J Mol Evol (1988) 27:17–28

Journal of Molecular Evolution

Cytoskeletal Actin Gene Families of *Xenopus borealis* and *Xenopus laevis*

Gareth S. Cross,[1] Clive Wilson,[1] Harry P. Erba,[2] and Hugh R. Woodland[1]

[1] MRC Animal Development Group, Department of Biological Sciences, University of Warwick, Coventry CV4 7AL, England
[2] The MEDIGEN Project, Department of Medicine, Stanford University School of Medicine and Veterans Administration Medical Center, Palo Alto, California 94304, USA

Summary. We have sequenced the coding and leader regions, as well as part of the 3′ untranslated region, of a *Xenopus borealis* type 1 cytoskeletal actin gene [defined according to the arrangement of acidic residues at the N-terminus; Vandekerckhove et al. (1981) J Mol Biol 152:413–426]. The encoded amino acid sequence is the same as the avian and mammalian β (type 1) cytoskeletal actins, except for an isoleucine at position 10 (as found in the mammalian γ cytoskeletal actins), and an extra amino acid, alanine, after the N-terminal methionine. Five introns were found, in the same positions as those of the rat and chicken β-actin genes. The 5′ and 3′ untranslated regions resemble those of the human γ (type 8) cytoskeletal actin gene more closely than the mammalian β genes.

Primer extension showed that this type 1 gene is transcribed in ovary and tadpole. Sequencing of primer extension products demonstrated two additional mRNA species in *X. borealis*, encoding type 7 and 8 isoforms. This contrasts with the closely related species *Xenopus laevis*, where type 4, 5, and 8 isoforms have been found. The type 7 isoform has not previously been found in any other species. The mRNAs of the *X. borealis* type 1 and 8 and *X. laevis* type 5 and 8 isoforms contain highly homologous leaders. The *X. borealis* type 7 mRNA has no leader homology with the other mRNA species and, unlike them, has no extra N-terminal alanine codon. The evolutionary implications of these data are discussed.

Key words: Actin genes — Evolution — 5′ and 3′ untranslated regions

Offprint requests to: G.S. Cross

Introduction

Actins are a family of highly conserved contractile proteins. Some isoforms are muscle specific, while the cytoskeletal actins are found in probably all eukaryotic cells. Of the four mammalian muscle-specific isoforms (Vandekerckhove and Weber 1978a), two are specific to striated muscles (the heart and skeletal muscle isoforms, differing in their relative abundance in the heart and in skeletal muscle), and two are specific to smooth muscle (the aortic and stomach isoforms, differing in their relative abundance in the aorta and the stomach). In mammals there are two different major cytoskeletal actins, the β and γ isoforms (Vandekerckhove and Weber 1978b), although there have been unconfirmed reports of additional isoforms in the nuclei of non-muscle cells (Bremer et al. 1981), and in brain synaptosomes (Marotta et al. 1978). The β and γ isoforms are also found in chickens (Vandekerckhove and Weber 1981), but, in addition, there has recently been a report that chickens have a gene encoding a type 5 (see Table 1) actin (Bergsma et al. 1985), although such a protein has so far not been detected in vivo. The cytoskeletal actins differ from the muscle-specific isoforms in 22–25 amino positions (Vandekerckhove and Weber 1979). For a given isoform, however, such as the β cytoskeletal actin, the amino acid sequence has been found to be identical in humans, rats, and chickens (Kost et al. 1983; Nudel et al. 1983; Ponte et al. 1984). The β and γ isoforms of mammals and chickens differ from each other in only four amino acids: the three acidic residues at the N-terminus and the amino acid at position 10 (Vandekerckhove and Weber 1978b).

Table 1. Distribution of the different vertebrate cytoskeletal actin isoforms

Actin type[a]	N-terminal amino acids	Mammals	*R. pipiens*	*T. cristatus*	*X. laevis*	*X. borealis*
1	asp-asp-asp	+ (= β)		+		+
3	asp-glu-asp			+		
4	glu-asp-asp		+		+	
5	asp-glu-glu		+		+	
7	glu-glu-asp					+
8	glu-glu-glu	+ (= γ)		+	+	+

[a] Nomenclature as suggested by Vandekerckhove et al. (1981)

The conservation of these different cytoskeletal isoforms in mammals and birds has suggested separate functions for each of them, although there is little evidence for this so far. It is therefore interesting that this conservation does not extend to amphibians, where Vandekerckhove et al. (1981) have partially sequenced the N-terminal tryptic peptides of cytoskeletal actins from several different species (*Xenopus laevis, Rana pipiens,* and *Triturus cristatus*) and found new isoforms that differ in their content and order of glutamate and aspartate residues in the first three amino acid positions. As Table 1 shows, different amphibian species have different sets of cytoskeletal actins. These differences may have occurred because of selective advantages of certain isoforms to species in particular environments, or they may be merely neutral mutations (i.e., there may be no functional difference between the isoforms). In this paper we show, by the sequencing of primer extension products, that *Xenopus borealis,* a species thought to have diverged from *X. laevis* only 8 million years (Myr) ago (Bisbee et al. 1977), contains mRNA species encoding two cytoskeletal isoforms, types 1 and 7 (see Table 1), which have not been found in *X. laevis,* and contains mRNA for a type 8 isoform as well, which is present in both species.

We also present the structure of a *X. borealis* type 1 cytoskeletal actin gene, demonstrating the conservation of its intron positions and amino acid sequence when compared with mammalian and avian β (type 1; see Table 1) cytoskeletal genes. This gene, however, shows greatest similarities in its leader and 3′ untranslated region sequences, as well as at amino acid position 10, to a human γ (type 8) cytoskeletal gene. We speculate on the evolutionary relationships of the *Xenopus* genes we have detected to the cytoskeletal actin genes of other organisms.

Materials and Methods

Materials. Clone λ35A was isolated from a partial Eco RI genomic library cloned in λgtWES (Bains 1982). Clone λ3H1 was isolated from a partial Sau 3A1 library in λ47.1 (a generous gift from Dr. P.C. Turner). Kenyan *X. borealis* frogs, used to make DNA for these libraries and RNA for primer extension analysis, were bred in this laboratory. *Xenopus laevis,* used to make RNA, were obtained from the South African Snake Farm, Fishhoek, South Africa. RNAs were kindly supplied by Ms. J.E.M. Ballantine and Dr. P.C. Turner, and were made as described previously (Woodland et al. 1984). [α-^{32}P]deoxynucleoside triphosphates (>2000 Ci/mmol) and [γ-^{32}P]ATP (>5000 Ci/mmol) were from Amersham International. Restriction enzymes were from Boehringer Mannheim, Bethesda Research Laboratories, and New England Biolabs. *Escherichia coli* DNA polymerase I (Klenow fragment and the Kornberg enzyme) and calf intestinal alkaline phosphatase were from Boehringer Mannheim. T4 DNA polymerase I, DNA ligase, and polynucleotide kinase were from Bethesda Research Laboratories. AMV reverse transcriptase was obtained from Life Sciences Inc.

Isolation of Actin Clones. Initially a *X. borealis* partial Eco RI library was screened by the procedure of Benton and Davis (1977). The probe was a nick-translated *Dictyostelium* actin cDNA clone (pcDd actin B1; Bender et al. 1978). This was hybridized to nitrocellulose in 4× standard saline citrate (SSC) (1× SSC = 0.15 M NaCl, 0.015 M sodium citrate), 1× Denhart's additives (0.2% Ficoll, 0.2% polyvinylpyrrolidone, 0.2% bovine serum albumin), 60 mg/l sonicated *E. coli* DNA, 0.1 mg/ml poly(A) at 60°C for 48 h, and washed at 55°C in 4× SSC for 1 h. Plaques that showed a positive signal were picked and screened a second time. These included the recombinant λ35A. Subsequently, a second library (a partial Sau 3A1 digest cloned in the vector λ47.1) was screened, resulting in the isolation of clone λ3H1, containing the 3.6-kb Eco RI fragment of λ35A in addition to the surrounding *X. borealis* sequence. For this screen, the hybridization was carried out in 3× SSC, 0.1× Denhart's additives, 0.1 mg/ml poly(A) at 65°C for 16 h, and the filters were washed in 3× SSC at 50°C. The probe used was the nick-translated insert of λ5AP, a clone obtained on screening the first library, which contains the sequence encoding the first 150 amino acids of a muscle-specific actin. The clone λ3H1 was purified by further plating and rescreening.

Mapping the Clones. We used double digests of nick-translated DNA and partial digests of DNA fragments labeled at one end with T4 polynucleotide kinase or T4 polymerase I (Smith and Birnstiel 1976). The resulting fragments were separated on agarose gels. To locate the coding regions, unlabeled restriction fragments were separated on agarose gels, blotted onto nitrocellulose (Southern 1975), and hybridized with the nick-translated *Dictyostelium* actin probe under the same conditions as used to isolate λ35A.

Sequencing of DNA. For this, as indicated in Fig. 1, we used both the chemical base modification procedure of Maxam and Gilbert (1978) and the enzymic method of Sanger et al. (1980)

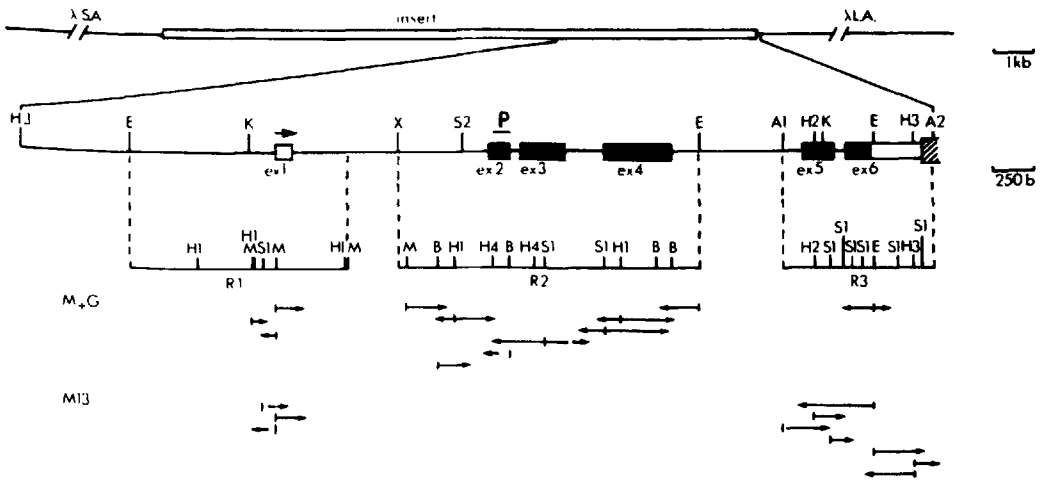


Fig. 1. Structural map and sequencing strategy for the *X. borealis* genomic clone λ3H1. λL.A and λS.A mark the long and short arms, respectively, of the vector. Three regions, R1, R2, and R3, have been mapped in greatest detail. Symbols for restriction enzymes: A1, Acc I; A2, Ava I; H1, Hinf I; H2, Hinc II; H3, Hind III; H4, Hha I; B, Bst NI; M, Msp I; E, Eco RI; K, Kpn I; X, Xba I; S1, Sau 3A1; S2, Sst I. Sequencing was in the region and direction represented by the arrows and was by the method of Maxam and Gilbert ("M+G," Maxam and Gilbert 1978) or of Sanger ("M13," Sanger et al. 1980). The solid blocks represent coding exons; the open bars, untranslated sections of the mRNA. The hatched region represents a portion of the λ cloning vector. The horizontal arrow over exon 1 indicates the start and direction of transcription. The region P is the Hha I/Bst NI fragment used for primer extension.

applied to single-stranded M13 subclones. For the former procedure, DNA restriction fragments were end-labeled, either with T4 polynucleotide kinase (after treatment with alkaline phosphatase) or with T4 DNA polymerase I, prior to asymmetrical cleavage with a second restriction enzyme. The products were separated on 5% nondenaturing acrylamide gels and extracted from the gel as described by Maxam and Gilbert (1978).

Primer Extension. The primer was prepared from a 450-bp Bst NI fragment, one end of which was situated in the first coding exon (= exon 2 in Fig. 1), and the other was situated 5′ to this exon. This Bst NI fragment was isolated from the 3.6-kb Eco RI fragment of λ35A (recloned into the plasmid pBR325 for easy preparation), treated with alkaline phosphatase, 5′ end-labeled with T4 polynucleotide kinase and [γ-^{32}P]ATP, cut with Hha I, and fractionated on an 8% acrylamide–7 M urea sequencing-type gel. The 103-nucleotide (nt) fragment of the noncoding strand, corresponding to residues 7–41 of the coding region, was extracted from the gel.

Hybridization of this primer to RNA and its subsequent extension were as described previously (Woodland et al. 1984), except that 8% sequencing-type gels were used to separate the primer extension products. In preliminary experiments, to optimize hybridization temperatures for example, 0.01–0.02 mg of total RNA was used in 10-μl hybridizations. To obtain enough extended material for sequencing, however, larger amounts of RNA were used. Thus, to sequence *X. borealis* primer extension products A and B (see Fig. 3), 0.75 mg of ovary RNA was hybridized to primer overnight in 0.35 ml of hybridization solution (0.4 M NaCl, 10 mM PIPES, pH 6.4) containing 50% formamide at 45°C (equivalent to a hybridization at 70°C without formamide). To sequence *X. borealis* product C, 0.2 mg of tadpole RNA was hybridized to primer for 3 h at 65°C in 0.13 ml of hybridization solution without formamide, but containing 1% sodium dodecyl sulfate. To sequence *X. laevis* products A and B, 0.4 mg of XTC cell (Pudney et al. 1973) RNA was hybridized in 0.2 ml of 1% SDS-containing hybridization buffer at 70°C for 3 h.

Results and Discussion

The Structure of a X. borealis Cytoskeletal Actin Gene

We have isolated a clone, λ35A, from a λgtWES *X. borealis* genomic library. The 3.6-kb Eco RI insert was found to contain the sequence encoding the first three exons of an actin gene. Subsequently, another clone, λ3H1, was isolated from a λ47.1 library (made from the DNA of a different individual). This clone contained an extra 1.4 kb of sequence in the 3′ portion of the gene and an extra 10 kb of sequence upstream from the 3.6-kb Eco RI fragment. This enabled us to determine the rest of the coding sequence, as well as 318 bases of the 3′ flanking sequence. The structure of this gene and the sequencing strategy are shown in Fig. 1.

Figure 2 shows the sequence of regions 2 and 3 from Fig. 1. As customary, we have numbered the amino acids according to the nomenclature suggested by Elzinga and Lu (1976) and Vandekerckhove and Weber (1978b) for rabbit skeletal muscle actin. Except for two positions, the amino acid sequence encoded by this gene (376 amino acids in total) is identical to that of the mammalian and avian β (type 1) cytoskeletal actin isoforms (Kost et al. 1983; Nudel et al. 1983; Ponte et al. 1984). This is a remarkable conservation of protein sequence, as the divergence of the mammalian and amphibian ancestral lines is estimated to have occurred 300–400 Myr ago (Wilson et al. 1977). One of the differences occurs at amino acid position 10, where,

a

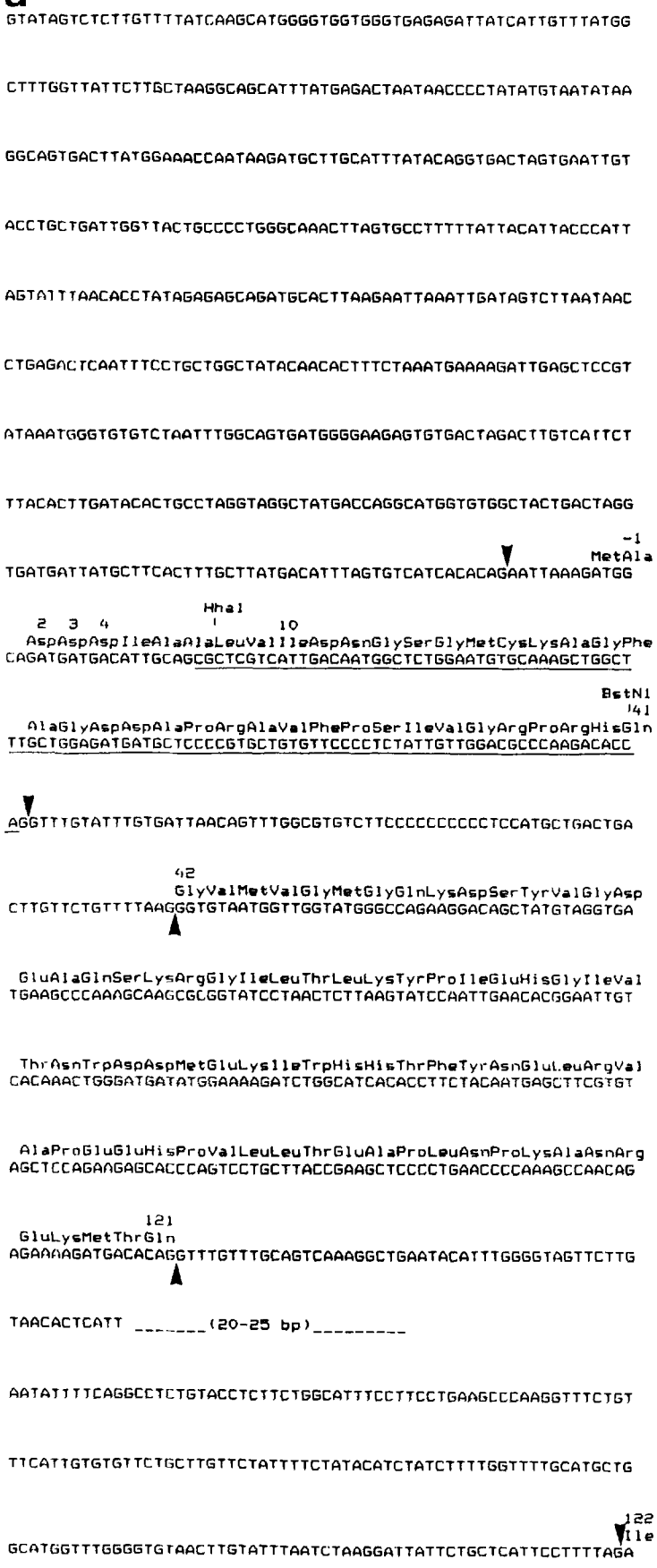

GTATAGTCTCTTGTTTTATCAAGCATGGGGTGGTGGGTGAGAGATTATCATTGTTTATGG
CTTTGGTTATTCTTGCTAAGGCAGCATTTATGAGACTAATAACCCCTATATGTAATATAA
GGCAGTGACTTATGGAAACCAATAAGATGCTTGCATTTATACAGGTGACTAGTGAATTGT
ACCTGCTGATTGGTTACTGCCCCTGGGCAAACTTAGTGCCTTTTTATTACATTACCCATT
AGTATTTAACACCTATAGAGAGCAGATGCACTTAAGAATTAAATTGATAGTCTTAATAAC
CTGAGACTCAATTTCCTGCTGGCTATACAACACTTTCTAAATGAAAAGATTGAGCTCCGT
ATAAATGGGTGTGTCTAATTTGGCAGTGATGGGGAAGAGTGTGACTAGACTTGTCATTCT
TTACACTTGATACACTGCCTAGGTAGGCTATGACCAGGCATGGTGTGGCTACTGACTAGG
-1
MetAla
TGATGATTATGCTTCACTTTGCTTATGACATTTAGTGTCATCACACAGAATTAAAGATGG
HhaI
2 3 4 10
AspAspAspIleAlaAlaLeuValIleAspAsnGlySerGlyMetCysLysAlaGlyPhe
CAGATGATGACATTGCAGCGCTCGTCATTGACAATGGCTCTGGAATGTGCAAAGCTGGCT
BstNI
41
AlaGlyAspAspAlaProArgAlaValPheProSerIleValGlyArgProArgHisGln
TTGCTGGAGATGATGCTCCCCGTGCTGTCTTCCCCTCTATTGTTGGACGCCCAAGACACC
AGGTTTGTATTTGTGATTAACAGTTTGGCGTGTCTTCCCCCCCCCCTCCATGCTGACTGA
42
GlyValMetValGlyMetGlyGlnLysAspSerTyrValGlyAsp
CTTGTTCTGTTTTAAGGGTGTAATGGTTGGTATGGGCCAGAAGGACAGCTATGTAGGTGA
GluAlaGlnSerLysArgGlyIleLeuThrLeuLysTyrProIleGluHisGlyIleVal
TGAAGCCCAAAGCAAGCGCGGTATCCTAACTCTTAAGTATCCAATTGAACACGGAATTGT
ThrAsnTrpAspAspMetGluLysIleTrpHisHisThrPheTyrAsnGluLeuArgVal
CACAAACTGGGATGATATGGAAAAGATCTGGCATCACACCTTCTACAATGAGCTTCGTGT
AlaProGluGluHisProValLeuLeuThrGluAlaProLeuAsnProLysAlaAsnArg
AGCTCCAGAAGAGCACCCAGTCCTGCTTACCGAAGCTCCCCTGAACCCCAAAGCCAACAG
121
GluLysMetThrGln
AGAAAAGATGACACAGGTTTGTTTGCAGTCAAAGGCTGAATACATTTGGGGTAGTTCTTG
TAACACTCATT _____(20-25 bp)_____
AATATTTTCAGGCCTCTGTACCTCTTCTGGCATTTCCTTCCTGAAGCCCAAGGTTTCTGT
TTCATTGTGTGTTCTGCTTGTTCTATTTTCTATACATCTATCTTTTGGTTTTGCATGCTG
122
Ile
GCATGGTTTGGGGTGTAACTTGTATTTAATCTAAGGATTATTCTGCTCATTCCTTTTAGA

Fig. 2. The sequence of regions **a** R2 and **b** R3 from Fig. 1. There are 60 bases on each line. Amino acids of the encoded actin protein have been indicated above their respective codons. Selected amino acids have been numbered above the sequence according to the numbering system derived from rabbit skeletal muscle actin (Elzinga and Lu 1976; Vandekerckhove and Weber 1978b). As the preceding amino acid is probably posttranslationally removed, it is numbered −1. The arrowheads indicate probable splice sites, following the GT/AG rule of Breathnach and Chambon (1981). The underlined sequence indicates the position of the Hha I/Bst NI fragment's noncoding strand used for primer extension.

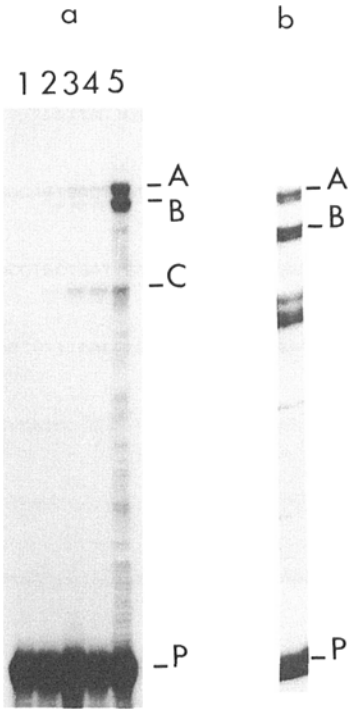


Fig. 3. Primer extension products resulting from hybridizing the primer from λ3H1 to **a** *X. borealis* ovary RNA and **b** *X. laevis* XTC cell RNA. Hybridizations were in 10 μl of hybridization solution containing 1% sodium dodecyl sulfate. In **a** 10 μg of RNA was hybridized overnight at 50°C (track 2), 55°C (track 3), 60°C (track 4), and 70°C (track 5). In **b** 20 μg of RNA was hybridized for 3 h at 70°C. The primer was then extended using reverse transcriptase and the products analyzed on 8% sequencing-type polyacrylamide gels. Electrophoresis was from top to bottom. In track 1 of **a** unhybridized primer was run. The extension products we have sequenced have been labeled. The primer is labeled P.

instead of a valine codon, an isoleucine codon was found. However, an isoleucine occurs at this position in the mammalian γ (type 8) isoform (Vandekerckhove and Weber 1978c). The second difference is the presence of an extra codon, for alanine, after the N-terminal methionine codon of the frog gene. This codon is not found in the human, rat, or chicken β-actin (Kost et al. 1983; Nudel et al. 1983; Ponte et al. 1984), or human γ-actin (Erba et al. 1986) genes that have been sequenced, although it has been found in a type 5 chicken gene (Bergsma et al. 1985). The significance of this is discussed below.

Introns interrupt the coding sequence of the *Xenopus* gene at positions 41/42 (i.e., between the codons for these amino acids), 121/122, 267 (i.e., within the codon for this amino acid), and 327/328. These are in the same positions as the introns in the rat and chicken β-actin genes (Kost et al. 1983; Nudel et al. 1983). No genomic sequences of avian or mammalian γ cytoskeletal actin genes have been published so far.

Confirmation that this gene truly encodes a cytoskeletal actin and not a muscle-specific isoform was provided by the primer extension experiments described below. These demonstrated the existence of transcripts for the gene in the oocyte. We have also found actin gene transcripts in RNA from embryos at all stages of development up to the tadpole, although the abundance of this mRNA species decreases relative to the other *X. borealis* cytoskeletal actin mRNAs (from the type 7 and type 8 genes—see Figs. 3 and 4) between the oocyte and the stage 12 embryo (unpublished data).

5′ Transcriptional Mapping

To map the start of transcription, a primer was made from a Hha I/Bst NI fragment in the first coding exon (see Figs. 1 and 2, and Materials and Methods). The 103-base, 5′-end-labeled, single-stranded primer was hybridized to *X. borealis* ovary or tadpole RNA at various temperatures, extended with reverse transcriptase, and the products separated on a sequencing-type polyacrylamide gel. As shown in Fig. 3, three major products were visualized by autoradiography of such gels. Surprisingly, two of these were obtained in greater amounts when the primer hybridization was at the highest temperature, 70°C in this case, while the optimum temperature for the smallest product was 65°C (data not shown). A similar phenomenon was noted by Woodland et al. (1984) while using a histone H4 primer. It is possible that higher temperatures can promote hybridization by the destruction of mRNA secondary structure. As with histone mRNAs, there are, in fact, leader and coding regions in the actin gene which could base pair with the primer-hybridizing part of the mRNA. The primer extension products obtained using the actin primer were eluted from the gels and sequenced by the method of Maxam and Gilbert (1978). The material used for this sequencing typically consisted of a doublet of bands. We believe these doublets result from extension on only one mRNA species, because the sequence obtained was unambiguous, and because such a phenomenon has previously been reported for a single species of globin mRNA (Luse et al. 1981).

The sequenced portions of the three primer extension products are shown in Fig. 4. Part of the sequence of the largest product exactly matches the coding sequence of the type 1 gene, and the first eight bases 5′ to this. The remaining 50 bases of product A sequence are identical to a region approximately 1200 bases upstream in the genomic clone. The 5′ untranslated leaders of all other sequenced vertebrate actin genes are also interrupted by an intron (Fornwald et al. 1982; Zakut et al. 1982; Kost et al. 1983; Nudel et al. 1983; Chang et al. 1985). To determine the start of transcription, primer

```
                                                    met ala asp asp asp ile ala
--CTTGCTAGCTTTGCTTTTTTTTTCCCCGTAAAGGAAACTTCCAGCACCAAGAATTAAAG ATG GCA GAT GAT GAC ATT GCA G    X.bor A = Type 1

  ***************   ****** ***** ******** *  ** ********** **** met ala glu glu glu ile ala
--CTTGCTAGCTTTGC---TTTTTTCCCGTTAAGGAAACCTTAAGTACCAAGAATAAAAG ATG GCA GAA GAA GAG ATT GCA G    X.bor B = Type 8

                  ** *******  ** **  ************** met ala glu glu glu ile ala
                --GTTAAGGAAATCTCTAGTTCCAAGAATTAAAG ATG GCA GAA GAA GAG ATT GCA G    X.lae A = Type 8

                  ** *************** *************** met ala asp glu glu ile ala
                --GTCAAGGAAACTTCCAGTACCAAGAATTAAAG ATG GCA GAT GAA GAG ATT GCA G    X.lae B = Type 5

                 *     *             ***   *  * *   met glu glu asp ile ala
              --GGAGACAATTCGTGTTTGCGTCCTACCCTCAGATCACA ATG GAA GAA GAT ATT GCT G       X.bor C = Type 7
```

Fig. 4. Partial sequence of *X. borealis* and *X. laevis* mRNAs derived from sequencing primer extension products. The sequences of *X. borealis* products A and B (see Fig. 3) were obtained using ovary RNA, and tadpole RNA was used to obtain the product C sequence. XTC cell RNA was used to obtain the sequence of *X. laevis* products A and B. In each case the sequence of the leader region is incomplete. We have indicated the type of actin encoded by each sequence, using nomenclature suggested by Vandekerckhove et al. (1981). We have compared the leader of the *X. borealis* type 1 mRNA with the other four mRNAs. An asterisk above a base indicates that it is the same as the equivalent base of the *X. borealis* type 1 leader. A gap has been introduced into the *X. borealis* type 8 mRNA sequence to optimize the upstream homology between this leader and that of the type 1 gene.

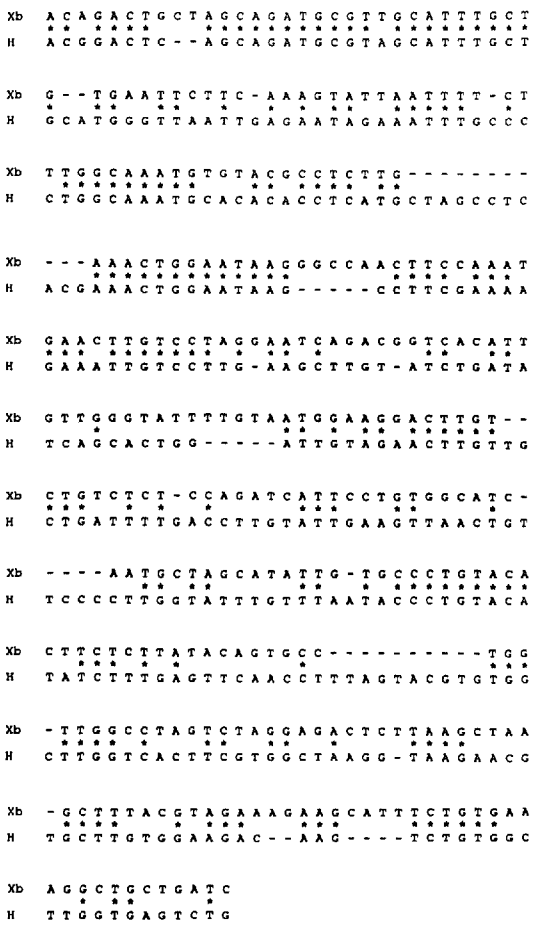
```
                -120              CAAT ?    1              2    -90
CCGGTACCGCCTCTTCTCTCAAGCAGGCGGCCAATCGCGAGGCGTGTTTCTGAAAGATGC

CAAT ?          -60                                          3   -30
CCATATTTGGTGATCTTCTATCTACGCNACGCCAGGCGGGGCTGGAAGCCTGTGACGTTA

TATA ?                          -1
AATAGAGATAACTGGCGCTATTAAATTCTCATTCAGCCGGTGGCAGTGGACCTGACGTTT

                        4
ATTTACTTTCTAGCTTCCTTGCTAGCTTTGCTTTTTTTTTCCCCGTAAAGGAAACTTCCA

GCACCAAGGTAAGACCTGCGGCTGTGTTCTTTTGGAGGGTGACTAACAAGAAGTCGGCGT

GTGGCTGTGCAATGGGTGGAAGAGCGCTTGGTGGCGG
```

Fig. 5. Sequenced portion of the region R1 from Fig. 1, including exon 1 and the region immediately upstream from the start of transcription. Each complete line consists of 60 bases. The large vertical arrowhead indicates the 5′ end of the leader intron. The horizontal arrow indicates the direction and start of transcription. The exact point of the latter is not known; the two small vertical arrows indicate possible positions. The bases 5′ to the start of transcription have been given negative numbers. Sequences representing possible promoter elements have been underlined. Sequences that are wholly or partly homologous to similar regions in other cytoskeletal genes have been overlined. Sequence 1 is found 100–110 nucleotides upstream from the start of transcription of the rat and chicken β-actin genes (Kost et al. 1983; Nudel et al. 1983). Sequence 2 has homology to the first 16 bases of the 25-nt homology block between the CAAT and TATA boxes of the rat and chicken β-actin genes (Kost et al. 1983; Nudel et al. 1983). The rat sequence is the same in 11 positions and the chicken sequence is the same in 10 positions. Seven of the eight bases of sequence 3 are found immediately upstream from the TATA box of the rat β gene (Nudel et al. 1983). Eleven of the 12 bases in sequence 4 are also found in part of the 5′ untranslated region of the rat β gene.

extension product A was sized by running it next to a set of sequencing tracks. We estimate the accuracy to be within two bases. As shown in Fig. 5, 29–30 bases upstream from the start of transcription (the "cap-site") lies the beginning of the sequence 5′-TTAAATA-3′. This resembles the "TATA box" promoter element [with consensus sequence $(TATA^{T}_{A}A^{T}_{A}A)$] found in this position in many eukaryotic polymerase II genes (Breathnach and Chambon 1981). Two potential "CAAT boxes" can also be found. One sequence, 5′-GCCCATA-3′, lies approximately 90 bases upstream from the start of transcription, and hence in a similar position to that found in many other genes (Breathnach and Chambon 1981). Another sequence, 5′-GGCCAATCG-3′, more closely resembles the consensus sequence for this promoter element, $GG^{C}_{T}CAATCT$ (Benoist et al. 1980; Efstratiadis et al. 1980), but is situated 120 bases from the cap site.

Previously, searches for regions of homology in the 5′ upstream regions of similar genes in different species have been made in order to identify sequences that may be involved in transcriptional control. Thus, a 20-nt-long region of homology around, and including, the CAAT box, has been found in the rat and chicken skeletal muscle actin genes (Ordahl and Cooper 1983). A block of 25 nucleotides between the CAAT and TATA boxes is very similar in the rat and chicken β-actin genes (Kost et al. 1983; Nudel et al. 1983), differing in the

a



b

Fig. 6. Sequence comparisons between **a** the 3′UTRs and **b** the 5′UTRs of the *X. borealis* (Xb) type 1 gene and a human (H) γ cytoskeletal cDNA clone (Erba et al. 1986). In **a** the comparison begins immediately after the termination codons and proceeds for the full 318 bases of downstream *Xenopus* sequence present in λ3H1. In **b** a portion of the leader of the *X. borealis* gene, beginning with the start of transcription, is compared with a similar sequence in the human gene. Outside this region there is little homology between the two leaders. Asterisks indicate positions where bases are the same. Gaps have been introduced in both sequences to maximize homology.

two genes by only four bases. This region lies between 50 and 75 bases upstream from the cap site of both genes.

We have compared the *X. borealis* type 1 gene with the rat and chicken β (type 1) genes and found several small regions of homology in the leader and 5′ flanking regions. Of the 16 most distal bases in the 25-nt homology block mentioned above, 11 bases of the rat sequence and 10 bases of the chicken sequence are the same within the region, 5′-GTTTCTGAAAGATGCC-3′. This lies between 102 and 87 bases upstream from the cap site (i.e., between the TATA box and most distal putative CAAT box of the *Xenopus* gene). The sequence 5′-CTGTGPuCG-3′ occurs a few bases upstream from the frog and rat TATA boxes, but not in that of the chicken. The sequence 5′-GCGAGGC-3′ is found in all three genes, however, at approximately the same distance from the cap site (100–110 nucleotides), and if the most proximal potential CAAT box is used for the frog gene, the same distance (approximately 20 nt) upstream from the CAAT boxes. In the leader region of the frog gene, a 12-base segment was found to have 11 bases in common with part of the rat β gene leader (see Fig. 5), but this sequence was not in the chicken leader. However, a comparison of the frog leader with that of a human γ (type 8) gene (Erba et al. 1986) revealed a high degree of homology in the first 20 bases (see Fig. 6b) with little homology elsewhere. It seems likely that this region of the leader has been conserved for functional reasons, but further comparisons with other species are required to ascertain if the smaller length homologies described above are meaningful or merely the result of chance.

3′ Untranslated Region (3′UTR)

It is probable that all of the 318 bases of cloned DNA 3′ to the termination codon of this *X. borealis* gene are transcribed, as two very similar genes of the closely related species *X. laevis* (see Fig. 4) both have 3′UTRs approximately 800 bases in length (T. Mohun, personal communication). Fig. 6a shows a sequence comparison between this 3′ region and the 3′UTR of a human γ (type 8) cytoskeletal actin cDNA clone (Erba et al. 1986). There are many blocks of complete homology in the first 120 bases with a regions of homology in the remainder. Apart from the first 15 bases, there is far less homology when the frog type 1 gene is compared with the rat, human, or chicken β-actin 3′UTRs (Kost et al. 1983; Nudel et al. 1983; Ponte et al. 1984). This is a remarkable example of noncoding sequence conservation, and it is presumably maintained for functional reasons. It may be relevant to its function that the region of 21 bases from bases 11 to 31 of the 3′UTR of the *X. borealis* gene is capable of forming a hairpin loop structure (see Fig. 7). In this region there is only one base difference between the *Xenopus* and human genes. Exceptionally high levels of 3′UTR conservation have also been observed in previous comparisons of other actin genes (Yaffe et al. 1985). Regions of high homology have been reported between the 3′UTRs of the human and *X.*

laevis cardiac muscle actin genes (Gunning et al. 1984), between the 3′UTRs of human, rat, and chicken skeletal muscle actin genes (Gunning et al. 1984), and between the 3′UTRs of the human, rat, and chicken β cytoskeletal actin genes (Kost et al. 1983; Nudel et al. 1983; Ponte et al. 1984). In contrast, apart from two small regions of homology between the rat cardiac and skeletal muscle gene 3′UTRs (Mayer et al. 1984), there is little homology between the 3′UTRs of the different isoforms within a species. This conservation of 3′UTR sequence, therefore, together with the 5′UTR conservation shown in Fig. 6b, suggests that the gene we have cloned is functionally, and presumably evolutionarily, related to the mammalian and avian γ cytoskeletal genes, rather than to the β genes (even though the three N-terminal amino acids are in the β configuration). This is also supported by the presence of an isoleucine codon as position 10, although this could be produced by only a single base substitution.

The function of the 3′UTR conservation in certain genes is not known. Yaffe et al. (1985) have suggested that in some cases, such as the c-fos gene (Miller et al. 1984), it is because this region is involved in translational control. Evidence has been found suggesting that the translational control of cytoskeletal actins takes place in the early embryonic stages of *X. laevis* (Ballatine et al. 1979). Alternative explanations are possible, however. For example, this region may be important in message stability.

Expression of Other Cytoskeletal Actins in Xenopus

The sequences of the other *X. borealis* primer extension products (see Fig. 3) are shown in Fig. 4. Product B represents a type 8 isoform, and like product A has an alanine after the N-terminal methionine, whereas product C represents a type 7 isoform with no alanine in this position (position −1). The type 7 isoform has not previously been found in any other species. Thus, in total, we have found mRNAs encoding isoform types 1, 7, and 8 in *X. borealis*.

By the partial sequencing of N-terminal tryptic peptides, Vandekerckhove et al. (1981) found the type 4, 5, and 8 isoforms in several nonmuscle tissues of *X. laevis*. Shown in Fig. 4 are the sequences of the two largest primer extension products produced as a result of hybridizing our *X. borealis* primer to XTC cell (an *X. laevis* fibroblast-like cell line; Pudney et al. 1973) RNA (see Fig. 3b). These sequences encode type 5 and type 8 actin isoforms, two of the isoforms found by Vandekerckhove et al. (1981). Primer extension bands probably representing at least one more mRNA species were also detected using *X. laevis* RNA, but it has not yet been

Fig. 7. Predicted hairpin loop structure formed from bases 11 to 31 of the 3′UTR of the *X. borealis* type 1 mRNA

possible to sequence these (although we would predict the existence of an mRNA species encoding a type 4 isoform). We have, therefore, partially confirmed the results of Vandekerckhove et al. (1981). This difference in actin isoform content between *X. borealis* and *X. laevis* is surprising, as these species are thought to have diverged only 8 Myr ago (Bisbee et al. 1977). Our ability to detect the type 5 isoform of *X. laevis* argues against the possibility that we were unable to detect an *X. borealis* type 5 mRNA because it hybridized poorly to our primer. Although more complex possibilities exist, it seems likely that we have detected all of the major cytoskeletal actin mRNAs in the *X. borealis* ovary and *X. laevis* XTC cells. We have also found the same primer extension bands using *X. borealis* tadpole and *X. laevis* ovary RNA, so these mRNA species are unlikely to be tissue specific.

Despite encoding different isoforms, the *X. laevis* type 5 and *X. borealis* type 1 mRNAs have very similar leader sequences. As shown in Fig. 4, these are also closely related to the *X. laevis* and *X. borealis* type 8 mRNA leaders. It is possible that these genes have all arisen recently from a common ancestral gene. Another, and not necessarily alternative, possibility is that the common ancestor to these two species possessed genes for all three of these isoforms, the type 1 being later silenced in *X. laevis*, and the type 5 in *X. borealis*. Within one species, the similarity of the leaders of the two genes may be the result of a gene conversion event or may reflect their formation by a recent duplication event. In this latter respect it is very relevant that *X. borealis* and *X. laevis* are believed to be tetraploid species, and are indeed more or less tetraploid relative to *X. tropicalis*, which diverged from the other two species approximately 30 Myr ago (Theibaud and Fischberg 1977; Tymowska and Fischberg 1977). It is not known if the duplication of the genome preceded the divergence of the two tetraploid species.

As can be seen in Fig. 4, the *X. laevis* type 5 and

type 8 mRNAs also have the extra alanine codon (at position 1) found in the *X. borealis* type 1 and type 8 mRNAs. Vandekerckhove et al. (1981) determined the N-terminal protein sequences of these *X. laevis* isoforms and found that they began with the acidic residues, blocked by an acetyl group. Therefore, the alanine must be posttranslationally removed, along with the first methionine.

Evolution of Actin Genes

The *X. borealis* cytoskeletal actin gene we have cloned contains introns at the same positions in the coding region as the β cytoskeletal actin genes of the rat and chicken (Kost et al. 1983; Nudel et al. 1983): at positions 41/42, 121/122, 267, and 327/328. The four mammalian and avian striated muscle-specific actin genes that have been sequenced, however, all contain introns at a different set of positions: at codons 41/42, 150, 204, 267, and 327/328 (Fornwald et al. 1982; Hamada et al. 1982; Zakut et al. 1982; Chang et al. 1985). A human aortic-type smooth muscle actin gene has been sequenced and found to contain a set of introns that is the same as those of the striated muscle-specific genes, but with additional introns at positions 84/85 and 121/122 (Ueyama et al. 1984). Our results suggest, therefore, that the cytoskeletal and muscle-specific genes became separate before the divergence of the amphibians from the rest of the vertebrates. In support of this, we have sequenced (unpublished data) part of a *X. borealis* muscle-specific actin gene, encoding amino acids from the N-terminus to position 150, and found it to contain introns in the same positions as the mammalian striated muscle-specific genes (i.e., at positions 41/42 and 150). Further, the amino acid sequence of this frog actin was identical to that of the chicken and human cardiac actins (Hamada et al. 1982; Chang et al. 1985). These results with amphibian genes support the evolutionary proposals of Vandekerckhove and Weber (1984). They found that the muscle-specific actins of invertebrates did not vary greatly in amino acid sequence from the cytoskeletal isoforms of vertebrates and invertebrates. However, they obtained the complete amino acid sequence of a lamprey muscle-specific actin and found it to resemble the sequence of mammalian and avian muscle-specific isoforms. This suggested that this type of muscle-specific actin gene arose prior to the divergence of jawless fish from the rest of the craniates.

The *X. borealis* type 1 and type 8 and the *X. laevis* type 5 and type 8 cytoskeletal actin mRNAs differ from the *X. borealis* type 7 mRNA, the β-actin genes of humans, rats, and chickens (Kost et al. 1983; Nudel et al. 1983; Ponte et al. 1984), and the human γ cytoskeletal actin gene (Erba et al. 1986) in possessing an extra codon, for alanine, after the N-terminal methionine. Have these different gene types [ala(+) or ala(−)] arisen from an ancestral gene by the deletion or addition of this extra alanine codon? The eight vertebrate muscle-specific actin genes sequenced so far (Fornwald et al. 1982; Hamada et al. 1982; Zakut et al. 1982; Hanauer et al. 1983; Chang et al. 1985; Stutz and Spohr 1986) have a cysteine codon following the N-terminal methionine codon. This is also true for all six *Drosophila* genes (Fyrberg et al. 1981), all four nematode genes (Files et al. 1983), and two sea urchin genes (Schuler et al. 1983). It seems likely, therefore, that the ancestral prechordate actin gene was of this cys(+) type. The cysteine codon was presumably retained in all invertebrate actin genes as well as vertebrate muscle-specific actin genes, but has been mutated to an alanine in some vertebrate cytoskeletal actin genes and deleted in others. Unless two point mutations occurred simultaneously, this change from TGPy to GCN would have involved an intermediate codon for a third type of amino acid.

Recently, Bergsma et al. (1985) have identified an ala(+) cytoskeletal actin gene in chickens encoding a type 5 (see Table 1) isoform. They found that this gene possessed just a single intron, in its leader, yet could find no evidence for it being a pseudogene. This contrasts with the five introns found in all other vertebrate cytoskeletal actin genes (Kost et al. 1983; Nudel et al. 1983), including the *X. borealis* ala(+) gene described here. Further information is needed to fit this gene into an evolutionary scheme. However, the fact that the soybean also contains an ala(+) actin gene (Shah et al. 1982) suggests that cysteine/alanine changes have probably occurred more than once in evolution.

In all four *Xenopus* ala(+) genes examined here (see Fig. 4), as well as the chicken type 5 and soybean genes, the same alanine codon, GCA, is found after the N-terminal methionine. Other synonymous alanine codons are more frequently used in the rest of the *X. borealis* type 1, chicken type 5, and soybean actin genes, however. GCA is used in only 4 out of 29 positions for each of the frog and chicken genes, and in 9 out of 28 positions in the soybean. This suggests that the role of the alanine codon following the N-terminal methionine may be in its contribution to the DNA or RNA sequence in this region, rather than simply encoding alanine. It is interesting in this respect that the six bases of leader sequence 5′ to the N-terminal methionine codon of the soybean and *X. borealis* type 8 genes are identical, giving a total run of 16 identical bases.

It seems likely that the role of the cysteine codon found after the N-terminal methionine of cys(+) genes is also in its contribution to the DNA or RNA sequence. This cysteine, like the alanine in the

equivalent position of ala(+) genes, must be removed posttranslationally, as it has not been found at the terminus of any mature actin protein (Vandekerckhove and Weber 1978a). In all the cys(+) actin genes that have been sequenced, except the human skeletal muscle gene (Hanauer et al. 1983), the more rarely used synonymous codon TGT is found in this position. Moreover, this is the only position in which this codon is used out of the six cysteine codons of the chicken skeletal muscle gene (Fornwald et al. 1982), the six cysteine codons each of the *X. laevis* cardiac and skeletal muscle genes (Stutz and Spohr 1986), the six cysteine codons of the sea urchin pSpG17 gene (Cooper and Crain 1982), the four cysteine codons of the sea urchin gene 1 (Schuler et al. 1983), and the seven cysteine codons of both the 79B and 88F genes in *Drosophila* (Sanchez et al. 1983). It is possible that structural demands favor certain sequences around the start of translation of actin mRNAs.

If conservation of such a sequence takes place, it does not seem to extend, at least in amphibians, to the acidic amino acid codons at positions 2–4. The difference in amino acids in these positions between the cytoskeletal actins of two such closely related species as *X. laevis* and *X. borealis* (which can form viable hybrids) as well as the differences found in other amphibians (Vandekerckhove et al. 1981) appear to indicate that there are no functional differences between such isoforms. It is also possible, though on the face of it less likely, that such differences are adaptive, as the actin genes may be subject to different selective pressures even in such similar animals as *X. laevis* and *X. borealis*.

Acknowledgments. We thank Elizabeth Ballantine for her technical assistance, Phil Turner and Bill Bains for advice and materials, Susan Easton for her preparation of the manuscript, and the Medical Research Council for financial support. Harry Erba is a predoctoral fellow of the Medical Scientist Training Program.

References

Baines W (1982) Thesis, University of Warwick, Coventry UK

Ballantine JEM, Woodland HR, Sturgess EA (1979) Changes in protein synthesis during the development of *Xenopus laevis*. J Embryol Exp Morphol 51:137–153

Bender W, Davidson N, Kindle KL, Taylor WC, Silverman M, Firtel RA (1978) The structure of M6, a recombinant plasmid containing *Dictyostelium* DNA homologous to actin messenger RNA. Cell 15:779–788

Benoist C, O'Hare K, Breathnach R, Chambon P (1980) The ovalbumin gene-sequence of putative control regions. Nucleic Acids Res 8:127–142

Benton WD, Davis RW (1977) Screening lambda-gt recombinant clones by hybridization to single plaques *in situ*. Science 196:180–182

Bergsma DJ, Chang KS, Schwartz RJ (1985) Novel chicken actin gene: third cytoplasmic isoform. Mol Cell Biol 5:1151–1162

Bisbee CA, Baker MA, Wilson AC, Hadji-Azimi I, Fischberg M (1977) Albumin phylogeny of the clawed toad *X. laevis*. Science 195:785–787

Breathnach R, Chambon P (1981) Organization and expression of eukaryotic split genes coding for proteins. Annu Rev Biochem 50:349–383

Bremer JW, Busch H, Yeoman LC (1981) Evidence for a species of nuclear actin distinct from cytoplasmic and muscle actins. Biochemistry 20:2013–2017

Chang KS, Rothblum KN, Schwartz RJ (1985) The complete sequence of the chicken alpha-cardiac gene: a highly conserved vertebrate gene. Nucleic Acids Res 13:1223–1237

Cooper AD, Crain WR (1982) Complete nucleotide sequence of a sea-urchin actin gene. Nucleic Acids Res 10:4081–4092

Efstratiadis A, Posakony JW, Maniatis T, Lawn RM, O'Connell C, Spritz RA, DeRiel JK, Forget BG, Weissman SM, Slightom JL, Blechl AE, Smithies O, Baralle FE, Shoulders CC, Proudfoot NJ (1980) The structure and evolution of the human beta-globin gene family. Cell 21:653–668

Elzinga M, Lu RC (1976) In: Perry SV, Margreth A, Adelstein RS (eds) Contractile systems in non-muscle tissues. North-Holland, Amsterdam, pp 29–37

Erba HP, Gunning P, Kedes L (1986) Nucleotide sequence of the human γ cytoskeletal actin mRNA; anomalous evolution of vertebrate non-muscle actin genes. Nucleic Acids Res 14: 5275–5294

Files JG, Carr J, Hirsh D (1983) The actin gene family of *Caenorhabditis elegans*. J Mol Biol 164:355–375

Fornwald JA, Kuncid G, Peng I, Ordahl CP (1982) The complete nucleotide sequence of the chick alpha-actin gene and its evolutionary relationship to the actin gene family. Nucleic Acids Res 10:3861–3876

Fyrberg EA, Bond BJ, Hershey ND, Mixter KS, Davidson N (1981) The actin genes of *Drosophila*: protein coding regions are highly conserved but intron positions are not. Cell 24: 107–116

Gunning P, Mohun T, Ng SY, Ponte P, Kedes L (1984) Evolution of human sarcomeric-actin genes: evidence for units of selection within the 3′ untranslated regions of the mRNAs. J Mol Evol 20:202–214

Hamada H, Perrine MG, Kakunaga T (1982) Molecular structure and evolutionary origin of human cardiac muscle actin gene. Proc Natl Acad Sci USA 79:5901–5905

Hanauer A, Levin M, Heilig R, Daegelen D, Kalm A, Mandel JL (1983) Isolation and characterization of cDNA clones for human skeletal muscle alpha-actin. Nucleic Acids Res 11: 3503–3516

Kost TA, Theodorakis N, Hughes SH (1983) The nucleotide sequence of the chick cytoplasmic alpha-actin gene. Nucleic Acids Res 11:8287–8301

Luse DS, Haynes JR, VanLeeuwen D, Schon EA, Cleary ML, Shapiro SG, Lingrel JB, Roeder RG (1981) Transcription of the beta-like globin genes and pseudogenes of the goat in a cell-free system. Nucleic Acids Res 9:4339–4354

Marotta CA, Strocchi P, Gilbert JM (1978) Microheterogeneity of brain cytoplasmic and synaptoplasmic actins. J Neurochem 30:1441–1451

Maxam AM, Gilbert W (1978) Sequencing end-labelled DNA with base-specific chemical cleavages. Methods Enzymol 65: 499–559

Mayer Y, Czosnek H, Zeelon PE, Yaffe D, Nudel U (1984) Expression of the genes coding for the skeletal muscle and cardiac actins in the heart. Nucleic Acids Res 12:1087–1100

Miller AD, Curran T, Verma M (1984) c-fos protein can induce cellular transformation: a novel mechanism of activation of a cellular oncogene. Cell 36:51–60

Nudel U, Zakur R, Shani M, Neuman S, Levy Z, Yaffe D (1983) The nucleotide sequence of the rat cytoplasmic beta-actin gene. Nucleic Acids Res 11:1759–1771

Ordahl CP, Cooper TA (1983) Strong homology in promoter and 3′-untranslated regions of chick and rat alpha-actin genes. Nature 303:348–349

Ponte P, Ng SY, Engel J, Gunning P, Kedes L (1984) Evolutionary conservation in the untranslated regions of actin mRNAs: DNA sequence of a human beta-actin cDNA. Nucleic Acids Res 12:1687–1696

Pudney M, Varma MGR, Leake CJ (1973) Establishment of a cell line XTC-2 from the South African clawed toad, *Xenopus laevis*. Experientia 29:466–467

Sanchez F, Tobin SL, Rdest U, Zulauf E, McCarthy BJ (1983) Two *Drosophila* actin genes in detail: gene structure, protein structure and transcription during development. J Mol Biol 163:533–551

Sanger F, Coulson AR, Barrell BG, Smith AJH, Roe BA (1980) Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. J Mol Biol 143:161–178

Schuler MA, McOsker P, Keller EB (1983) DNA sequence of two linked actin genes of the sea urchin *S. purpuratus*. Mol Cell Biol 3:448–456

Shah DM, Hightower RC, Meagher RB (1982) Complete nucleotide sequence of a soybean actin gene. Proc Natl Acad Sci USA 79:1022–1026

Smith HO, Birnstiel ML (1976) A simple method for DNA restriction site mapping. Nucleic Acids Res 3:2387–2398

Southern EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 98:503–517

Stutz F, Spohr G (1986) Isolation and characterisation of sarcomeric actin genes expressed in *Xenopus laevis* embryos. J Mol Biol 187:349–361

Theibaud CH, Fischberg M (1977) DNA content in the genus *Xenopus*. Chromosoma 59:253–257

Tymowska J, Fischberg M (1973) Chromosome complements of the genus *Xenopus*. Chromosoma 44:335–342

Ueyama H, Hamada H, Battula N, Kakunaga T (1984) Structure of a human smooth muscle actin gene (aortic type) with a unique intron site. Mol Cell Biol 4:1073–1078

Vandekerckhove J, Weber K (1978a) At least six different actins are expressed in a higher mammal: an analysis based on the amino acid sequence of the amino-terminal tryptic peptide. J Mol Biol 126:783–802

Vandekerckhove J, Weber K (1978b) Actin amino acid sequences. Comparison of actins from calf thymus, bovine brain, and SV40-transformed mouse 3T3 cells with rabbit skeletal muscle actin. Eur J Biochem 90:451–462

Vandekerckhove J, Weber K (1978c) Mammalian cytoplasmic actins are the products of at least two genes and differ in primary structure in at least 25 identified positions from skeletal muscle actins. Proc Natl Acad Sci USA 75:1106–1110

Vandekerckhove J, Weber K (1979) The complete amino acid sequences of actins from bovine aorta, bovine heart, bovine fast skeletal muscle, and rabbit slow skeletal muscle. Differentiation 14:123–133

Vandekerckhove J, Weber K (1981) Actin typing on total cellular extracts: a highly sensitive protein-chemical procedure able to distinguish different actins. Eur J Biochem 118:595–603

Vandekerckhove J, Weber K (1984) Chordate muscle actins differ distinctly from invertebrate muscle actins. J Mol Biol 179:391–413

Vandekerckhove J, Franke WW, Weber K (1981) Diversity of expression of non-muscle actin in Amphibia. J Mol Biol 152: 413–426

Wilson AC, Carlson SS, White TJ (1977) Biochemical evolution. Annu Rev Biochem 46:573–639

Woodland HR, Warmington JR, Ballantine JEM, Turner PC (1984) Are there major developmentally regulated H4 gene classes in *Xenopus*? Nucleic Acids Res 12:4939–4958

Yaffe D, Nudel U, Mayer Y, Neuman S (1985) Highly conserved sequences in the 3′ untranslated region of mRNAs coding for homologous proteins in distantly related species. Nucleic Acids Res 13:3723–3737

Zakut R, Shani M, Givol D, Neuman S, Yaffe D, Nudel U (1982) The nucleotide sequence of the rat skeletal muscle actin gene. Nature 298:857–859

Received July 24, 1986/Revised May 2, 1987