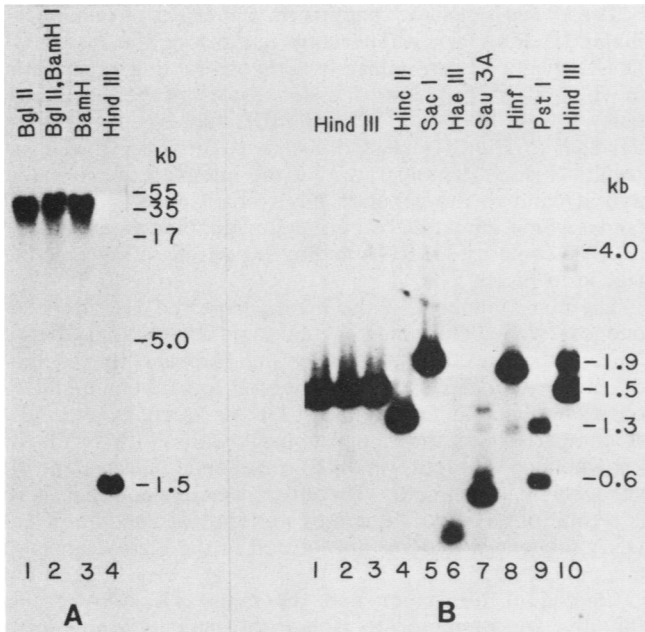


FIG. 1. Characterization of *X. laevis* genomic U1 DNA sequences. *X. laevis* liver DNA was digested with different restriction enzymes and analyzed by Southern blot hybridization. In addition to *Bam*HI and *Bgl*III (lanes 1 to 3 of A), the restriction enzymes *Eco*RI, *Kpn*I, *Xba*I, *Xho*I, *Sal*I, and *Sac*II were all found to produce 30 kb of U1 DNA fragments (data not shown). Lanes 1 to 3 in B show the *Hind*III digests of genomic DNAs from three additional frogs and lane 10 shows a partial *Hind*III digest. Genomic DNA (10  $\mu$ g) was loaded in each lane, and the digestion products were fractionated in 0.7% agarose gels. After transfer to nitrocellulose filters, hybridization to  $^{32}$ P-labeled human U1 DNA (pU1.15b) was at 42C in 50% formamide. Autoradiograms are shown.

Southern blot hybridization) were pooled, and the DNA was redigested with *Sac*I or *Hind*III to generate unit-length fragments of the U1 repeat DNA. A phage  $\lambda$  library was constructed by the addition of *Eco*RI linkers to the *Sac*I fragments and insertion of these fragments into the *Eco*RI site of  $\lambda$ gt10 DNA. Six independent isolates of the major U1 RNA genes ( $\lambda$ X1U1-S-1 through -6) were obtained by screening 5,000 plaques from this library. Eleven additional isolates of these genes (pX1U1-H-1 through -11) were obtained by cloning in pBR322. The 1.5-kb *Hind*III U1 DNA fragments were purified from the *Hind*III-digested enriched DNA by preparative agarose gel electrophoresis and cloned into pBR322 by transformation of strain HB101. Ampicillin-resistant cells were screened by colony hybridization (12), with human U1 cDNA probes (see above).

**Expression of *X. laevis* U1 RNA genes.** *X. laevis* oocyte injections were carried out as previously described (19) except that DNA templates were injected as supercoiled plasmid DNAs (pX1U1-H clones) or as full-length  $\lambda$ DNAs ( $\lambda$ X1U1-S and  $\lambda$ X1U1-G clones). Total nucleic acids were extracted from the oocytes 20 h after injection and analyzed by polyacrylamide gel electrophoresis either directly or after preparative hybridization to filter-bound human U1 DNA as described elsewhere (19). Total RNAs were analyzed in 8% (30:0.8) polyacrylamide gels containing 7 M urea (19), whereas hybrid-selected U1 RNAs were displayed on 15% (19:1) nondenaturing polyacrylamide gels (10a, 13). Purified RNAs were analyzed by RNase T<sub>1</sub> fingerprinting (25) with homomix C (1).

## RESULTS

**The major U1 RNA genes in *X. laevis* DNA are organized in large tandem arrays.** When the genomic DNA of *X. laevis* was assayed by Southern blot analysis, with human U1 RNA genes as probes, a single strongly hybridizing band (U1 DNA) was observed (Fig. 1). The U1 DNA migrated either as undigested (>50 kb) fragments or as a relatively small (<2.0 kb) fragment, depending on which enzymes were used to digest the genomic DNA. This pattern is indicative of a tandemly repeated sequence. In this case, the size of the repeat unit was estimated to be at least 1.9 kb, based on the mobility of the fragments produced by complete digestion with *Sac*I or by partial digestion with other enzymes such as *Hind*III (Fig. 1B, lanes 5 and 10).

The tandem organization of the major U1 RNA gene repeat was established by partial digestion of the genomic DNA with either *Sac*I (Fig. 2) or *Hind*III (data not shown). In both cases, a ladder of bands of U1 DNA was obtained, and the mobilities of the partial digestion products indicated that they were multimers of the 1.9-kb *Sac*I fragment. The results (Fig. 2) demonstrated that the *X. laevis* genome contained multiple, tandemly repeated 1.9-kb units of U1 DNA. Similar analysis of partial *Hind*III digests in a 0.4% agarose gel indicated that the tandem array(s) contained at least 20 repeat units (data not shown).

Several of the fragments (Fig. 1B) were smaller than 1.9 kb, presumably as a result of more than one cleavage of the repeat unit. In most cases only one of the resulting fragments was capable of hybridizing the U1 RNA gene probe, but both of the enzymes *Pst*I (lane 9) and *Pvu*II (data not shown) generated two fragments of U1 DNA, the sizes of which added up to 1.9 kb (Fig. 1B). The fact that the two fragments generated by either *Pst*I or *Pvu*II each hybridized approximately half as well as the larger fragments (see also Fig. 3) suggested that the repeat contained two U1 RNA genes which were separated by *Pst*I and *Pvu*II cleavage sites. This conclusion was confirmed by restriction mapping of cloned *X. laevis* U1 DNA fragments, as described below.

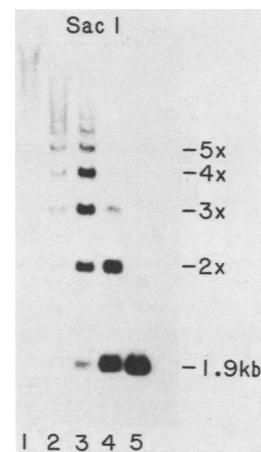


FIG. 2. Tandem organization of the major repeated U1 RNA genes of *X. laevis*. *X. laevis* liver DNA was subjected to partial digestion with restriction endonuclease *Sac*I and analyzed by Southern blot hybridization as in Fig. 1. Samples containing 10  $\mu$ g of DNA were withdrawn after 0, 1, 3, 10, or 30 min of digestion with 2 U of *Sac*I per  $\mu$ g of genomic DNA (lanes 1 to 5). The single major band in lane 5 corresponds to 1.9-kb-long U1 DNA fragments, and the ladder of bands observed in lanes 2 to 4 corresponds to integral multimers of the basic 1.9-kb repeat DNA, as indicated.

In addition to the major intense band of hybridization, each digest of genomic DNA contained several minor bands which were detectable only upon longer exposures of the autoradiograms. Since each restriction enzyme produced a distinct pattern of minor bands (Fig. 1 and 3B), it appeared that the corresponding U1 DNA sequences were dispersed in the genome rather than being tandemly arranged. A few of these bands might represent U1 DNA fragments derived from the junctions between the ends of the tandem arrays and the adjacent sequences. However, other minor bands were found to correspond to U1 RNA genes which differed from the major U1 repeat DNA (see below). Few, if any, of these minor bands represented U1 RNA pseudogenes, since the patterns of bands were not influenced by the stringency of hybridization, i.e., no differences were observed between patterns generated by hybridization at 37 or 55°C (data not shown).

The same patterns and size classes of fragments were produced by analysis of genomic DNAs isolated from the liver (Fig. 1) or erythrocytes (Fig. 3). Furthermore, no differences were observed between either the major or minor bands of U1 DNA in genomic DNAs obtained from individual frogs (cf. lanes 1, 2, and 3 of Fig. 1B).

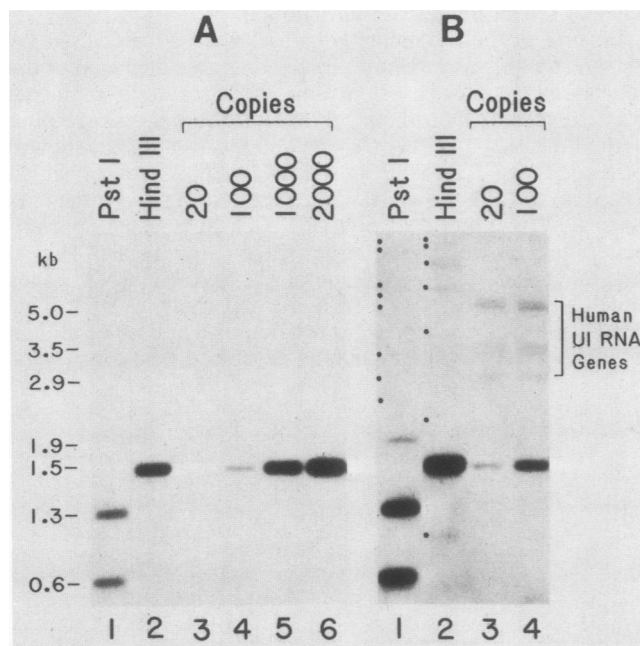


FIG. 3. Determination of gene copy numbers of the major and minor U1 RNA genes of *X. laevis*. Samples (4  $\mu$ g) of *X. laevis* erythrocyte were digested with *Pst*I (lanes 1) or *Hind*III (lanes 2) and fractionated in a 1.2% agarose gel. Copy number standards for the major U1 DNA repeat were provided by adjacent lanes (3 to 6) containing different amounts of *Hind*III-digested pX1U1-H1 DNA (cf. Fig. 4C and 5). Autoradiograms of the nitrocellulose filter exposed for 20 (A) or 100 (B) h are shown. The dots in lanes 1 and 2 of B indicate the position of minor U1 hybrid bands. Assuming that the haploid genome of *X. laevis* contains  $3 \times 10^9$  base pairs (6), the amounts of pX1U1-H1 DNA loaded corresponded to 20, 100, 1,000, and 2,000 copies of U1 RNA coding region sequences per haploid genome (lanes 3 to 6). Lanes 3 to 6 also contained 4  $\mu$ g of *Hinc*II-digested human fibroblast DNA as a carrier and as an internal control for low copy numbers of U1 RNA genes. The human DNA fragments indicated corresponded to 5 to 15 copies of U1 RNA genes per haploid genome (14). The nitrocellulose filter was probed with  $^{32}$ P-labeled human U1 DNA (HU1-1C) in 50% formamide at 50°C.

The *X. laevis* genome contains ca. 500 copies of each of the major U1 RNA genes. The copy number of the major U1 DNA repeat was determined by a reconstruction experiment in which various amounts of a cloned copy of the repeat (see below) were digested with *Hind*III and electrophoresed adjacent to *Hind*III-digested *X. laevis* genomic DNA. The results (Fig. 3) demonstrate that the intensity of the major hybrid band in the genomic DNA corresponds to ca. 500 copies of the repeat DNA per haploid genome equivalent or to 1,000 copies of U1 RNA coding region sequences (see the legend to Fig. 3).

The copy numbers of the minor dispersed U1 DNA sequences were determined in the same experiment after a longer exposure of the autoradiogram. It appears that each of the minor *X. laevis* bands corresponds to between 5 and 20 copies per haploid genome of U1 DNA sequences (Fig. 3B). A comparison of the minor hybrid bands with bands of known copy numbers in human genomic DNA (lane 4) supports these estimates. Therefore, we conclude that there is a total of ca. 40 to 50 copies per haploid genome of U1 DNA sequences that are not located in the highly repeated tandem array.

**Cloning of the major and the minor U1 RNA genes.** Initially, we attempted to isolate the major 1.9-kb repeat DNA from a genomic bank of *X. laevis* embryo DNA cloned in bacteriophage  $\lambda$  (28), but from ca. 500,000 plaques screened, no representatives of the highly repeated U1 DNA were isolated. We did, however, obtain five clones ( $\lambda$ X1U1-G) which upon digestion with *Hind*III produced a 5-kb U1 DNA fragment (Fig. 4A), presumably corresponding to one of the minor fragments observed in *Hind*III-digested genomic DNA (cf. Fig. 3). Zeller et al. (32), who reported the isolation of a clone very similar (or identical) to these clones, also were unable to obtain clones of the major U1 repeat from this genomic library. Analysis of total DNA isolated from this particular genomic library indicated that it was deficient in recombinant phages carrying the major U1 RNA gene repeat (data not shown).

To clone the major DNA repeat, we prepared *X. laevis* DNA enriched in these sequences by size fractionation of total genomic DNA which had been digested with enzymes that do not cleave the tandem array. Unit-length U1 repeat DNA was then released from the large (>50 kb) DNA fragments by cleavage with *Sac*I or *Hind*III and was cloned into  $\lambda$  or pBR322 DNA, respectively.

The cloned *Sac*I fragments of *X. laevis* U1 DNA were identical in size to that of the 1.9-kb repeat units of genomic DNA (Fig. 4B). In each  $\lambda$  phage DNA preparation, however, a various amount of DNA was observed that contained a smaller insert; apparently, this smaller insert resulted from the loss (by homologous recombination) of 0.6 kb of DNA between two U1 RNA coding regions (data not shown).

Cloning of the 1.5-kb *Hind*III fragments of the U1 repeat DNA in plasmid pBR322 resulted in the isolation of stable recombinant plasmids that showed no evidence of accumulating deletions upon propagation. Figure 4C shows four representative isolates of these clones. Although multiple independent isolates of these sequences appeared homogeneous in size (Fig. 4C), fine structure analyses revealed several regions of minor sequence heterogeneity (Fig. 4D). As indicated on the physical map of the U1 repeat unit DNA, these heterogeneities corresponded to small deletions (15 to 25 base pairs) located both outside and within the coding region of one of the U1 RNA genes (thin vertical arrows in Fig. 5).

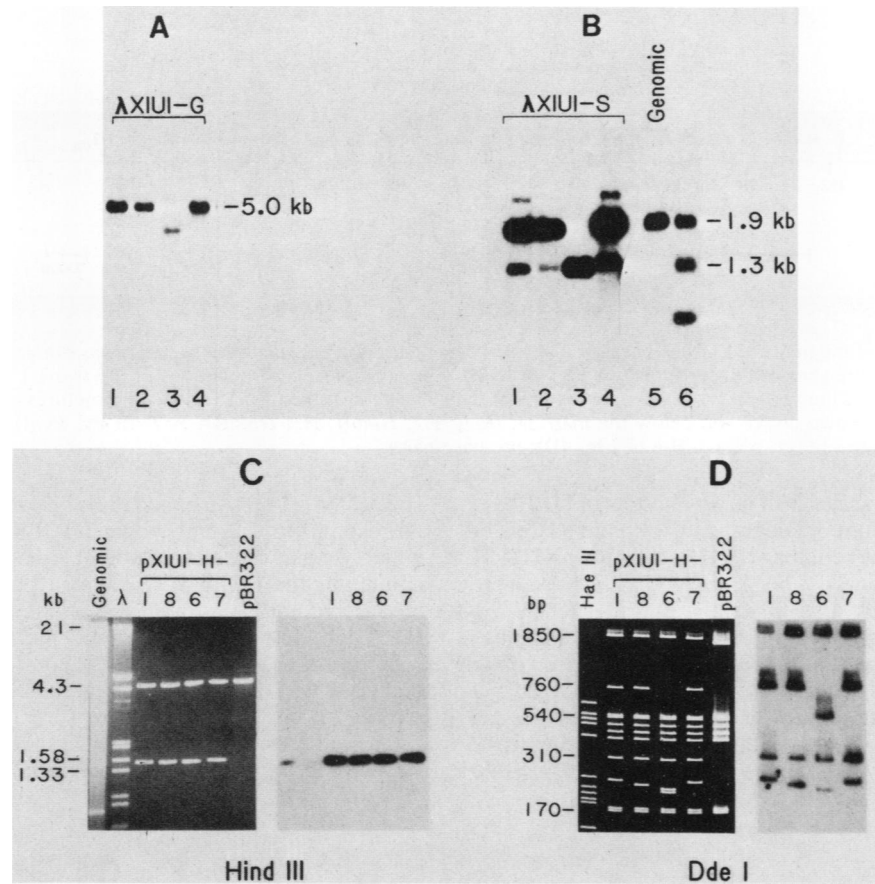


FIG. 4. Characterization of cloned copies of the major and minor U1 RNA genes of *X. laevis*. Three types of clones carrying *X. laevis* U1 DNA fragments were characterized by restriction enzyme analyses: (A) recombinant phages ( $\lambda$ X1U1-G) from the genomic library of Wahli and Dawid (28); (B)  $\lambda$  recombinant phages ( $\lambda$ X1U1-S) from a bank enriched in the full-length 1.9kb-long U1 DNA repeat; and (C and D) pBR322 recombinant plasmids (pX1U1-H) carrying 1.5 kb of the major U1 DNA repeat. (A) Analysis of *Hind*III digests of  $\lambda$ X1U1-G DNAs obtained from plaque-purified isolates of three independent clones (lanes 1 to 4). Southern blot hybridization was performed as in Fig. 1; an autoradiogram is shown. The different sizes of the U1 DNA fragments in independent isolates of one clone (lanes 3 and 4) indicate that DNA sequences within this fragment were lost upon phage propagation. (B) Analysis of *Eco*RI digests of  $\lambda$ X1U1-S DNAs obtained from four independent clones (lanes 1 to 4). A *Sac*I digest of genomic *Xenopus* liver DNA (lane 5) and a limit *Eco*RI-*Pst*I digest of  $\lambda$  DNA from the recombinant phage of lane 1 (lane 6) are shown. The U1 DNA fragments were detected as in Fig. 1A; an autoradiogram is shown. The presence of smaller U1 DNA inserts in all phage DNA preparations suggests that sequences within the U1 repeat DNA are frequently lost. (C and D) Analyses of four independent isolates of pX1U1-H clones. Ethidium-bromide stained gels (left) and corresponding autoradiograms (right) of the U1 DNA fragments are presented. Southern blot hybridizations were performed as in Fig. 3 with the human U1 RNA coding region probe. (C) *Hind*III digests of DNAs from clones 1, 8, 6, and 7 are shown. *Hind*III-digested *X. laevis* erythrocyte DNA (genomic), *Hind*III plus *Eco*RI-digested  $\lambda$ DNA ( $\lambda$ ), and *Hind*III-digested pBR322 DNA (pBR322) were included as size markers in the 1% agarose gel. (D) *Dde*I digests of the same four DNAs, analyzed in a 6% polyacrylamide gel. The size markers were pBR322 DNA digested with *Hae*III (*Hae*III) or *Dde*I (pBR322). The observed sequence heterogeneities define three classes of 1.5-kb *Hind*III U1 repeat DNA fragments (clones 1 and 7 belong to the same class). The presence (and the sizes) of the four *Dde*I U1 DNA fragments indicates that the U1 repeat DNA contains two U1 RNA genes.

Southern blot analysis, with a variety of restriction enzymes, confirmed that each repeat unit contained two different U1 RNA gene coding regions (Fig. 4D and data not shown). The orientations of the two U1 RNA coding regions were determined by the relative positions of restriction enzyme cleavage sites known to be located within the coding regions of U1 RNA genes from RNA sequence data (3, 10a, 14).

The U1 RNA genes within the repeat DNA were tentatively identified as coding for xU1a and xU1b embryonic RNAs (10a) on the basis of their restriction cleavage sites. Specifically, both U1 RNA coding regions contained a *Dde*I cleavage site (CTCAG) at the position corresponding to nucleotides 59 to 63 in the RNAs, and sequences corresponding to this site were found to be hallmarks of the two

embryonic U1 RNAs (cf. Fig. 3 and Table 1 of reference 10a). Furthermore, the *Sal*I cleavage site (GTCGAC, corresponding to positions 115 to 120 of xU1b RNA) unique to one of the coding regions (gene B, Fig. 5) was indicative of a gene for xU1b RNA.

The embryonic U1 RNAs, xU1a and xU1b, are encoded by the U1 RNA genes of the major repeat unit. The suggestion that both embryonic U1 RNA species were encoded by the major U1 repeat DNA was tested directly by injection of the cloned DNA into the nuclei of mature *X. laevis* oocytes. In this experiment, the cloned DNAs from seven independent isolates of the 1.5-kb *Hind*III fragment were coinjected with [ $\alpha$ - $^{32}$ P]GTP; 20 h later, total RNA was isolated and analyzed by polyacrylamide gel electrophoresis (Fig. 6). Clearly, injection of each of the cloned DNAs of the *X. laevis* U1

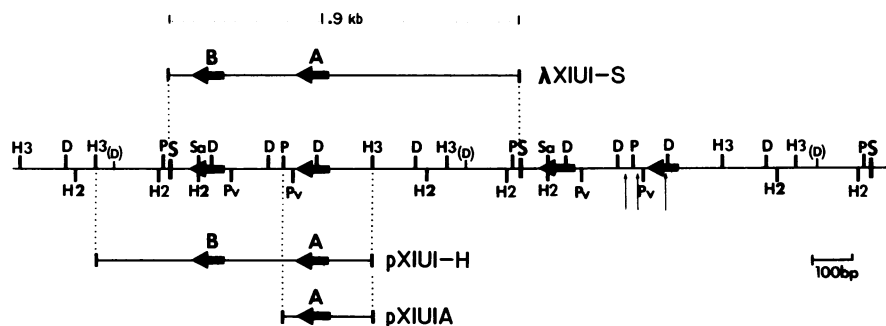


FIG. 5. Physical map of the major U1 DNA repeat of *X. laevis*. The restriction enzyme cleavage map was obtained by analyses of cloned copies of the repeat DNA by standard procedures (cf. Fig. 4B to D). The positions and orientations of the two U1 RNA coding regions are indicated by heavy arrows. Thin vertical arrows denote regions of sequence heterogeneities. The general structures of the X1U1-DNA clones discussed in the text are shown above and below the map. D, *DdeI*; H2, *HincII*; H3, *HindIII*; P, *PstI*; Pv, *PvuII*; S, *SacI*; Sa, *Sall*; (D), polymorphic *DdeI* cleavage site (cf. pX1U1-H-6 of Fig. 4D); bp, base pairs.

repeat resulted in the synthesis of large amounts of U1 RNA, comparable to the amount obtained by injection of a clone of a human U1 RNA gene (compare pHU1-1D with pX1U1-H [lanes 1 to 7] in Fig. 6A). The *X. laevis* U1 RNAs were further characterized by analysis of the hybrid-selected U1 RNAs in a nondenaturing 15% polyacrylamide gel. (The use of this gel system for the separation of different species of *X. laevis* U1 RNAs is discussed in detail elsewhere [10a].) From electrophoretic mobilities of the U1 RNAs in this gel (Fig. 6B) and from the RNase T<sub>1</sub> fingerprints of the individual gel-purified RNAs (Fig. 6C), we concluded that the two

species of U1 RNAs encoded by the repeat unit did, in fact, correspond to the embryonic U1 RNAs, xU1a and xU1b, which we had recently identified (10a). The λX1U1-S clone, containing the 1.9-kb *SacI* fragment, was also transcriptionally active and was found to encode embryonic U1 RNAs (data not shown). As predicted, injection of the DNA of a subclone of the major U1 repeat containing only the A gene coding region (pX1U1A; cf. Fig. 5) resulted in the synthesis of only xU1a RNA (Fig. 6B).

**Oocyte and somatic U1 RNAs are encoded by less abundant U1 RNA genes.** Injection of the cloned DNA containing the

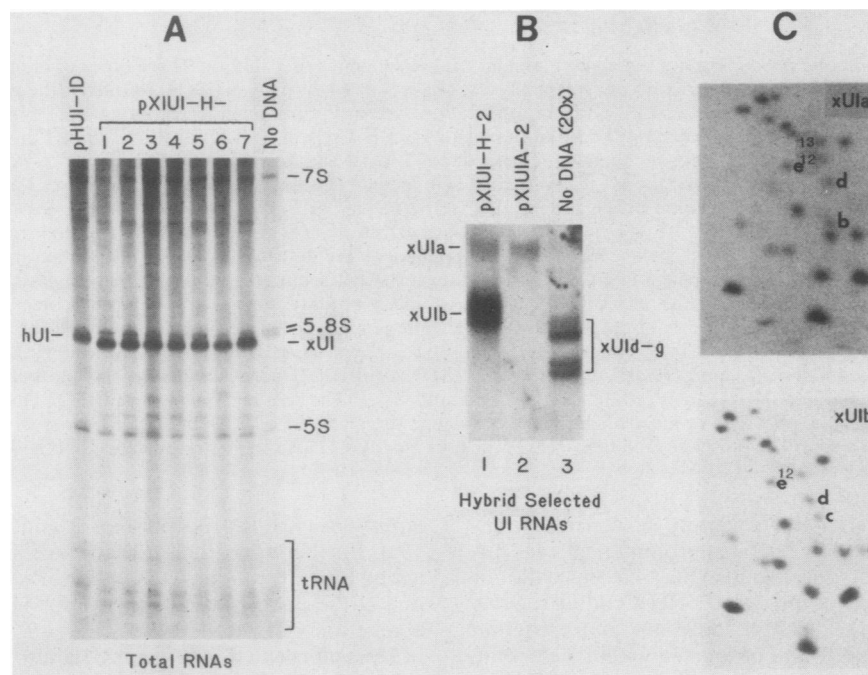


FIG. 6. Expression of the embryonic U1 RNAs encoded by the major U1 DNA repeat in *X. laevis* oocytes. Cloned copies of the 1.5-kb *HindIII* fragments (pX1U1-H DNAs, cf. Fig. 5) were coinjected with [ $\alpha$ -<sup>32</sup>P]GTP into the nuclei of *X. laevis* stage V to VI oocytes. Total nucleic acids were prepared after 20 h of incubation and analyzed by electrophoresis in an 8% (30:0.8) polyacrylamide gel containing 7 M urea (A). Each lane contains total RNA of two oocytes; the injected DNAs were pX1U1-H-1 through -7 and pHU1-1D, a cloned human U1 RNA gene (19). (B) Hybrid-selected U1 RNAs from 1 to 2 oocytes injected with pX1U1-H-2 (lane 1) or pX1U1A-2 (lane 2) or from 20 oocytes which received no DNA (lane 3) were fractionated in a 15% nondenaturing polyacrylamide gel. (C) The individual U1 RNAs of lanes 1 and 2 were eluted and analyzed by RNase T<sub>1</sub> fingerprinting, revealing that the two U1 RNAs encoded by the major U1 repeat DNA were the embryonic species, xU1a and xU1b, as indicated. The lower case letters (a to e) and numbers (12 and 13) refer to RNase T<sub>1</sub> oligonucleotides which were found to be characteristic of the two embryonic U1 RNAs as shown in Fig. 2 and in Table 1 of reference 10a.

minor 5-kb *Hind*III U1 DNA fragment (i.e., DNA from  $\lambda$ X1U1-G, cf. Fig. 4A) also led to the synthesis of U1 RNA. This U1 RNA was shown by RNase T<sub>1</sub> fingerprint analysis to be one of the minor U1 RNA species synthesized in *X. laevis* tissue culture cells and stage VI oocytes rather than either of the embryonic species xU1a and xU1b (data not shown). This finding is in agreement with the sequences of two minor U1 RNA genes from a similar clone as determined by Zeller et al. (32).

#### DISCUSSION

We have demonstrated that *X. laevis* contains ca. 1,000 copies of the genes for U1 RNA per haploid genome and that two different embryonic U1 RNA genes make up more than 90% of the U1 DNA sequences. These embryonic U1 RNA genes are organized in one or more large arrays, with a total of ca. 500 tandemly repeated units. Each 1.9-kb repeat unit encodes both of the two U1 RNAs that are synthesized during early embryogenesis (10a).

The major and minor gene families encode different species of U1 RNAs. In addition to the major family of highly repeated U1 RNA genes, the *X. laevis* genome also contains several minor families of dispersed U1 RNA genes that are present in less than 50 copies per haploid genome. Similar results have been obtained independently by Zeller et al. (32), who recently reported the isolation of one of the minor 5-kb *Hind*III fragments of *X. laevis* U1 DNA. That clone, which contains three genes for different species of U1 RNAs, might well correspond to the  $\lambda$ X1U1-G clones described here (cf. Fig. 4A). Thus it appears that at least one of the minor families of U1 RNA genes encodes somatic U1 RNAs. We infer that the dispersed genes must be responsible for the synthesis of all of the U1 RNAs specific to late oocytes and somatic cells since the repetitive U1 RNA genes encode the embryonic U1 RNAs. A possibility, which cannot be ruled out at the present time however, is that a small fraction (<10%) of the major U1 DNA repeat encodes oocyte or somatic U1 RNA species. This latter possibility seems highly unlikely in view of the sequence homogeneity of the cloned repeats. In particular, we note that the U1 RNAs encoded by seven different clones of the repeat appeared identical by RNase T<sub>1</sub> fingerprinting (data not shown) and that restriction enzyme cleavage sites were highly conserved between independent isolates of the repeating unit.

*X. laevis* U1 and U2 RNA genes show striking homology in organization. Recently, Mattaj and Zeller (17) described the organization of the U2 RNA genes in *X. laevis*. Again, the major U2 RNA genes were found to be tandemly repeated (in this case with a single copy of the U2 DNA per 0.85-kb repeating unit), and minor U2 RNA genes appeared to be dispersed rather than tandemly arranged. Because of the similar organization of the U1 and U2 RNA genes, it is appealing to speculate that the expression of the highly repeated genes is coordinately controlled during development. Unfortunately, there is no evidence for this idea since little information is available about the synthesis of U2 RNA in *X. laevis*. However, RNase T<sub>1</sub> fingerprint analysis of U2 RNA synthesized during early embryogenesis supports the hypothesis that the embryonic U2 RNA is encoded by the abundant reiterated U2 RNA genes (E. Lund and D. Forbes, unpublished data).

**Differential control of expression of RNA genes in *X. laevis*.** For several years the oocyte and somatic 5S RNA genes of *Xenopus* spp. have provided a model system for studying developmental control of the expression of two classes of

reiterated RNA genes (4, 11, 22). The presence of 20,000 tandemly arranged oocyte-specific 5S RNA genes allows for the high rate of 5S RNA synthesis that is required during oogenesis.

It is possible that a high copy number of the embryonic U1 RNA genes is required in a similar way for the very efficient U1 RNA synthesis at the onset of transcription in the 4,000-cell blastula embryo. At this stage of embryogenesis, the midblastula transition (20), the cells are still dividing rapidly (approximately once per hour), and the stockpile of previously synthesized U1 RNAs has been depleted (10). Therefore, a high rate of U1 RNA transcription might be required at this time; that may be accomplished most efficiently by activation of the tandemly reiterated U1 RNA genes. Whether the mechanism of differential control of U1 RNA expression is similar to that regulating the 5S gene system (29, 31) remains to be determined.

**Transcription of embryonic genes in *X. laevis* oocytes.** Although transcription of the highly reiterated U1 RNA genes is not normally observed during late oogenesis, efficient transcription of these genes does occur upon injection of cloned copies of the repeat DNA into stage V and VI oocytes. The relative levels of xU1a and xU1b synthesis in oocytes injected with pXU1-H DNAs, however, are not identical to those in embryos. More specifically, xU1a RNA is the minor product in such injected oocytes (cf. Fig. 6B), whereas it is the major U1 RNA species in the embryo; the reverse is true for xU1b RNA (cf. Fig. 2 of reference 10a). One explanation for this quantitative discrepancy is that the injected clones of the 1.5-kb *Hind*III fragment lack some of the 5 flanking region sequences normally located far upstream of the xU1a genes (see the map of Fig. 5); the absence of such far-upstream sequences could affect the level of transcription of the xU1a RNA. A more interesting possibility, however, is that the difference in xU1a and xU1b RNA accumulation might accurately reflect the control of expression of the two embryonic genes in the oocyte. In that case, the low rate of transcription of the xU1a gene (which is very actively transcribed in the embryo) might demonstrate some of the cell specificity observed during development. Additional experiments utilizing our clone of the full-length U1 repeat DNA should allow us to distinguish between these possibilities.

Although it is unknown how the major class of U1 RNA genes is maintained in a quiescent state until the onset of embryonic transcription, it is possible that the clustering of the embryonic U1 RNA genes could facilitate the simultaneous activation of many repeating transcription units. It will be of interest to determine how differences in flanking region sequences affect the differential expression of the various U1 RNA genes.

#### ACKNOWLEDGMENTS

We thank J. Mitchen for technical assistance, D. Stalsberg for typing the manuscript, and E. Schenborn, J. Skuzeski, and J. Newport for critical comments. D. F. thanks M. Kirschner and T. Kornberg for generous support and encouragement.

This work was supported by Public Health Service grant GM30220 from the National Institutes of Health and by grant PCM 83-09618 from the National Science Foundation to J. E. D. and E. L. and by an American Cancer Society Senior Postdoctoral Fellowship-California Division (S-14-82) to D.F.

#### LITERATURE CITED

1. Barrell, B. G. 1971. Fractionation and sequence analysis of radioactive nucleotides, p. 751-779. In G. L. Cantoni and D. R.

- Davies (ed.), Procedures in nucleic acid research. Vol. 2. Harper & Row, Publishers, New York.
2. **Benton, W. D., and R. W. Davis.** 1977. Screening  $\lambda$ gt recombinant clones by in situ hybridization to single plaques in situ. *Science* **196**:180–182.
  3. **Branlant, C., A. Krol, J.-P. Ebel, E. Lazar, B. Gallinaro, M. Jacob, J. Sri-Widada, and P. Jeanteur.** 1980. Nucleotide sequences of nuclear U1A RNAs from chicken, rat and man. *Nucleic Acids Res.* **8**:4143–4154.
  4. **Brownlee, G. G., E. M. Cartwright, and D. D. Brown.** 1974. Sequence studies of the 5S DNA of *Xenopus laevis*. *J. Mol. Biol.* **89**:703–718.
  5. **Card, C. O., G. F. Morris, D. T. Brown, and W. F. Marzluff.** 1982. Sea urchin small nuclear RNA genes are organized in distinct tandemly repeating units. *Nucleic Acids Res.* **10**:7677–7688.
  6. **Dawid, I. B.** 1965. Deoxyribonucleic acid in amphibian eggs. *J. Mol. Biol.* **12**:581–599.
  7. **Denhardt, D. T.** 1966. A membrane-filter technique for the detection of complementary DNA. *Biochem. Biophys. Res. Commun.* **23**:641–646.
  8. **Denison, R. A., S. W. Van Arsdell, L. B. Bernstein, and A. M. Weiner.** 1981. Abundant pseudogenes for small nuclear RNAs are dispersed in the human genome. *Proc. Natl. Acad. Sci. U.S.A.* **78**:810–814.
  9. **Denison, R. A., and A. M. Weiner.** 1982. Human U1 RNA pseudogenes may be generated by both DNA- and RNA-mediated mechanisms. *Mol. Cell. Biol.* **2**:815–828.
  10. **Forbes, D. F., T. B. Kornberg, and M. W. Kirschner.** 1983. Small nuclear RNA transcription and ribonucleoprotein assembly in early *Xenopus* development. *J. Cell Biol.* **97**:62–72.
  - 10a. **Forbes, D. J., M. W. Kirschner, D. Caput, J. E. Dahlberg, and E. Lund.** 1984. Differential expression of multiple U1 small nuclear RNAs in oocytes and embryos of *Xenopus laevis*. *Cell* **38**:681–689.
  11. **Ford, J. P., and E. M. Southern.** 1973. Different sequences for 5S RNA in kidney cells and ovaries of *Xenopus laevis*. *Nat. New Biol.* **241**:7–12.
  12. **Grunstein, M., and D. S. Hogness.** 1975. Colony hybridization: a method for the isolation of cloned DNAs that contain a specific gene. *Proc. Natl. Acad. Sci. U.S.A.* **72**:3961–3965.
  13. **Lund, E., C. Bostock, M. Robertson, S. Christie, J. L. Mitchen, and J. E. Dahlberg.** 1983. U1 small nuclear RNA genes are located on human chromosome 1 and are expressed in mouse-human hybrid cells. *Mol. Cell. Biol.* **3**:2211–2220.
  14. **Lund, E., and J. E. Dahlberg.** 1984. True genes for human U1 small nuclear RNA: copy number, polymorphism and methylation. *J. Biol. Chem.* **259**:2013–2021.
  15. **Manser, T., and R. F. Gesteland.** 1982. Human U1 loci: genes for human U1 RNA have dramatically similar genomic environments. *Cell* **29**:257–264.
  16. **Marzluff, W. F., D. T. Brown, S. Lobo, and S. S. Wang.** 1983. Isolation and characterization of two linked mouse U1b small nuclear RNA genes. *Nucleic Acids Res.* **11**:6255–6270.
  17. **Mattaj, I. W., and R. Zeller.** 1983. *Xenopus laevis* U2 snRNA genes: tandemly repeated transcription units sharing 5' and 3' flanking homology with other RNA polymerase II transcribed genes. *EMBO J.* **2**:1883–1891.
  18. **Monstein, H.-F., K. Hammerström, G. Westin, J. Zabielski, L. Philipson, and U. Pettersson.** 1983. Loci for human U1 RNA—structural and evolutionary implications. *J. Mol. Biol.* **167**:245–257.
  19. **Murphy, J. T., R. R. Burgess, J. E. Dahlberg, and E. Lund.** 1982. Transcription of a gene for human U1 small nuclear RNA. *Cell* **29**:265–274.
  20. **Newport, J. W., and M. W. Kirschner.** 1982. A major developmental transition in early *Xenopus* embryos. I. Characterization and timing of cellular changes at the midblastula stage. *Cell* **30**:675–686.
  21. **Nojima, H., and R. Kornberg.** 1983. Genes and pseudogenes for mouse U1 and U2 small nuclear RNAs. *J. Biol. Chem.* **258**:8151–8155.
  22. **Peterson, R. C., J. L. Doering, and D. D. Brown.** 1980. Characterization of two *Xenopus* somatic 5S DNAs and one minor oocyte-specific 5S DNA. *Cell* **20**:131–141.
  23. **Rice, P. W., and J. E. Dahlberg.** 1982. A gene between *polA* and *gluA* retards growth of *Escherichia coli* when present in multiple copies: physiological effects of the gene for spot 42 RNA. *J. Bacteriol.* **152**:1196–1210.
  24. **Roop, D. R., P. Kristo, W. E. Stumph, M. J. Tsai, and B. W. O'Malley.** 1981. Structure and expression of a chicken gene coding for U1 RNA. *Cell* **23**:671–680.
  25. **Sanger, F., G. G. Brownlee, and B. G. Burrell.** 1965. Two-dimensional fractionation procedure for radioactive nucleotides. *J. Mol. Biol.* **13**:373–398.
  26. **Smith, G. E., and M. D. Summers.** 1980. The bidirectional transfer of DNA and RNA to nitrocellulose or diazobenzyloxymethyl-paper. *Anal. Biochem.* **109**:123–129.
  27. **Southern, E. M.** 1976. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**:503–517.
  28. **Wahli, W., and I. B. Dawid.** 1980. Isolation of two closely related vitellogenin genes, including their flanking regions, from a *Xenopus laevis* gene library. *Proc. Natl. Acad. Sci. U.S.A.* **77**:1437–1441.
  29. **Wakefield, L., and J. B. Gurdon.** 1983. Cytoplasmic regulation of 5S RNA genes in nuclear-transplant embryos. *EMBO J.* **2**:1613–1619.
  30. **Watanabe-Nagasu, N., Y. Itoh, T. Tani, K. Okano, N. Koga, N. Okada, and Y. Ohshima.** 1982. Structural analysis of gene loci for rat U1 small nuclear RNA. *Nucleic Acids Res.* **11**:1791–1801.
  31. **Wormington, W. M., M. Schlissel, and D. D. Brown.** 1983. Developmental regulation of *Xenopus* 5S RNA genes. *Cold Spring Harbor Symp. Quant. Biol.* **47**:879–884.
  32. **Zeller, R., M.-T. Carri, I. W. Mattaj, and E. M. DeRobertis.** 1984. *Xenopus laevis* U1 snRNA genes: characterization of transcriptionally active genes reveals major and minor repeated gene families. *EMBO J.* **3**:1075–1081.