# Stage-specific Keratins in *Xenopus laevis* Embryos and Tadpoles: The XK81 Gene Family

**Seiji Miyatani, Jeffrey A. Winkles, Thomas D. Sargent, and Igor B. Dawid**

Laboratory of Molecular Genetics, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland 20892. Dr. Winkles' present address is Division of Cell Biology, Biotechnology Research Center, Meloy Laboratories, Inc., Rockville, Maryland 20850.

*Abstract.* This report describes the isolation and characterization of genomic and cDNA clones which define a subfamily of type I keratins in *Xenopus laevis* whose expression is restricted to embryonic and larval stages. The XK81 subfamily, named after the prototype cDNA clone DG81, contains four members arranged in two pairs of closely homologous loci; they were named 81A1, A2, B1, and B2. Genomic clones were obtained representing all of these regions. The A1 gene has been completely sequenced together with ~1 kb of flanking sequences at each end; this gene corresponds to the previously reported cDNA clone 8128 (Jonas, E., T. D. Sargent, and I. B. Dawid, 1985, *Proc. Natl. Acad. Sci. USA*, 82:5413–5417). The B2 gene is represented by a partial cDNA clone, DG118. Upstream sequences and about half of the coding regions have been sequenced for the B1 and B2 genes, whereas the A2 locus has been identified on the basis of hybridization data and could be a gene or pseudogene. Genomic Southern blotting indicates that all members of the subfamily have been isolated. The keratin proteins encoded by the B1 and B2 genes are 96% homologous in the central rod domain, whereas A/B gene homology in this region is 81%. During development mRNAs derived from A and B genes accumulate coordinately during gastrula and neurula stages; in the tadpole, 81A mRNA decays rapidly, whereas 81B mRNA shows a second abundance peak, persists for most of tadpole life, and decays by metamorphosis. RNAs derived from the XK81 keratin subfamily are undetectable in the adult, where different type I keratin genes are expressed.

INTERMEDIATE filaments in epidermal and other epithelial cells are composed of keratins, members of a family of proteins that share a structural design as well as primary sequence homologies (reviewed in reference 17). Keratins can be classified into two distinct groups, type I keratins that are generally smaller and more acidic, and type II keratins that are generally larger and more basic. Proteins of each type are more homologous to each other than they are to those of the other type (reviewed in references 17 and 18).

In the frog *Xenopus laevis*, gastrulation is accompanied by vigorous expression of new genetic information (14). Prominent among the genes activated at this stage of development are keratin genes whose products accumulate rapidly in the ectoderm and in the differentiating epidermis of the developing embryo (10, 15, 22; Jamrich, M., T. D. Sargent, and I. B. Dawid, unpublished observations). The timing of expression of these keratin genes varies, but all share the property that expression is restricted to the embryonic and larval periods; no RNA transcribed from these embryonic genes can be detected in adult skin where distinct keratins are expressed (8, 9). We have undertaken a detailed analysis of the keratin genes expressed in the frog embryo with the aim to analyze the developmental program of expression in terms of regulatory controls and functional implications for epidermal development.

The embryonic keratin genes characterized so far in *Xenopus* include representatives of type I and type II polypeptides. Two predicted type I polypeptide sequences were derived from cDNAs named DG81 and DG70, which are quite different in sequence and in time of expression during development (10, 22). Further analysis of embryonic keratin genes has led to the realization that *Xenopus* contains several genes closely related to the original cDNA DG81. In the present paper we report the isolation and analysis of a genomic clone corresponding to the originally reported cDNA clone, and of clones representing additional related genes; the results define a subfamily of type I keratins in *Xenopus* which we call the XK81 family, with XK signifying the predicted proteins. Despite their close sequence similarity, member genes of this family are differentially expressed in frog development.

## Materials and Methods

### Genomic Clones

One million recombinants from a lambda library of genomic DNA sequences isolated from wild-caught animals (4) were screened by hybridization with nick-translated DG81 insert DNA (1), followed by washing at high

stringency (0.2× SSPE[1] at 65°C; 1× SSPE is 150 mM NaCl, 10 mM $Na_2HPO_4$, pH 7.0, 1 mM EDTA). Several different clones were isolated, including G8123, G8103, and G8132, which were analyzed in detail.

Restriction enzyme sites were mapped by standard procedures, and where possible, were confirmed by inspection of nucleotide sequence data. The XbaI site at 14 kb in G8103 overlaps with an MboI site, and will be methylated and resistant to digestion with XbaI in DNA prepared from *dam*$^+$ bacteria.

## Sequencing

***The 81A1 Locus.*** The 6.0-kb EcoRI fragment from G8132 containing the A1 locus was purified from an agarose gel and digested with exonuclease III and S1 to generate a series of partially deleted fragments. These were treated with DNA polymerase I to repair imperfectly blunt ends and cloned into pUC18 that had been digested with SmaI and treated with calf intestine alkaline phosphatase. A set of nested deletion derivatives was thereby generated. DNA from these subclones was prepared by a rapid alkaline lysis method (2) and further purified by chromatography on Sepharose CL2B. This DNA was digested with either BamHI or Asp718, dephosphorylated and labeled with T4 kinase and $\gamma$-$^{32}$P ATP, followed by digestion with either HindIII (for the BamHI) or EcoRI (for Asp718), which cleaved the polylinker sequence yielding a large labeled 81A1 fragment and a very small labeled polylinker fragment. The latter was removed by CL2B chromatography, and this was followed by chemical sequencing according to Maxam and Gilbert (11). Most of the 81A1 locus was sequenced on one strand only, except for 600 bp of 5′ flanking sequence, which was read on both strands.

***The 81B2 Locus.*** The 5′ half of the B2 locus, along with ~1,000 bp of 5′ flanking DNA, was sequenced by the chemical modification method (11) in conjunction with restriction endonuclease site mapping.

***The 81B1 Locus.*** A 2.7-kb EcoRI fragment from G8123 containing the 5′ half of the B1 locus, along with ~1,000 nucleotides of 5′ flanking DNA, was cloned into the EcoRI site of pUC18. This subclone was digested with BamHI and PstI and subjected to the directed exonuclease/S1 procedure of Henikoff (7). The deletion derivatives thereby obtained were sequenced by a primer extension method using double-stranded DNA template (21). Both the B1 and B2 loci were sequenced on one strand only.

***DG118.*** The cDNA clone, DG118, was sequenced by the chemical modification method of Maxam and Gilbert (11). Both strands were sequenced completely.•

## RNA Blot Analysis

Total RNA was isolated from whole embryos, subjected to electrophoresis on gels containing methylmercury hydroxide, transferred to nylon membrane, and hybridized as described in Sargent et al. (15). A final wash step was included using 0.1× SSPE at 70°C. This high stringency wash was shown to eliminate >90% of cross-reaction between 81A and 81B sequences in RNA/DNA hybridizations (data not shown).

## Primer Extension

A 17-residue long oligodeoxyribonucleotide was labeled with $\gamma$-$^{32}$P-ATP and T4 kinase. 50 ng primer was annealed with 0.5 μg poly(A)$^+$ RNA, prepared as described in Sargent et al. (15) from late neurula embryos, and extended by incubation with M-MLV reverse transcriptase (Bethesda Research Laboratories, Gaithersburg, MD) at 43°C for 20 min. The product was subjected to electrophoresis on a 6% acrylamide/urea gel, along with chain termination sequencing products using the identical primer and double-stranded DNA from G8132 as a template (13, 21).

# Results

## The XK81 Subfamily of Epidermal Keratins

To facilitate orientation we present the overall relationships among XK81 (*Xenopus* keratin 81) family genes at the outset before giving the evidence on which this summary is based. There are four loci in this subfamily, arranged in two pairs of closely related sequences. Following the nomenclature in the similar case of the four *Xenopus* vitellogenin genes (19),

1. *Abbreviation used in this paper*: 1× SSPE, 150 mM NaCl, 10 mM $Na_2HPO_4$, pH 7.0, 1 mM EDTA.

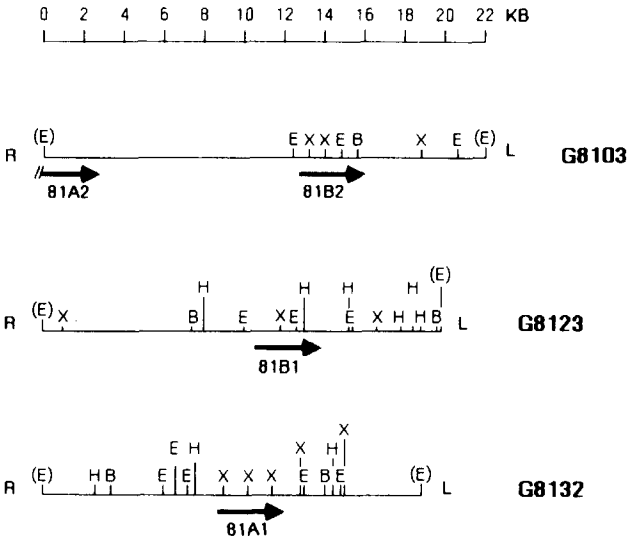


*Figure 1.* Maps of genomic clones containing XK81 family keratin genes. The position and direction of transcription of each gene is indicated by arrows. The 81A2 locus is incomplete. *R* and *L* signify the right and left arms of the lambda J1 cloning vector. Restriction enzymes are shown as follows: *B*, BamHI; *E*, EcoRI; *H*, HindIII; *X*, XbaI. EcoRI sites in parentheses correspond to the vector polylinker sequence and do not appear in the *Xenopus* genome.

we named the loci XK81A1, A2, B1, and B2. The A1 gene corresponds to the previously published cDNA clone 8128, which is a full-length homologue of the cDNA clone DG81 (10, 14), and the B2 gene is represented by the cDNA clone DG118 (see below). Therefore it can be concluded that these two genes are genuinely expressed, i.e., are not pseudogenes. Although there is nothing in the structure of the B1 locus to indicate otherwise, we have no direct evidence that it is transcribed. Less is known about the A2 locus; a portion of its 5′ region including the first exon has not been cloned, and no sequence data have been obtained. The close relationship between the A1 and A2 sequences is indicated by hybridization data.

## Genomic Clones from the XK81 Family

Phages homologous to cDNA 8128 were isolated from a *Xenopus* genomic DNA library, and three different phages containing homologous regions are shown in Fig. 1. The segments corresponding to keratin genes are indicated, together with the names assigned on the basis of subsequent analysis and the direction of transcription of each locus. As noted above, the A2 locus is incomplete: The region homologous to the cDNA probe (8128) was mapped to be very near the *Xenopus*-lambda boundary, and furthermore, no part of the A2 region could hybridize to a labeled probe corresponding to the first exon of the A1 gene, a probe that includes highly conserved sequences.

Aside from the absent 5′ end of A2, it appears that the three genomic clones shown in Fig. 1 include the entire XK81 subfamily. Evidence supporting this conclusion is shown in Fig. 2. Nuclear DNA samples from a single animal and single-copy equivalents of the three genomic clones were digested with EcoRI and analyzed by Southern blotting and hybridization with cDNA 8128, the A1 gene product. In the genomic DNA, four bands are visible, including a doublet of

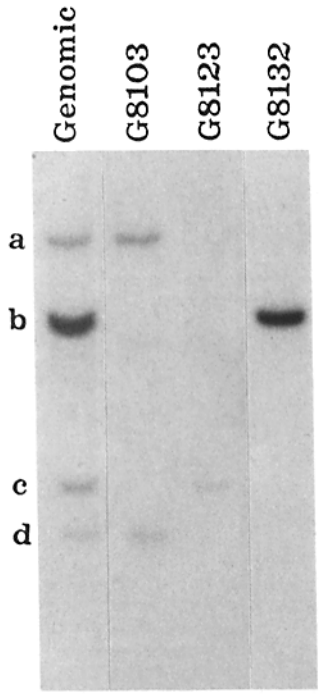


*Figure 2.* Genomic Southern blots. DNA from a single frog and genomic clones were digested with EcoRI, subjected to electrophoresis on an agarose gel, transferred to nitrocellulose membrane, and hybridized with labeled probe from cDNA clone 8128. The final wash was in 0.2× SSPE at 65°C. 5 μg of genomic DNA, or a single copy equivalent of phage clone DNA (83 pg) diluted with EcoRI-digested *Drosophila melanogaster* DNA, was loaded on each lane. Bands are labeled for reference in the text.

6.4 kb (band *b*). The intensity of the signals in genomic DNA samples is similar to that observed with single-copy equivalents of cloned DNA, indicating that the sequences corresponding to these bands are present only once per haploid genome. All of the bands visible in the genomic DNA samples are represented in the genomic clones. The doublet band is probably due to population polymorphism; we have isolated several genomic clones that are very similar to G8132 except that they contain a slightly larger EcoRI fragment homologous to 8128 (data not shown). Bands *a* and *d* are present in G8103, band *b* in G8132, and band *c* in G8123.

### *The XK81A1 Gene Sequence*

The keratin gene contained on phage G8132 was shown to correspond to the previously sequenced cDNA 8128; it was named XK81A1 and was completely sequenced together with ~1 kb of flanking sequence at each end (Fig. 3). The gene is composed of eight exons and seven introns whose placement will be discussed further below. The exon sequence shows only three differences compared with the sequence of cDNA 8128 (10; see legend to Fig. 3), all of which are silent. These minor differences are almost certainly due to polymorphisms in the frog population from which the various clones have been isolated; thus we conclude that cDNA 8128 is derived from the region cloned in G8132, and both of these clones represent the 81A1 gene.

The 5′ end of the mRNA was determined by primer extension (Fig. 4). The major extension product is actually one nucleotide shorter than the cDNA clone 8128. It appears likely that the extension product terminated one position downstream of the cap site; we have assigned position +1 to the 5′-terminal nucleotide of 8128.

### *The XK81B2 Gene*

The subtracted gastrula cDNA library prepared in our laboratory (14) was screened with a probe derived from late neurula stage epidermal poly(A)$^+$ RNA. Several cDNA clones were identified that hybridized strongly to this probe, and one, named DG118, proved to represent a keratin mRNA related to DG81. Sequence analysis showed that DG118 is precisely homologous to the keratin gene region in the center of the phage clone G8103 (Fig. 1); this region was named the 81B2 gene. The cDNA clone DG118 is ~80% of full length, missing sequences at both the 5′ and 3′ ends. The sequence information available from DG118 was augmented with sequences derived from the first exon of the B2 locus and the result is shown in Fig. 5. In this figure the nucleotide sequence of the 81B2 gene is translated into the XK81B2 protein and is compared with the predicted sequence of the XK81A1 protein (10, and Fig. 3). The B2 protein has a typical keratin structure. In the central rod domain (see reference 17), the homology between the A1 and B2 proteins is 81%, but the $NH_2$- and COOH-terminal domains have little significant similarity.

### *Sequence Relationships in the XK81 Keratin Family*

A keratin gene closely related to the B2 gene is contained on phage G8123 (Fig. 1); we have named it the XK81B1 gene. Part of its sequence, including 1 kb of 5′ flanking sequence and the first four exons and three introns, has been determined, and a portion of the upstream sequence is presented in Fig. 8. In its rod domain the predicted XK81B1 protein is 96% identical with the B2 protein, and even in the nonhelical $NH_2$-terminal domain the homology between the two proteins is 83%.

The intron sequences of B1 and B2, which are not shown, and all other available sequence data from the XK81 family, are being submitted to the Genbank Database. Analysis of the available intron sequences shows no significant homology between sequences in XK81A1 introns and those of the B1 or B2 loci. There is considerable homology between the introns in the closely related B1 and B2 genes; segments of high homology are interrupted by unrelated sequence stretches, suggesting evolutionary changes by frequent events of insertion or deletion. As in numerous other cases studied, the introns of different members of the XK81 keratin gene family show much lower levels of homology than the exons.

Relationships in the XK81 family were studied also by hybridization, especially to obtain information about the upstream homology region on phage G8103 (eventually named the A2 locus) for which no sequence data are available. Fig. 6 shows the results of an experiment in which digests of the three genomic clones (Fig. 1) were hybridized at moderately stringent criterion with probes derived from the rod-encoding regions of the B2 (G8103), B1 (G8123), and A1 (G8132) genes. The A1 probe hybridizes most intensely to the fragment of clone G8132 containing this gene (band *b*); the second most intense signal was seen with the upstream fragment from G8103 (band *a*). Weaker signals were obtained with fragments containing the B1 and B2 sequences. The corresponding intensity relationships were seen with the B gene probes: strongest hybridization to self, moderate hybridization to the other B gene fragments, and weak signal with A sequences. These results suggest that A1 and A2 have a similar level of homology as B1 and B2 for which sequence information is available (see above). On the basis of this information we classified the four XK81 loci into a subfamily with two pairs of closely related members, as introduced at the be-


10 20 30 40 50 60 70 80 90 100
E1 I1
I1. E2
E2 I2
I2 E3.
E3 I3
I3 E4
E4 I4
E5. I5
I5. E6
E6 I6
I4 E5

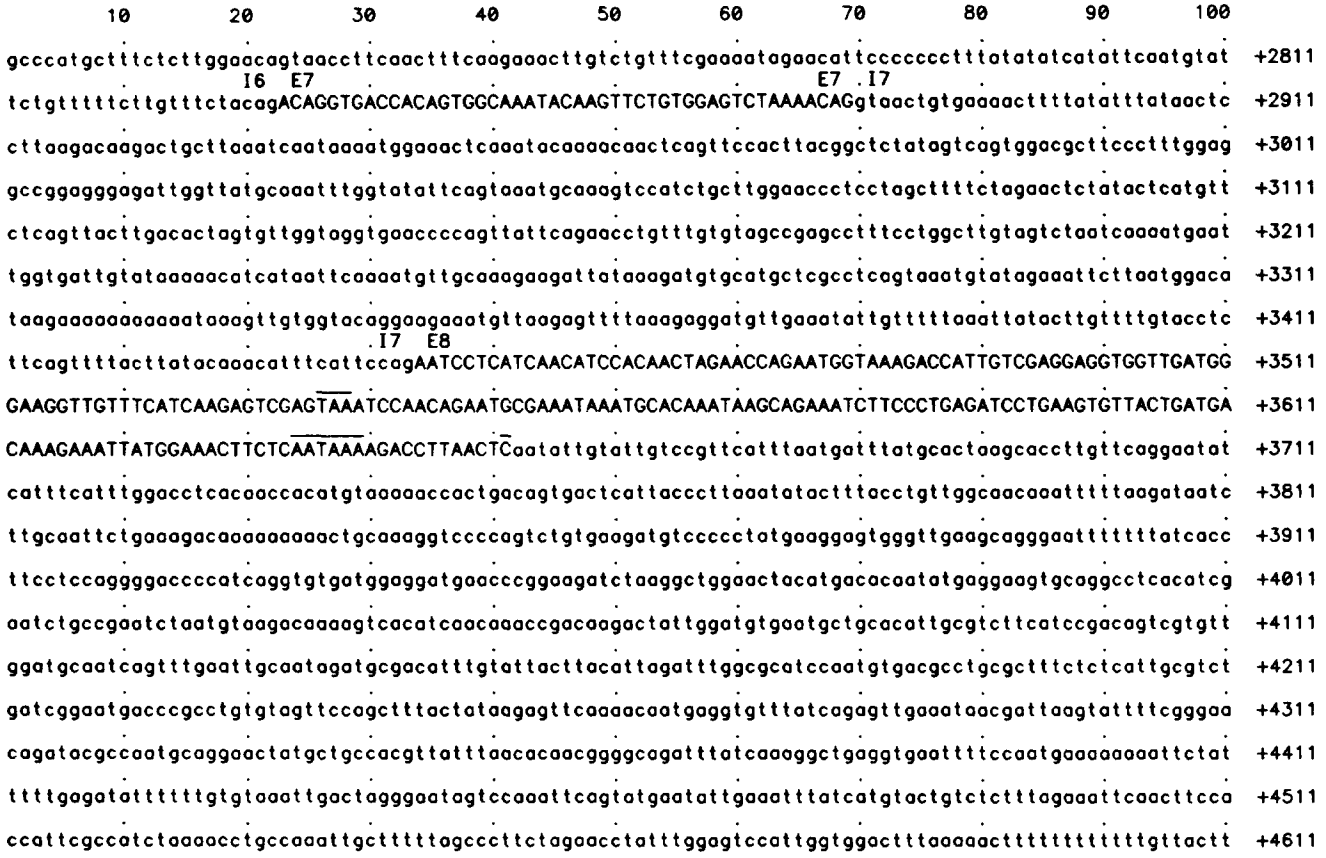


*Figure 3.* Sequence of the XK81A1 gene. Flanking sequences and introns are shown in lower case letters, exons in upper case. The exons (*E*) and introns (*I*) are also numbered above each boundary. The 5′-terminal nucleotide of the mRNA is assigned the number +1. Regions of interest are overlined: The TATA box at −23, the ATG translation initiation codon at +68, the termination codon TAA at +3,537, the polyadenylation signal at +3,635, and the poly(A) addition site at +3,652. The following changes are observed in the exon sequences compared with the sequence of cDNA 8128 (see reference 10): The T residues at positions +223, +466, and +1,538 are G, C, and C, respectively, in 8128. None of these changes alter the amino acid sequence of XK81A1.

ginning of the Results section. As already mentioned, the A1 and B2 regions are known to be expressed and thus represent genes, the B1 region has the structural properties of an active gene, while the status of the A2 region as a gene or pseudogene is undecided.

The four members of the XK81 family are not polymorphic variants of the same gene. Though the clones we analyzed were derived from libraries made from wild-caught animals that are known to harbor many polymorphisms, even the highly similar B1 and B2 genes have very different intron sequences and are contained within larger genomic segments of quite distinct restriction maps (Fig. 1).

### *Intron/Exon Maps of XK81 Family Keratin Genes*

The complete sequence of the 81A1 gene and the partial sequences of the B1 and B2 genes yield intron/exon maps that are summarized in Fig. 7 and Table I; the position of introns is also indicated on the exonic sequence of Fig. 5. The first four introns, for which information is available for all three genes, occur at homologous positions of the coding sequences. The placement of all the introns in 81A1 is similar to that in keratin genes of other animals, but some variation in placement occurs (18). In the XK81 family genes as in all type 1 keratin genes studied so far, the first exon is relatively large and includes the 5′ untranslated region, the entire sequence encoding the $NH_2$-terminal nonhelical domain of the protein, and a portion of the rod domain.

### *Sequence Homologies in the 5′ Flanking Regions*

Fig. 8 shows ~300 bp of the upstream flanking regions of the 81A1, B1, and B2 genes. These data document the great similarity of the B1 and B2 genes. Further upstream, the B1 and B2 loci are still homologous, but differences occur that suggest insertion or deletion events, similar to the situation in the introns of the same loci (see above).

Substantial homologies were also observed between the A and B genes, particularly in four regions. The first homology region is constituted by the TATA box, located 23 bp upstream of the 5′ end of the A1 mRNA. Whereas the A1 and B2 TATA boxes are quite standard, the B1 locus has an unusually long stretch of alternating A and T residues. Since this locus is not represented by a cDNA clone we cannot be certain that it is transcribed in vivo.

Three homology regions upstream of the TATA box, identified by overlining in Fig. 8, were observed. In the sequence up to 1 kb upstream no additional homology regions between A and B genes occur. The initiation site-proximal region with the sequence $CA_7G$ is precisely represented in all three genes. The other two regions include some mismatched bases but still constitute clearly identifiable segments of ho-

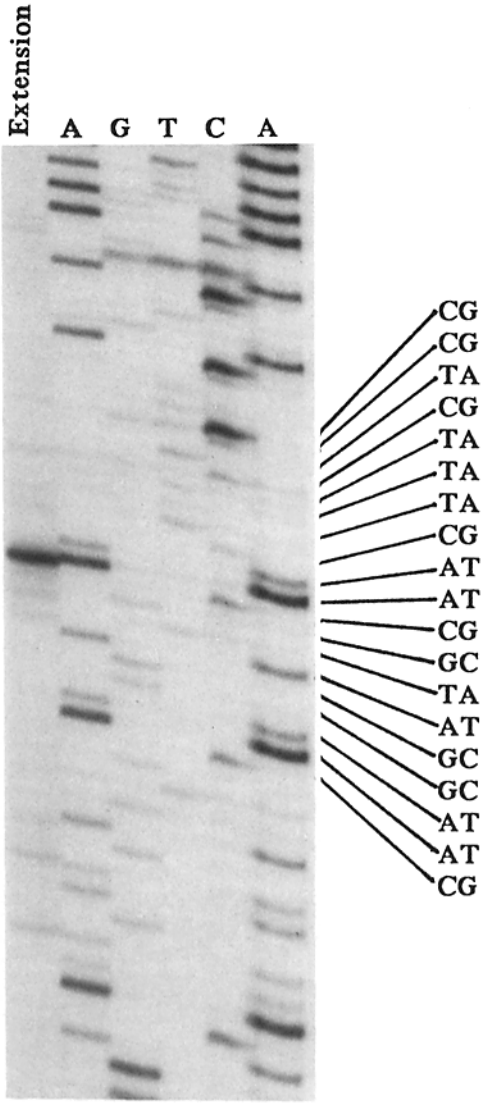


*Figure 4.* Determination of the cap site of XK81A1 mRNA by primer extension. A synthetic oligodeoxyribonucleotide (TGAAGTTTCCACATCAC) was annealed with mRNA, extended with reverse transcriptase, and separated on a sequencing gel next to chain termination reactions generated with the same primer and cloned genomic DNA (see Materials and Methods). As explained in the text the 5'-terminal nucleotide of the mRNA is assumed to correspond to the T residue immediately upstream from the point at which primer extension terminated.

mology. It may be speculated that these regions contain signals for the temporally and regionally controlled activation of the keratin genes during frog embryogenesis.

### *Differential Expression of A and B Genes in Development*

Earlier experiments showed that RNA homologous to the cDNA clone DG81 begins to accumulate at early gastrula and eventually decays in later tadpoles to become undetectable after metamorphosis (10, 16). Since the DG81 probe is now known to be homologous to different RNAs, we repeated the experiment under conditions that distinguish between A and B gene expression. Gel blot experiments showed that the XK81A and B mRNAs have the same size which is known to be 1,478 nucleotides plus the poly(A) tail (10). Fig. 9 shows the results of relative quantitations of XK81 mRNAs: Transcripts derived from the A and the B genes accumulate in parallel through gastrulation, neurulation, and early tadpole stages. After a peak in the 3-d tadpole, both types of RNA decay to about half their peak level during the next day. After this, the 81A mRNA continues to decay rapidly, while the 81B RNA increases again to reach a second peak in tadpoles during their later stage of development. As reported earlier all XK81-type mRNA reaches very low to undetectable levels in the postmetamorphic frog skin.

## *Discussion*

### *The Complexity of Keratin Genes in Xenopus laevis*

To date there is direct evidence, in the form of cDNA or genomic clones, for 11 keratin loci in *X. laevis*. Of these, three type II and one type I genes are expressed in adult skin (8, 9), five are embryonic/larval specific, and two, which are of the embryonic/larval class, are not definitely known to be expressed. These genes can be classified into subfamilies: The three adult type II genes are very closely related, whereas the embryonic sequences include one type II gene and six type I genes that define two subfamilies. These two type I subfamilies were named after the prototype cDNA clone identifying them; the XK70 subfamily (22) is represented by two known genes, whereas four members make up the XK81 group. Homology between the predicted XK81A1 or B2 and XK70 proteins is not particularly high, only 53 and 57% in the rod domain, respectively. Divergence between these two subfamilies is thus as high or somewhat higher than between either and a human type I keratin (see 22). This suggests that divergence between XK70 and XK81 genes began early in the evolution of vertebrates.

The total complexity of the keratin family in any one organism is not known. While the clones we have obtained represent the more abundant keratin mRNAs expressed in early frog development, data from protein gel electrophoresis suggest that additional keratins occur during oogenesis and in later tadpole stages (5, 6; Winkles, J. A., T. D. Sargent, M. Jamrich, and I. B. Dawid, unpublished results).

### *Evolution of the XK81 Family*

Karyotypic and molecular evidence suggest that a genome duplication has taken place during the evolution of the *X. laevis* species (3, 19). The properties of the XK81 keratin family, which are reminiscent of those of the vitellogenin family in the same animal (19, 20), are consistent with such an interpretation. We suggest that an XK81A and B gene were generated by duplication some time before the genome duplication; this duplication was probably a local one since the A2/B2 pair is still closely linked. After some divergence between these genes had taken place the entire *X. laevis* genome duplicated, generating the four-locus XK81 family. Since this time relatively little divergence has occurred, leading to the closely related A1/A2 and B1/B2 pairs.

### *Differential Expression of 81A and 81B mRNAs*

The A and B RNAs accumulate coordinately but decline differentially; we do not know whether this is due entirely

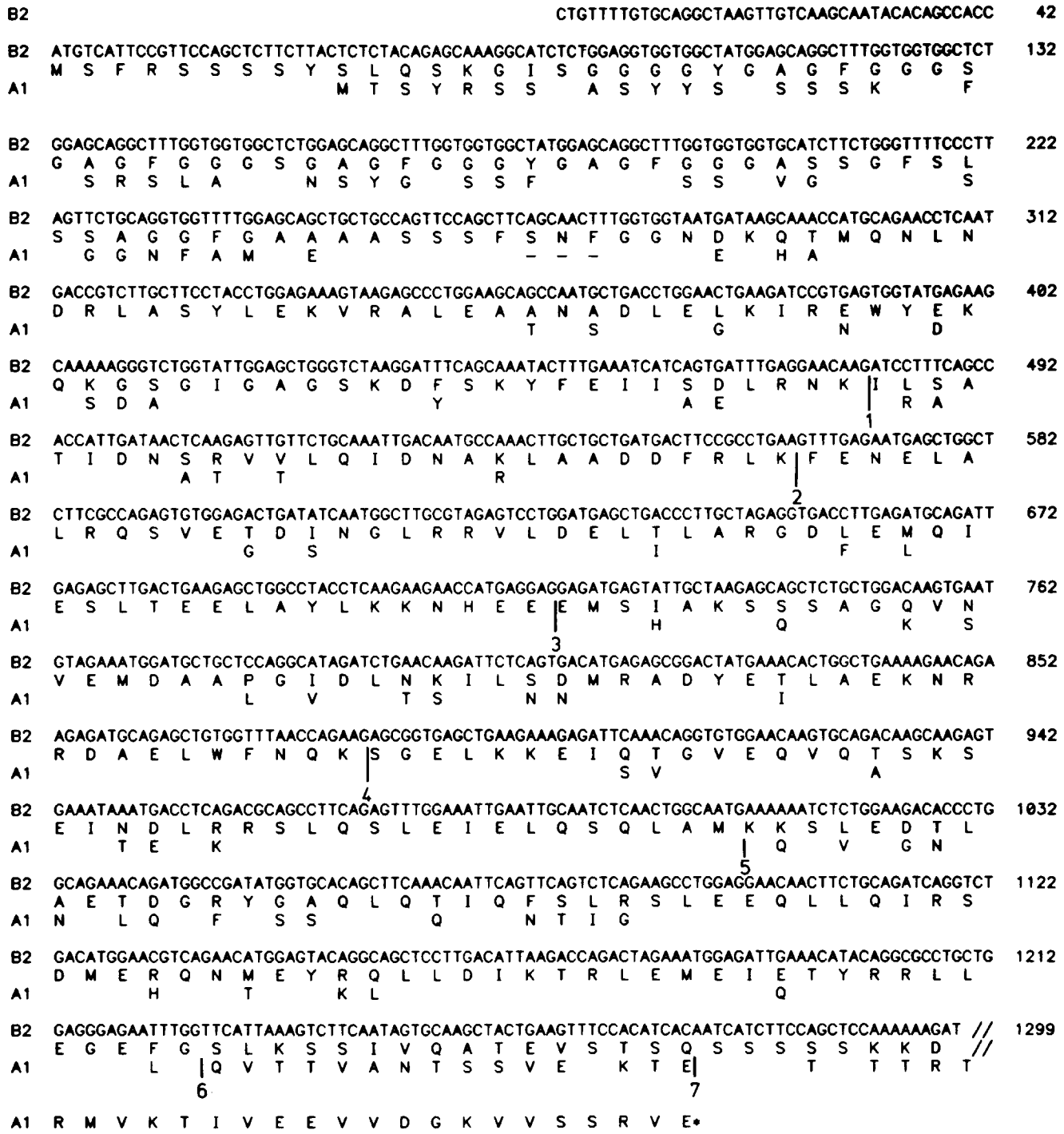

*Figure 5.* Sequence of the XK81B2 mRNA and comparison of the predicted XK81A1 and XK81B2 proteins. The nucleotide sequence of cDNA clone DG118 is represented by nucleotides 107–1,299. Exon sequence of the B2 gene corresponding to nucleotides 1–106 was used to provide the 5′ region of the mRNA. The position of nucleotide +1 was inferred by homology to the A1 gene (Fig. 8), and the polypeptide sequence was assumed to begin with the first available ATG codon. In the sequence overlap of 776 nucleotides available between cDNA DG118 and genomic B2 exons on phage G8103 there were no differences. The predicted XK81B2 protein is compared with the A1 protein taken from Jonas et al. (10). Blank spaces signify identity, and dashes are gaps introduced for alignment. The positions at which introns interrupt the sequence are shown by vertical lines; these positions are identical in A1 and B2 for the first four introns. B2 genomic sequence is not available beyond this point.

to transcriptional regulation or also to differences in RNA stability. If the A and B genes originated by duplication, the sequences involved in differential regulation must have evolved since that time. Sequence comparisons between the A1 and B2 genes revealed some similarities as well as differences in the upstream flanking and untranslated regions; which of these regions may be involved in the common activation of these genes in early gastrula or their different behavior later is presently unknown.

In our experiments we did not distinguish between A1 and

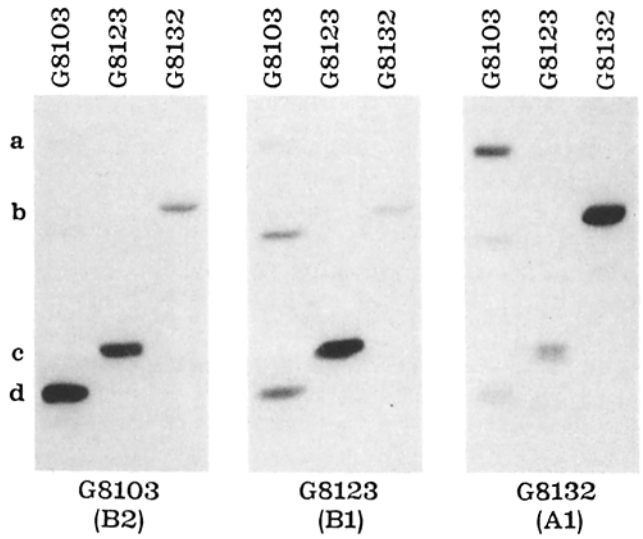


*Figure 6.* Cross-hybridization of XK81 family clones. Phage DNA (Fig. 1) was digested with EcoRI, separated on an agarose gel, and transferred to nylon membrane filters. Replicate filters were hybridized and washed as described (15). The final wash buffer has a $Na^+$ concentration of 200 mM and was used at 65°C. Probes were derived from the indicated genes as follows: The A1 probe was a 2.6-kb PstI fragment from G8132, from 8.4 to 11 kb in Fig. 1 (PstI sites not shown in this figure). The B1 probe was the 2.7-kb EcoRI fragment from G8123, and the B2 probe was the 2.1-kb EcoRI fragment from G8103. The major bands are indicated as in Fig. 2 for reference in the text.

*Table I. Exon and Intron Sizes in XK81 Genes*

| | *A1* | *B1* | *B2* |
|---|---|---|---|
| E1 | 469 | 490 | 482 |
| I1 | 452 | 601 | 679 |
| E2 | 83 | 83 | 83 |
| I2 | 283 | 129 | 112 |
| E3 | 157 | 157 | 157 |
| I3 | 130 | 129 | 114 |
| E4 | 162 | 162 | 162 |
| I4 | 379 | – | – |
| E5 | 126 | – | – |
| I5 | 100 | – | – |
| E6 | 218 | – | – |
| I6 | 274 | – | – |
| E7 | 47 | – | – |
| I7 | 564 | – | – |
| E8 | 208 | – | – |

The size, in base pairs, of exons (E1, E2, etc.) and introns (I1, I2, etc.) was deduced from available sequence data from the A1, B1, and B2 genes. The 5′ boundary of exon 1 (E1) for the A1 gene corresponds to the first nucleotide of cDNA clone 8128 (Fig. 4; reference 10). The equivalent positions of the B1 and B2 genes were predicted by homology to the A1 gene (Fig. 8). Ambiguities in the location of splice sites were resolved according to the assumption that introns begin with the residues GT and end with AG.

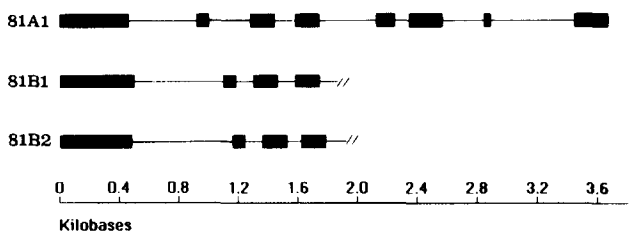


*Figure 7.* Intron/exon map of XK81 family genes. The maps summarize the available sequence information. The positions of introns and their lengths are also given in Table I and indicated in Fig. 5.

A2 or between B1 and B2 RNAs. Given the great similarity between transcribed and flanking sequences of the B1 and B2 genes we might expect them to be co-regulated. If this were found to be the case it would strengthen the interpretation of the origin of the ½ pairs by genome duplication: Their existence would have no functional basis but would be a consequence of a global evolutionary event in the history of the species.

While differential regulation alone does not provide conclusive evidence that the 81A and 81B gene products have different functions, such an interpretation is consistent with the expression data. XK81A mRNA is present only during early development when the epidermis is first established in the embryo; development beyond 4 d after fertilization does not require this mRNA. The situation is even more pronounced with XK70 mRNA which encodes a type I keratin of rather distinct sequence: This mRNA accumulates in the gastrula and neurula stages but decays rapidly thereafter and

```
A1  TTGCTGAAGGAAGCCTGGAGCAAGGAGAGAGTGGTGCTGCAACATCAAGTGAAAATCTTCCAAGAATTTGCCCCAACCAGTTTGTAACCAAGTTTTTGTTT  -170
B1  TTCAGCAGGGTTGCCTGGGGCAGGTTAGAGAGAGAAGGGTGAAAACTGTAAAGAAAAG----AAGAATTTGCCTCTCCCAGTTTCTAACTAATACCTTTAGG
B2  TTCAGCAAAGTTGCCTGGGGCAAGGT-GAGAGGGACGGCTGAAAACTGGAAAGAAAAG----AAGAATTTGCCTCTCCCAGTTTCTAACCAATACCTTCAGG

A1  AACAAACACCCTGAGGCTACGTAACTGAATCAAGGCGATTATCCATCTCTCGATCAGTGCAGTGACCTGGCCAAAAAAAGGAATTTTCCATATAGGAGGAC  -69
B1  GCTATTTCATTTATTTAAATATGAACTGATTATGATGGTGATTTTCTGGAAACAGATGCTGACCTGT----CAAAAAAAGTATTTACTCTGGGAGGCTGTC
B2  GCTATATCTTTTATGTAAATATGGACTGATTATGATGGTGATTTTCTGGAAACAGATGCTGGCCTGG----CAAAAAAAGTATTTACTCTGGGAGGCTGTC

                                                                                    1-->
A1  GTCCGGTTT-GGGAAA-A-GGACCACCCACCCCAGCATGTGTATATATGTGCATGGA-CTGGAAGGAGAAAGTTGCATCCTTGCACTCTTCTTTGTCCTAG  +29
B1  CCCAGTGTAACGGAAACACCCACCACTCTTTCTACATACATTATATATATATATATATATAAAGCCTGTCAAAGCTGCATGCATCCATTCTGTTTTGTGCAGG
B2  CCTGGTGTAACTGAAACACCCACCACTCTTTGTGTATCCACTATATATATA----------GGCTGTCAAAGCTGCAGGCACCCATTCTGTTTTGTGCAGG
```

*Figure 8.* Upstream flanking sequences in XK81 family genes. Gaps have been introduced for optimal alignment. Sequences of interest are overlined and described in the text.

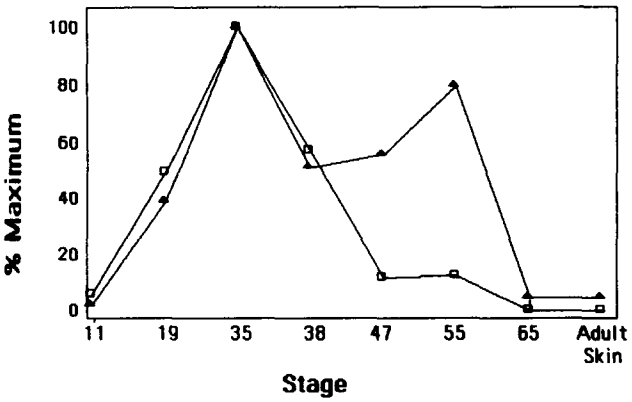


*Figure 9.* Differential expression of XK81A and 81B RNA. Relative mRNA abundance was measured by RNA gel blotting of equal amounts of RNA from different stages as described in Materials and Methods. The hybridization and wash conditions used permit ≤10% cross-hybridization between A and B sequences. Hybridization signals were measured by densitometry of multiple exposures in the linear range of the film, and normalized to 100% maximal values for each probe. ▲, DG118 probe, XK81B RNA; □, 8128 probe, XK81A RNA. Stages are according to Nieuwkoop and Faber (12).

has effectively disappeared at the end of day 2 (22). The morphology of tadpole epidermis does not change dramatically until metamorphosis (12), yet the keratin composition is modified as indicated by the changing mRNA population and by electrophoretic analysis of proteins (5). These data suggest that functionally distinct types of intermediate filaments are elaborated during tadpole epidermal development, but different explanations based on co-regulation of gene sets can be envisioned as discussed earlier (10).

Received for publication 21 April 1986, and in revised form 12 July 1986.

### *References*

1. Benton, W. D., and R. W. Davis. 1977. Screening gt recombinant clones by hybridization to single plaques in situ. *Science (Wash. DC).* 196:180–182.

2. Birnboim, H. C., and J. Doly. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* 7:1513–1523.

3. Bisbee, C. A., M. A. Baker, A. C. Wilson, I. Hadji-Adzimi, and M. Fischberg. 1977. Albumin phylogeny for clawed frogs (Xenopus). *Science (Wash. DC).* 195:785–787.

4. Chien, Y.-H., and I. B. Dawid. 1984. Isolation and characterization of calmodulin genes from *Xenopus laevis. Mol. Cell. Biol.* 4:507–513.

5. Ellison, T. R., P. M. Mathisen, and L. Miller. 1985. Developmental changes in keratin patterns during epidermal maturation. *Dev. Biol.* 112:329–337.

6. Franz, J. K., L. Gall, M. A. Williams, B. Picheral, and W. W. Franke. 1983. Intermediate-size filaments in a germ cell: expression of cytokeratins in oocytes and eggs of the frog *Xenopus. Proc. Natl. Acad. Sci. USA.* 80:6254–6258.

7. Henikoff, S. 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. *Gene.* 28:351–359.

8. Hoffmann, W., and J. K. Franz. 1984. Amino acid sequence of the carboxy-terminal part of an acidic type I cytokeratin from *Xenopus laevis* epidermis as predicted from the cDNA sequence. *EMBO (Eur. Mol. Biol. Organ.) J.* 3:1301–1306.

9. Hoffmann, W., J. K. Franz, and W. W. Franke. 1985. Amino acid microheterogeneity of basic (type II) cytokeratins of *Xenopus laevis* epidermis and evolutionary conservativity of helical and non-helical domains. *J. Mol. Biol.* 184:713–724.

10. Jonas, E., T. D. Sargent, and I. B. Dawid. 1985. Epidermal keratin gene expressed in embryos of *Xenopus laevis. Proc. Natl. Acad. Sci. USA.* 82:5413–5417.

11. Maxam, A. M., and W. Gilbert. 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* 65:499–560.

12. Nieuwkoop, P. D., and J. Faber. 1967. Normal Table of Xenopus laevis (Daudin). 2nd ed. Elsevier North-Holland Biomedical Press, Amsterdam.

13. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* 74:5463–5467.

14. Sargent, T. D., and I. B. Dawid. 1983. Differential gene expression in the gastrula of *Xenopus laevis. Science (Wash. DC).* 222:135–139.

15. Sargent, T. D., M. Jamrich, and I. B. Dawid. 1986. Cell interactions and the control of gene activity during early development in *Xenopus laevis. Dev. Biol.* 114:238–246.

16. Sargent, T. D., E. Jonas, M. Jamrich, G. S. Michaels, S. Miyatani, J. A. Winkles, and I. B. Dawid. 1986. Cytokeratins and their genes in Xenopus laevis: structural and developmental aspects. *In* Intermediate Filaments. R. D. Goldman and P. M. Steinert, editors. Plenum Publishing Corp., New York. In press.

17. Steinert, P. M., and D. A. D. Parry. 1985. Intermediate filaments: conformity and diversity of expression and structure. *Annu. Rev. Cell. Biol.* 1:41–65.

18. Steinert, P. M., A. C. Steven, and D. R. Roop. 1985. The molecular biology of intermediate filaments. *Cell.* 42:411–419.

19. Wahli, W., I. B. Dawid, G. U. Ryffel, and R. Weber. 1981. Vitellogenesis and the vitellogenin gene family. *Science (Wash. DC).* 212:298–304.

20. Wahli, W., J. E. Germond, B. ten Heggeler, and F. E. B. May. 1982. Vitellogenin genes A1 and B2 are linked in the *Xenopus* genome. *Proc. Natl. Acad. Sci. USA.* 79:6832–6836.

21. Wallace, R. B., M. J. Johnson, S. V. Suggs, K. Miyoshi, R. Bhatt, and K. Itakura. 1981. A set of synthetic oligodeoxyribonucleotide primers for DNA sequencing in the plasmid vector pBR322. *Gene.* 16:21–26.

22. Winkles, J. A., T. D. Sargent, D. A. D. Parry, E. Jonas, and I. B. Dawid. 1985. Developmentally regulated cytokeratin gene in *Xenopus laevis. Mol. Cell. Biol.* 5:2575–2581.