J Mol Evol (1994) 39:347–356

Journal of Molecular Evolution

Evolution of Actin Gene Families of Sea Urchins

Hung Fang, Bruce P. Brandhorst

Institute of Molecular Biology and Biochemistry, Simon Fraser University, Burnaby, B.C., V5A 1S6, Canada

Received: 20 January 1994 / Accepted: 26 April 1994

Abstract. The actin gene family of the sea urchin *Lytechinus pictus* includes a single muscle actin gene, LpM, and four cytoskeletal actin genes: LpC1, LpC2, LpC3, and LpC4. The origin and relationship of these actin genes to members of the actin gene family of the sea urchin *Strongylocentrotus purpuratus* were considered. Comparison of deduced amino acid sequences suggested a close relationship between LpC1 and the CyI-CyII subfamily of *S. purpuratus* actin genes, and between LpC2 and the CyIII subfamily of *S. purpuratus* actin genes; the muscle actin genes were orthologous. It is proposed that two divergent cytoskeletal actin genes of the common ancestral sea urchin gave rise by duplication to the extant cytoskeletal actin genes of these species, some of which have changed 3′ noncoding sequences while others have maintained a terminus highly conserved among sea urchin actin genes.

Key words: Actin gene — Gene evolution — Sea urchin — Echinoids — Echinoderms — Codon bias

Introduction

Actin genes were among the first to be cloned, sequenced, and compared for a wide variety of phylogenetically diverse species. The number of genomic organization of actin genes are highly variable. Amino acid replacements accumulate in actin proteins at a very low rate, and there is a high degree of similarity of nucleotide coding sequence among actin genes (Hightower and Meagher 1986). The conservation of the actin gene family provides an opportunity to investigate its evolution. All chordates, echinoderms, and insects examined to date include at least one actin gene exclusively expressed in muscle and multiple cytoskeletal (nonmuscle) actin genes expressed in other types of cells. The vertebrate muscle and cytoskeletal actin genes can be distinguished on the basis of a set of diagnostic amino acid differences (Vandekerckhove and Weber 1984). The muscle actin genes of chordates and echinoderms apparently arose in a common deuterostome ancestor, probably from an echinoderm cytoskeletal actin gene (Kovilur et al. 1993). The muscle actin genes of insects arose independently (Mounier et al. 1992).

The actin gene family of the sea urchin *Strongylocentrotus purpuratus* is the most completely characterized among echinoderms (Cooper and Crain 1982, Davidson et al. 1982). It consists of at least eight nonallelic actin genes (Lee et al. 1984). A single functional muscle (M) actin gene has been identified, as well as five functional cytoskeletal actin genes: the closely linked CyI, CyIIa, and CyIIb genes and the closely linked CyIIIa and CyIIIb genes (Durica et al. 1980; Scheller et al. 1981; Lee et al. 1984; Akhurst et al. 1987; Flytzanis et al. 1989). The transcripts of these cytoskeletal actin genes are distinguishable by sequence elements present in the 3′ noncoding regions of each mRNA. Differences in 3′ noncoding sequences and organization of linkage groups have been used to classify these cytoskeletal actin genes into three subfamilies (Lee et al. 1984). The congeneric species *S. drobachiensis* and *S. franciscanus* also have CyI, linked CyIIa and CyIIb, and two CyIII actin genes (Lee et al. 1984). This indicates that the actin gene sub-

Correspondence to: B.P. Brandhorst

families of *S. purpuratus* were present before these three species diverged. The actin genes of sea stars have been partially characterized and include muscle and multiple cytoskeletal forms which differ in number between asteroid species (Kovesdi and Smith 1985; Kowbel and Smith 1989).

The timing and spatial domains of expression of the actin genes are differentially regulated during the *S. purpuratus* life cycle (Crain et al. 1981; Garcia et al. 1984; Shott et al. 1984; Cox et al. 1986; Lee et al. 1986). Regulation of *S. purpuratus* actin genes has been investigated in transgenic sea urchin embryos microinjected with actin promoter-reporter fusion genes (Hough-Evans et al. 1987; Flytzanis et al. 1989; Franks et al. 1988; Collura and Katula 1992; Niemeyer and Flytzanis 1993) and in reciprocal hybrids of *S. purpuratus* and *Lytechinus pictus* (Bullock et al. 1988; Nisson et al. 1989). The mechanism of transcriptional regulation of some homologous genes appears to be different in embryos of these sea urchin species (Brandhorst and Klein 1992).

We are interested in the extent to which mechanisms regulating the spatial specificity of gene expression are conserved among sea urchin species and among related genes. We have characterized the expression *L. pictus* actin genes in embryos and find similarities and differences in comparison with actin gene expression in *S. purpuratus* embryos (our unpublished observations). *S. purpuratus* (family Echinometridae; subfamily Strongylocentrodinae) and *L. pictus* (family Echinometridae; subfamily Toxopneustinae) diverged 30–40 mya (Smith 1988). For interpretation of comparative investigations of actin gene expression in these species it is important to know whether the common ancestor had the same subfamilies of actin genes present in the *S. purpuratus* genome. Is the spatially restricted pattern of differential actin gene expression in *S. purpuratus* embryos an ancient or recent phenomenon? We thus sought to identify and characterize the expression of orthologous members of the *S. purpuratus* actin gene family in *L. pictus* embryos. We have recently cloned all five nonallelic genes in the *L. pictus* actin gene family (our unpublished data). In this report we compare the sequences of these genes and consider the relationships among sea urchin actin genes. We propose a model to account for the extant cytoskeletal actin genes of these sea urchins in which the common ancestral species already had two cytoskeletal actin genes which, after divergence of the species, duplicated to give rise to the CyI/CyII and CyIII subfamilies of *S. purpuratus* actin genes.

Materials and Methods

There are five actin genes of the sea urchin *L. pictus* including four cytoskeletal genes LpC1–C4 and a single muscle actin gene LpM; we have cloned and sequenced cDNAs corresponding to LpC1 and LpC2 and cloned and sequenced genomic DNA corresponding to the 3′ coding and noncoding regions of LpC3, LpC4, and LpM (manuscript in preparation). All other actin gene sequence data involved in this study were obtained from the Genbank database release 77.0 (Benson et al. 1993). The accession numbers are, for the sea star *Pisaster ochraceus,* M26500 for the muscle actin gene PoM and M26501 for the cytoskeletal actin gene PoCy (Kowbel and Smith 1989); for the sea urchin *S. purpuratus,* V01350 for CyI (Schuler et al. 1983), V01349 for CyIIa (Schuler et al. 1983), M35323 for CyIIb (Durica et al. 1988), M29808 and M30511 for CyIIIa (Akhurst et al. 1987), M35324 for CyIIIb (Durica et al. 1988), and X05739–X05744 for the muscle actin gene SpM (Crain et al. 1987); for the sea urchin *S. francisacanus,* X03075 for the Sf15A and X03076 for the Sf15B cytoskeletal actin genes (Foran et al. 1985); and L13787 for the ascidian *Styela clava* alpha-muscle actin gene ScM (Beach and Jeffery 1992).

Sequences were aligned and compared with the aid of the eyeball sequence editor, ESEE (Cabot and Beckenbach 1989) and the PCGene software package (IntelliGenetics, version 6.5). The numbering of deduced amino acid residues was according to the convention of Lu and Elzinga (1977). Rates of synonymous and nonsynonymous nucleotide substitutions were computed with the LWL program (Li et al. 1985).

The maximum parsimony algorithm was used to generate phylogenetic trees using the DNAPARS and PROTPARS functions of PHYLIP (version 3.5; Felsenstein 1989). Phylogenetic analysis was also done by the neighbor-joining method (Saitou and Nei 1987) using the NEIGHBOR program of PHYLIP. Bootstrap analyses were performed by using the SEQBOOT, DNADIST, PROTDIST, NEIGHBOR, and CONSENSE functions of the PHYLIP package.

Results

Comparison of Coding Sequences of Echinoderm Actin Genes

Table 1 shows pairwise comparisons of the coding sequences and deduced amino acid sequences corresponding to the *L. pictus* actin genes LpC1 and LpC2; the *S. purpuratus* actin genes CyI, CyIIb, CyIIIa/b, and SpM; the *S. franciscanus* actin genes Sf15A and Sf15B; and the actin genes PoCy and PoM of the sea star *P. ochraceus.* The data in Table 1 confirm that cytoskeletal actins are extremely conserved within sea urchins, with more than 90% similarity at the nucleotide level and more than 96% at the amino acid level. The nucleotide sequence divergences indicate that the two *L. pictus* cytoskeletal actin genes LpC1 and LpC2 are more closely related to each other and to CyI than to other *S. purpuratus* actin genes. However, at the amino acid level, there are only five amino acid replacements between LpC1 and CyI actin but 13 amino acid replacements between LpC1 and LpC2 actins; there are 11–13 amino acid replacements between LpC2 and *S. purpuratus* cytoskeletal actins. *The S. purpuratus* muscle actin gene SpM and the two sea star actin genes PoCy and PoM are most divergent from sea urchin cytoskeletal actin genes. Table 1 indicates that the sea star *P. ochraceus* muscle actin PoM is more similar to cytoskeletal actins than is the *S. purpuratus* muscle actin SpM.

The 3′ Coding Region of Echinoderm Actin Genes

The cytoskeletal actin genes of *S. purpuratus* have three exons encoding amino acids which are interrupted by an

Table 1. Pairwise comparisons of cytoskeletal and muscle actin genes from the sea urchins *L. pictus, S. purpuratus, S. franciscanus,* and the sea star *P. ochraceus*[a]

	S.p				*S.f*		*L.p*			Muscle	
	CyI	CyIIb	CyIIIa	CyIIIb	Sf15a	Sf15b	LpC1	LpC2	*P.o* PoCy	SpM	PoM
CyI		1.62/6.5	7.78/35.9	6.86/30.1	3.19/12.5	2.92/13.0	3.66/15.5	6.06/24.2	12.74/66.7	12.90/52.6	13.91/68.1
CyIIb	1		7.78/37.2	7.25/31.4	2.82/12.3	2.54/12.4	3.84/15.7	6.34/24.9	13.13/68.2	12.86/49.8	14.63/71.4
CyIIIa	9	8		3.10/10.7	6.97/31.0	7.37/34.5	8.48/39.1	9.50/44.6	16.31/93.4	15.65/67.7	15.38/76.6
CyIIIb	10	11	7		6.36/25.3	6.84/29.0	7.44/32.2	8.25/37.5	15.66/86.5	15.25/63.9	14.95/70.8
Sf15a	3	2	10	12		1.71/6.8	3.94/15.2	6.44/24.6	13.46/67.9	13.93/55.2	14.28/64.5
Sf15b	1	0	9	11	2		4.41/18.7	7.03/28.9	13.67/70.8	13.72/56.8	14.29/66.2
LpC1	5	6	13	13	7	6		4.60/16.2	12.61/62.9	13.72/53.4	14.08/65.3
LpC2	12	13	12	11	14	13	13		14.88/72.5	15.61/64.1	15.97/76.5
PoCy	7	8	13	13	11	8	10	17		16.39/75.8	11.37/52.7
SpM	26	27	30	29	29	27	31	33	24		12.78/56.7
PoM	14	15	17	17	17	15	19	21	13	16	

[a] Percentage DNA sequence divergence using the Kimura two-parameter correction (above) and synonymous site substitution rates (below) are given above the diagonal. Below the diagonal are the number of amino acid replacements. All amino acid sequences were translated from corresponding nucleotide sequences. There are three uncertain amino acids in each of the sequences of CyI, CyIIIa, and SpM; since these amino acids are very conserved in all known actin amino acid sequences of sea urchins and sea stars, we did not include these sites as replacements

intron between codons 121 and 122, and an intron within codon 204 (Schuler et al. 1983; Durica et al. 1988; Akhurst et al. 1987). The third exon includes codons corresponding to amino acids 204–374. This region is interrupted by an intron at codon 267 for the known muscle actin genes of echinoderms (Crain et al. 1987; Kowbel and Smith 1989). The intron positions for most of the *L. pictus* actin genes have not been established. For simplicity, we refer to DNA sequences encoding amino acid residues 204–374 as "the 3′ coding region." For the known sea urchin actin genes, this region has the highest frequency of nonsynonymous substitutions. Table 2 compares the sequences of the 3′ coding region of all five *L. pictus* actin genes with other echinoderm actin genes. Above the diagonal are corrected pairwise nucleotide sequence divergences. Below the diagonal are rates of nucleotide substitutions for synonymous and nonsynonymous sites. The rates of nucleotide substitution at nonsynonymous sites of sea urchin cytoskeletal actin genes are very low, ranging from 0.5% (CyI–CyIIb) to 4.6% (CyIIIb–LpC4).

The 3′ Noncoding Regions of Echinoid Actin Genes

The 3′ noncoding sequences of the LpC1 and LpC2 actin genes were compared to the *S. purpuratus* CyI and CyIIa/b actin genes. Because of the incomplete sequence data available for the *S. purpuratus* actin genes, only 138 nt immediately following the termination codon were aligned in Fig. 1. Pairwise comparisons show that LpC1 and CyI are most similar (81.9%) while LpC2 and CyI are also quite similar (77.5%). LpC1 and LpC2 are 62.4% and 58.2% similar to CyIIa, respectively. The first 40–50 nt sequence proximal to the stop codon is more conserved than the remaining 3′ noncoding sequence among the five cytoskeletal actin genes aligned in Fig. 1. Comparisons of the 3′ noncoding regions of the LpC1 or LpC2 genes to the CyIIb, CyIIIa, and CyIIIb actin genes show an average similarity of 53.4%, of marginal significance (data not shown). The entire 3′ noncoding sequences of LpC1 and LpC2 were also compared (data not shown). The two sequences are 89% similar when ten gaps are introduced for LpC1 and 11 for LpC2; the alignment gaps range from 1 to 9 bp. The alignment of the 3′ noncoding regions of the muscle genes SpM and LpM is shown in Fig. 2 and showed a 73.6% similarity. Both sequences are incomplete; 280 bp of sequence were compared.

Amino Acid Replacements Deduced from the 3′ Coding Regions

The amino acid sequences of all the actins compared were deduced from the nucleotide sequences based on the universal code. Table 3 shows amino acid replacements in the 3′ coding region for echinoderm actins compared to the *S. purpuratus* CyI actin. Several informative features were noted: (1) All four cytoskeletal actins of *L. pictus* (LpC1–LpC4) and two *S. purpuratus* cytoskeletal actins (CyIIIa and CyIIIb) have serine at position 264. (2) The muscle actins of the two sea urchins and the sea star share the same distinctive amino acids at positions 259 (threonine), 266 (isoleucine), 277 (threonine), 286 (isoleucine), 302 (threonine), 304 (serine), and 305 (tyrosine). Among these amino acid replacements, those at positions 259 and 266 have been suggested as diagnostic for distinguishing muscle actins from cytoskeletal actins in deuterostomes (Collins and Elzinga 1975; Kovilur et al. 1993). Four additional distinctive amino acids were shared by muscle actins of the two sea urchins *S. purpuratus* and *L. pictus:* alanine at positions 228 and 232 and serine at positions 322 and 323. (3) Like the muscle

Table 2. Pariwise comparisons of the third exon of actin genes of the sea urchins *L. pictus, S. purpuratus, S. franciscanus,* and the sea star *P. ochraceus*[a]

	S.p				*S.f*	
	CyI	CyIIb	CyIIIa	CyIIIb	Sf15a	Sf15b
CyI		1.67	8.13	7.42	4.92	3.82
CyIIb	6.2/0.5		7.90	5.59	4.04	2.96
CyIIIa	33.7/2.7	34.3/2.1		5.59	7.19	6.73
CyIIIb	29.6/2.4	32.0/2.9	18.1/2.7		7.40	7.16
sf15a	19.0/1.3	17.5/0.8	27.1/2.7	24.1/3.4		2.30
sf15b	17.9/0.5	15.1/0.0	27.0/2.2	25.2/2.9	8.3/0.8	
LpC1	20.6/0.3	21.1/0.8	39.6/2.4	32.6/2.1	21.1/1.3	24.9/0.8
LpC2	28.3/1.8	29.3/2.4	52.8/3.0	45.6/1.8	32.7/2.9	34.9/2.4
LpC3	19.0/1.3	18.6/1.8	43.1/3.5	36.7/3.2	22.6/2.4	23.8/2.1
LpC4	25.4/3.0	24.6/2.4	47.0/3.6	39.0/4.6	31.4/3.0	32.7/2.4
PoCy	58.6/3.7	56.9/4.1	104.6/5.9	90.0/5.5	71.7/4.7	73.7/4.1
SpM	61.1/7.5	53.6/8.3	73.7/9.7	74.9/9.2	69.2/9.2	68.1/8.3
LpM	72.0/5.9	66.0/6.8	93.9/8.1	105.6/7.6	77.8/7.6	75.0/6.8
PoM	85.9/6.2	89.0/6.9	83.7/7.3	83.0/7.2	80.8/7.7	82.0/6.9

[a] Percentage DNA sequence divergence using the Kimura two-parameter correction are given above the diagonal. Below the diagonal are fractions (in percentage) of synonymous (above) and nonsynonymous (below) substitutions

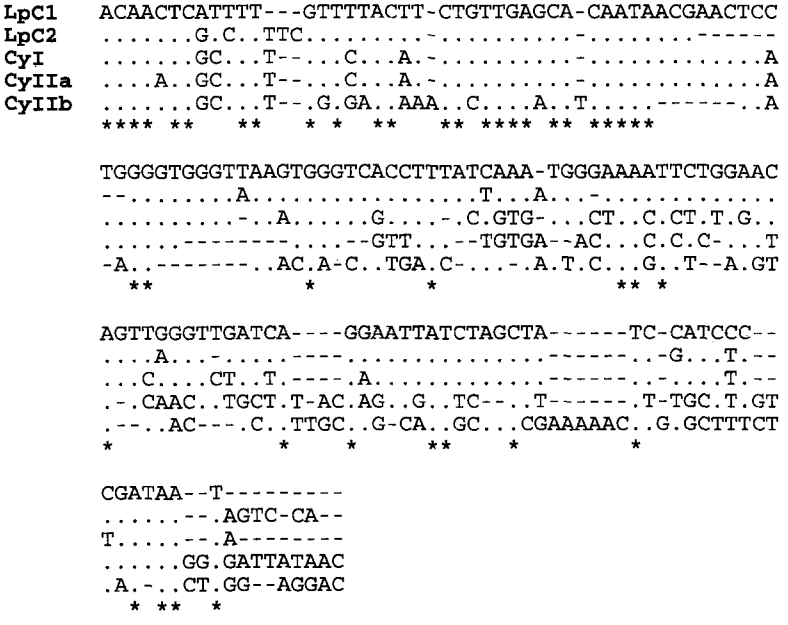

Fig. 1. Comparison of the 3′ noncoding regions of cytoskeletal actin genes. The *L. pictus* actin genes LpC1 and LpC2 were compared to the *S. purpuratus* cytoskeletal actin genes CyI, CyIIa, and CyIIb; 138 nt immediately following the termination codon of each actin genes were aligned. *Asterisks* (*) indicate nucleotides conserved in all five genes. *Dashes* indicate gaps introduced to optimize alignment.

actins, the cytoskeletal actins CyIIIa, CyIIIb, and LpC2 have two musclelike "diagnostic" amino acids: threonine at the position 259 and isoleucine at codon 266. Similar to the sequences of muscle actins, the LpC4 and PoCy actins have a threonine at position 277 and CyIIa has a tyrosine at position 305. (4) A replacement of cysteine with alanine at position 256 is found in the three *S. purpuratus* actins CyIIa, CyIIb, and CyIIIa, the two *S. franciscanus* actins Sf15a and Sf15b, and the *L. pictus* actin LpC4.

Among the four cytoplasmic actins of *L. pictus,* the amino acid sequence encoded by the 3′ coding region of the LpC1 actin gene is the most similar to *S. purpuratus* CyI actin, having only one amino acid replacement at position 264 (serine in place of alanine). All of the *L. pictus* cytoskeletal actins, as well as the *S. purpuratus* actins CyIIIa and CyIIIb, share this serine.

Among 12 amino acid differences between the actins LpC2 and CyI (Table 1), eight are encoded by the 3′ coding region. Five conservative replacements were observed at positions 212 (valine for isoleucine), 261 (leucine for phenylalanine), 266 (isoleucine for leucine), 317 (serine for threonine), and 314 (arginine for lysine). Three additional replacements involved changes between polar and nonpolar amino acids: leucine for glutamine at position 228, threonine for alanine at position 259, and serine for alanine at position 264. A striking observation is that four identical amino acid replacements were also

Table 2. Extended

L.p					Muscle		
LpC1	LpC2	LpC3	LpC4	*P.o PoCy*	SpM	LpM	PoM
4.04	6.71	4.70	7.15	12.30	15.68	16.10	17.11
4.68	7.38	5.13	6.69	12.76	15.89	16.30	18.39
8.59	10.97	10.26	10.96	17.75	19.45	20.42	18.27
7.65	9.53	9.30	10.90	17.37	19.64	21.20	18.65
5.13	8.31	6.26	8.30	15.01	18.51	18.11	18.34
5.35	8.30	6.25	8.07	14.76	17.46	17.09	17.83
	5.13	2.52	5.78	14.74	19.64	18.13	18.63
20.4/1.6		6.03	8.97	12.78	17.21	17.63	16.53
8.3/1.0	20.5/2.6		7.14	15.03	18.79	19.47	18.64
18.6/2.7	30.1/4.0	21.0/3.8		12.76	18.00	18.42	17.04
66.2/3.4	66.4/5.1	56.6/4.4	71.9/5.0		17.97	18.36	14.32
69.4/8.0	89.1/8.3	63.6/9.2	69.1/9.5	61.1/9.8		8.22	16.21
81.1/6.5	81.2/6.6	74.6/7.9	86.7/7.8	71.4/8.5	37.5/1.9		16.84
76.7/76.7	96.8/7.0	68.2/7.7	82.0/8.1	61.1/5.8	69.2/6.4	74.9/6.5	

Fig. 2. Alignment of the 3′ noncoding termini of muscle actin genes LpM and SpM; 280 nt of LpM and 277 nt of SpM immediately distal to the stop codon were compared.

were also found in the sequences of CyIIIa actin at positions 259, 264, 266, and 317 and of CyIIIb actin at positions 259, 261, 264, and 266.

Compared to CyI actin, the *L. pictus* cytoskeletal actin LpC4 has eight replacements of amino acids encoded by the 3′ coding region. There are four changes of alanine to serine at positions 259, 264, 271, and 318. The replacement of alanine for serine at position 264 is a common feature for all cytoskeletal actins of *L. pictus* and for the CyIIIa and CyIIIb cytoplasmic actins of *S. purpuratus*. The serines at positions 272 and 319, and the replacements of threonine for isoleucine at position 249 and tyrosine for phenylalanine at position 261, are unique for the LpC4 actin among echinoderm actins characterized.

The 3′ coding region of the LpC3 gene encodes five amino acid differences compared to the CyI gene. Three replacements are unique to this echinoid actin: phenylalanine for aspartic acid at position 291, phenylalanine for leucine at position 319, and valine for isoleucine at position 326. The replacement of an aspartic acid with asparagine at position 211 also occurs in the *L. pictus* muscle actin.

Compared with the CyI actin, the muscle actin LpM has 13 amino acid replacements, SpM has 16, and PoM has only 10. Muscle actin genes of echinoderms are thought to have evolved by duplication from a cytoskeletallike ancestral actin gene (Kovilur et al. 1993). The muscle actin PoM of the sea star has retained some amino acid sequence features of cytoskeletal actins, while the *S. purpuratus* muscle actin SpM is the most divergent from the cytoskeletal actins.

Molecular Phylogenetic Analysis

A phylogenetic tree of the echinoderm actin was predicted by application of the maximum parsimony algo-

Table 3. Comparison of deduced amino acids 204–374 actins[a]

Actin	Amino acid position number																			
	211	212	219	222	225	227	228	230	231	232	249	250	256	257	259	261	264	266	271	277
CyI	D	I	V	D	Q	M	Q	A	A	S	I	G	C	P	A	F	A	L	A	C
CyIIa	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.
CyIIb	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.
CyIIIa	.	.	.	.	.	.	.	.	.	.	.	.	A	S	T	.	S	I	.	.
CyIIIb	.	.	.	.	E	.	.	.	.	.	.	.	.	S	T	L	S	I	.	.
Sf15a	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	P	.	.	.
Sf15b	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.
LpC1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	S	.	.	.
LpC2	.	V	.	.	.	.	L	.	.	.	.	.	.	.	T	L	S	I	.	.
LpC3	N	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	S	.	.	.
LpC4	.	.	.	.	.	.	.	.	.	.	T	.	A	.	S	Y	S	.	S	T
PoCy	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	S	.	.	T
LpM	N	.	.	.	.	.	A	P	.	A	.	.	.	.	T	.	.	I	.	T
SpM	.	.	T	N	R	.	A	.	.	A	.	.	.	.	T	.	.	I	.	T
PoM	.	.	T	.	.	.	.	.	S	.	.	.	.	.	T	.	.	I	.	T

[a] Standard abbreviations of amino acids are used. Dots represent amino acids which are identical to those of CyI actin. Since the sequence of CyIIa is incomplete, amino acids corresponding to positions 306–374 were not compared and are shown as blanks in the table. Shaded boxes indicate shared features discussed in Results

rithm to the deduced amino acid sequences of the 3′ coding regions and is shown in Fig. 3a. A muscle actin of the ascidian *S. clava* (ScM) was used as an outgroup. The formation of a clade including LpC2 and CyIIIa,b actins separated from other sea urchin cytoskeletal actins was supported by 96% of bootstrap replicates. The two CyIII actins are encoded by genes which are closely linked and share a very similar 3′ noncoding terminus in *S. purpuratus* (Lee et al. 1984; Akhurst et al. 1987); this terminus is not shared with any *L. pictus* actin gene. The LpC1 and LpC3 actins were on a branch supporting the *S. purpuratus* CyI and CyIIb actins and *S. franciscanus* Sf15a and Sf15b actins.

Figure 3b shows the maximum parsimony trees generated by comparison of the entire deduced amino acid sequences of actins; the lack of complete sequences for the *L. pictus* LpC3, LpC4, and LpM genes precluded their inclusion. Again, LpC2 actin was grouped with the CyIII actins, and LpC1 actin was on a different branch including the CyI and CyIIb actins of *S. purpuratus* and the 15a and 15b actins of *S. franciscanus*. The separation of these two branches was supported by 73% of bootstraps. The application of the neighbor-joining distance-matrix algorithm (Saitou and Nei 1987) to the amino acid sequences deduced from the 3′ coding region or the entire actin gene supported the same formation of two branches of echinoid cytoskeletal actins.

In the two trees shown in Fig. 3, the muscle actins of echinoderms formed a clade separated from all the cytoskeletal actins (supported by 100% of bootstraps). This indicates that the ancestral muscle actin gene existed in the common ancestor of these echinoderms, and the muscle actin genes LpM and SpM are orthologous.

Discussion

Based on sequence comparisons and patterns of expression in embryos (our unpublished data), the *L. pictus* actin genes LpC1, LpC2, LpC3, and LpC4 are cytoskeletal, while LpM is a muscle actin gene; there is no evidence that any of them are closely linked (Johnson et al. 1983; our unpublished observations). The five expressed cytoskeletal actin genes of *S. purpuratus* have been categorized into three subfamilies (CyI, CyII, and CyIII) primarily based on similarities of the 3′ noncoding sequences and linkage (Lee et al. 1984). CyI, CyIIa, and CyIIb actin genes are linked and share sequence similarities in their first intron; the 3′ noncoding sequences of the CyII genes are very similar to one another and somewhat similar to CyI (Scheller et al. 1981; Durica et al. 1988). The CyIIIa and CyIIIb actin genes of *S. purpuratus* are linked and have similar 3′ noncoding sequences and similarities in deduced amino acids which distinguish them from other *S. purpuratus* actin genes (Lee et al. 1984; Akhurst et al. 1987; Flytzanis et al. 1989).

The cytoskeletal actin genes of *L. pictus* do not easily fit such a categorization. For instance, the 3′ coding regions of the genes LpC1 and LpC3 are more similar to each other than to other *L. pictus* actin genes, but their 3′ noncoding regions are quite different. On the other hand, the 3′ noncoding regions of LpC1 and LpC2 are highly (89%) similar, but the two genes encode actins differing by 13 amino acid replacements (Table 1). The 3′ noncoding sequences of the LpC1 and LpC2 actin genes are also very similar to the 3′ noncoding sequence of the *S. purpuratus* CyI gene; this is the only 3′ terminal sequence of an *S. purpuratus* cytoskeletal actin gene which

Table 3. Extended

Amino acid position number																				
280	283	286	291	294	296	301	302	303	304	305	314	316	317	318	319	322	323	324	326	366
S	K	V	D	A	T	G	S	T	M	F	K	I	T	A	L	P	T	M	I	P
.	.	.	.	.	S	.	.	.	.	Y										
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
R	.	.	.	V	.	.	.	.	.	.	.	.	S	.	.	.	.	.	.	.
.	.	.	.	.	.	A	.	.	.	.	.	.	V	.	.	.	.	.	.	.
.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	R	.	S	.	.	.	.	.	.	.
.	.	.	F	.	.	.	.	.	.	.	.	.	.		F	.	.	.	V	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	S	.	.	.	.	.	
.	.	.	.	.	.	.	.	.	.	.	.	V	.	.	.	.	.	.	.	.
.	.	I	.	.	.	.	T	S	.	Y	.	.	.	.	.	S	S	.	.	.
.	.	I	.	.	.	.	T	S	.	Y	.	.	.	.	.	S	S	V	.	A
.	.	I	.	.	.	.	T	S	.	Y	.	.	Q	.	.	.	.	.	.	.

cross-hybridizes with *L. pictus* DNA (Lee et al. 1984). We have confirmed that this is the only 3′ noncoding sequence of *S. purpuratus* actin genes extensively shared with *L. pictus* actin genes (our unpublished data). Consistent with Southern blots of Lee et al. (1984) we have identified two *L. pictus* actin genes which have a 3′ noncoding sequence similar to that of the CyI actin gene.

Sequence Bias and Codon Usage in Echinoderm Genes

A strong bias for the use of particular synonymous codons have been noted for *Strongylocentrotus* actin genes by Foran et al. (1985) and for sea star actin genes by Kovesdi and Smith (1989). We have observed similar bias in codon usage for *L. pictus* actin genes: 46% of the bases at the third positions of each codon are C. Britten (1993) compared the sequences of actin and other genes from diverse organisms and concluded that highly conserved genes having extreme G + C base compositions at the third sites of codons are constrained in the extent of synonymous base substitutions at these sites. Bias in codon usage is presumably responsible for the maintenance of unusual G + C contents of third sites, at least in part. Britten also found that some synonymous substitutions are never observed, apparently forbidden by unknown selective mechanisms. Codon usage for actin genes of *L. pictus* and *S. purpuratus* is similar (our unpublished observations).

Evolution of Cytoskeletal Actin Genes in Sea Urchins

Data presented in Tables 1 and 2 indicate low rates of silent site substitutions between most of the cytoskeletal actin genes of these sea urchins. However, the silent site substitutions may be approaching saturation for some actin genes due to the constraints mentioned above; this would limit the utility of comparing nucleotide coding sequences in assessing the phyologenetic relationships of these genes.

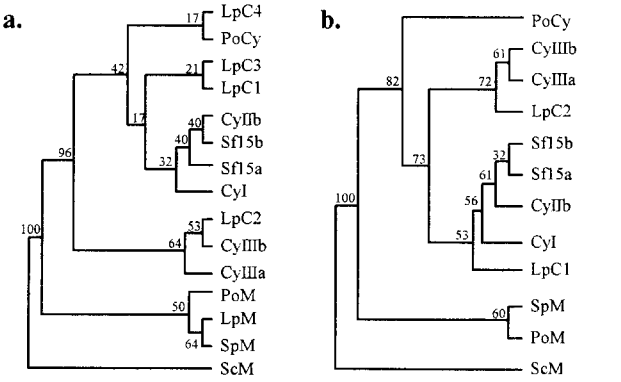

Fig. 3. Molecular phylogenetic analysis of the actins of echinoderms. Trees shown were constructed by applying the maximum parsimony algorithm to **(a)** the deduced amino acid sequence of the 3′ coding region and **(b)** the deduced amino acid sequence of the entire coding region. In all trees, an ascidian muscle actin gene (ScM) was included as an outgroup. Numbers shown are percentages of 100 bootstrap replicates in which the same internal branch was recovered. See Materials and Methods for terminology.

Neighbor-joining and maximum-parsimony algorithms were used to analyze the phylogeny of the coding sequences of echinoid actin genes. All the *L. pictus* cytoskeletal actin genes were included in a clade separated from the *Strongylocentrotus* actin genes. The simplest interpretation of these trees is that all extant cytoskeletal actin genes of *L. pictus* arose from a single gene present in the common ancestor of *L. pictus* and *S. purpuratus*.

On the other hand, the similarity observed in the coding sequences of the cytoskeletal actin genes within each sea urchin species might be a consequence of intraspecific homogenization of nucleotide sequences. Studies on globin gene families suggest that concerted evolution

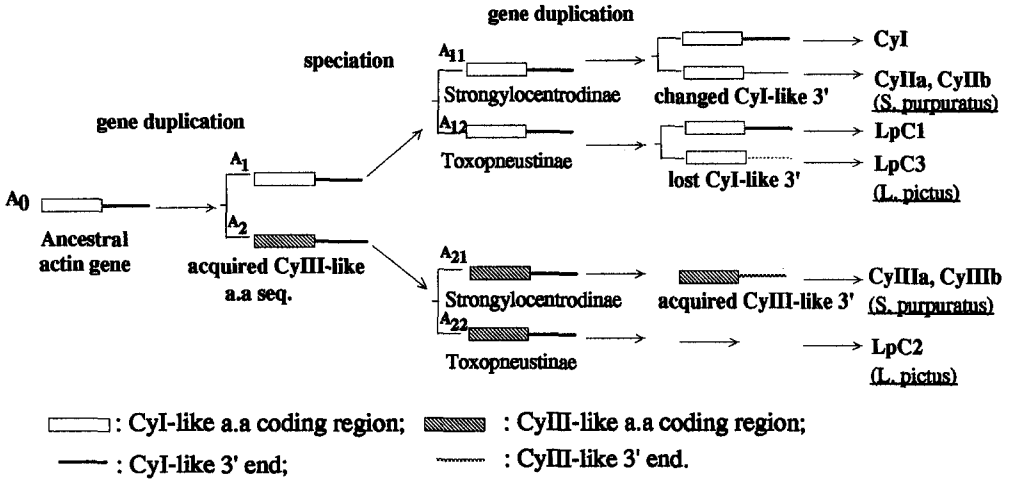

Fig. 4. Model for the evolution of the actin gene families of the sea urchins *L. pictus* and *S. purpuratus*. A pathway is presented for the generation of extant CyI-like and CyIII-like genes in contemporary sea urchins from that of a putative ancestral actin gene (A_0). This model accounts for the presence of a single CyI-like 3′ noncoding terminus among *S. purpuratus* actin genes, while two such termini are found among *L. pictus* actin genes. Variations are possible on this model having two ancestral actin genes in the common ancestral species, but we believe this version to be most parsimonious.

events may partly or completely obscure the evolutionary history of divergence between duplicated genes (Li and Graur 1991, pp. 168–169). Constraints on allowable nucleotide substitutions within a species and/or gene conversion events can obscure phylogenetic relationships. Crain et al. (1987) have provided evidence for a gene conversion event between a muscle and cytoskeletal actin gene in *S. purpuratus*.

The trees based on the deduced amino acid sequence of the 3′ coding region (Fig. 3a) or the entire coding region (Fig. 3b) present a quite different picture of actin gene evolution which we believe is more compelling. In these trees, the grouping of LpC1 with CyI-CyII and LpC2 with CyIIIa,b was strongly supported in bootstrap replicates. Comparison of amino acid sequences may be most informative if one focuses on rare features which are shared. As described in Results, comparison of the 3′ coding regions supports a close relationship of LpC1 and CyI actins, which differ by a single nonconservative amino acid replacement at position 264 (one which is common to all *L. pictus* cytoskeletal actins). The LpC2 and CyIIIb actins share an unusual complex of amino acid replacements at positions 259, 261, 264, and 266. Thus LpC2 actin indeed appears to be closely related to CyIIIb, and, to a lesser extent, CyIIIa actins. The LpC2 actin gene has a pattern of expression in embryos similar to that of the CyIIIa and CyIII genes, while the LpC1 gene has a pattern of expression most similar to that of the CyIIa gene, as well as the CyI and CyIIb genes (our unpublished data). These similarities in unusual actin gene sequences and spatial patterns of expression consistent with LpC2 actin being related to the CyIII subfamily and LpC1 being related to the CyI/CyII subfamily of actins.

In Fig. 4 we propose a model, based on similarities of the deduced amino acid sequences and 3′ noncoding sequences, which can account for the origin of the extant cytoskeletal actin genes of *L. pictus* and *S. purpuratus*. We propose that the ancestral cytoskeletal actin gene had a CyI-like third exon and 3′ noncoding sequences, since these have been conservatively retained in several echinoid species (Lee et al. 1984); we refer to this gene as A_0. A_0 duplicated and one copy (A_1) retained CyI-like coding sequences while the other copy (A_2) diverged to become CyIII-like in coding sequence. The conserved complex of amino acid replacements encoded by the 3′ coding region of the CyIII genes may be the result of selection for regained functionality after initial mutational events resulted in impaired function and/or the result of selection for specialized function. In this model, there were at least two ancestral cytoskeletal actin genes in the common ancestor of *L. pictus* and *S. purpuratus*. The origin of the LpC4 gene is not accounted for by our model. It may have arisen by duplication after the divergence of the species or may be derived from a primitive cytoskeletal actin gene now absent in *S. purpuratus*. If there was a single ancestral cytoskeletal actin gene at the time of species divergence, convergent evolution probably accounts for the similarities in the amino acid sequences encoded by the LpC2 and CyIIIa,b actin genes and their spatial pattern of expression.

Constraints on the 3′ Noncoding Sequences of Actin Genes

The 3′ noncoding sequences of the LpC1 and LpC2 actin genes are very similar. The many small gaps required for alignment indicate that these sequences did not arise recently; thus the 3′ noncoding sequences are likely to be selectively constrained. This possibility is supported by the observation that this conserved 3′ noncoding se-

quence is retained by one or more actin genes of all species of sea urchins examined to date (Lee et al. 1984). Slipped-strand mispairing (Levinson and Gutman 1987; Li and Graur 1991) has been invoked to explain the insertion/deletion events which occur in DNA regions containing contiguous short repeats, such as the intron and 3′ noncoding regions of the linked actin genes of *S. franciscanus* 15A and 15B (Foran et al. 1985), as well as *S. purpuratus* actin genes CyI and CyIIa (Schuler et al. 1983) and CyIIIa and CyIIIb (Durica et al. 1988). The same mechanism may explain the insertion/deletion events in the A-T-rich 3′ noncoding regions of LpC1 and LpC2 genes. Vertebrate actin genes also have highly conserved, isoform-specific 3′ noncoding sequence elements (Ng et al. 1985; Yaffe et al. 1985; Erba et al. 1986, 1988). Several functions have been proposed for the highly conserved noncoding regions sometimes observed among members of multigene families, including roles in localization and activity of the mRNAs (Lloyd and Gunning 1993). A conserved 40-base-pair sequence in the 3′ noncoding terminus of vertebrate β-actin genes is involved in transcriptional regulation (DePonti-Zilli et al. 1988).

While the function of the conserved 3′ noncoding sequence of the LpC1 and LpC2 actin genes is unknown, it is unlikely to be related to the spatial patterns of expression of these genes in the embryos, since these are quite distinct for the two genes. Using sequence-specific hybridization probes we have found that all five *L. pictus* actin genes are expressed in embryos, each having a unique pattern of expression (our unpublished data). The LpC1 gene is expressed in several different spatial territories of the embryo, showing combined features of the CyI and CyIIa,b gene expressions. The LpC2 gene is predominantly expressed in aboral ectoderm cells, a pattern similar to that of the CyIIIa,b genes of *S. purpuratus*.

Acknowledgments. We thank Dr. M.J. Smith for help with the data analysis. M.J. Smith, R. Davenport, and B. Crawford provided useful comments on the manuscript. This research was supported in part by grants from the Natural Sciences and Engineering Research Council of Canada and the B.C. Health Research Foundation.

References

Akhurst RJ, Calzone FJ, Lee JJ, Britten RJ, Davidson EH (1987) Structure and organization of the CyIII actin gene subfamily of the sea urchin, *Strongylocentrotus purpuratus*. J Mol Biol 194:193–203

Beach RL, Jeffery WR (1992) Multiple actin genes encoding the same α-muscle isoform are expressed during ascidian development. Dev Biol 151:55–66

Benson D, Lipman DJ, Ostell J (1993) Genbank. Nucleic Acids Res 21:2963–2965

Brandhorst BP, Klein WH (1992) Territorial specification and control of gene expression in the sea urchin embryo. Semin Dev Biol 3:175–186

Britten RJ (1993) Forbidden synonymous substitutions in coding regions. Mol Biol Evol 10:205–220

Bullock BP, Nisson PE, Crain WR (1988) The timing of expression of four actin genes and an RNA polymerase III-transcribed repeated sequence is correct in hybrid embryos of the sea urchin species *Strongylocentrotus purpuratus* and *Lytechinus pictus*. Dev Biol 130:335–347

Cabot EL, Beckenbach AT (1989) Simultaneous editing of multiple nucleic acid and protein sequences with ESEE. Comput Appl Biosci 5:233–234

Collins JH, Elzinga M (1975) The primary structure of actin from rabbit skeletal muscle. J Biol Chem 250:5915–5920

Collura R, Katula KS (1992) Spatial pattern of expression of CyI actin-β-galactosidase fusion genes injected into sea urchin eggs. Dev Growth Differ 34:635–647

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

Cox KH, Angerer LM, Lee JJ, Davidson EH, Angerer RC (1986) Cell lineage-specific programs of expression of multiple actin genes during sea urchin embryogenesis. Mol Biol 188:159–172

Crain WR, Boshar MF, Cooper AD, Durica DS, Nagy A, Steffen D (1987) The sequence of a sea urchin muscle actin gene suggests a gene conversion with a cytoskeletal actin gene. J Mol Evol 25:37–45

Crain WR, Durica DS, Doren KV (1981) Actin gene expression in developing sea urchin embryos. Mol Cell Biol 1:711–720

Davidson EH, Thomas TL, Scheller RH, Britten RJ (1982) The sea urchin actin genes, and a speculation on the evolutionary significance of small gene families. In: Dover GA, Flavell RB (eds) Genome evolution. Systematics Association, Academic Press, New York, pp 177–191

DePonti-Zilli L, Seiler-Yuyns A, Paterson BM (1988) A 40-base-pair sequence in the 3′ end of the β-actin gene regulates β-actin mRNA transcription during myogenesis. Proc Natl Acad Sci USA 85: 1389–1393

Durica DS, Garza D, Restrepo MA, Hryniewicz MM (1988) DNA sequence analysis and structural relationships among the cytoskeletal actin genes of the sea urchin *Strongylocentrotus purpuratus*. J Mol Evol 28:72–86

Durica DS, Schloss JA, Crain WR (1980) Organization of actin gene sequences in the sea urchin: molecular cloning of an intron-containing DNA sequence coding for a cytoplasmic actin. Proc Natl Acad Sci USA 77:5683–5687

Erba HP, Eddy R, Shows T, Kedes L, Gunning P (1988) Structure, chromosome location, and expression of the human g-actin gene: differential evolution, location, and the expression of the cytoskeletal β-and γ-actin genes. Mol Cell Biol 8:1775–1789

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

Felsenstein J (1989) PHYLIP—phylogeny inference package (version 3.2). Cladistics 5:164–166

Flytzanis CN, Bogosian EA, Niemeyer CC (1989) Expression and structure of the CyIIIb actin gene of the sea urchin *Strongylocentrotus purpuratus*. Mol Reprod Dev 1:208–218

Foran DR, Johnson PJ, Moore GP (1985) Evolution of two actin genes in the sea urchin. *Strongylocentrotus franciscanus*. J Mol Evol 22: 108–116

Franks RR, Hough-Evans BR, Britten RJ, Davidson EH (1988) Direct introduction of cloned DNA into the sea urchin zygote nucleus, and fate of injected DNA. Development 102:287–299

Garcia R, Paz-Aliaga B, Ernst SG, Crain WR (1984) Three sea urchin actin genes show different patterns of expression: muscle specific, embryo specific, and constitutive. Mol Cell Biol 4:840–845

Hightower RC, Meagher RB (1986) The molecular evolution of actin. Genetics 114:315–332

Hough-Evans BR, Franks RR, Cameron RA, Britten RJ, Davidson EH (1987) Correct cell-type-specific expression of a fusion gene injected into sea urchin eggs. Dev Biol 121:576–579

Johnson PJ, Foran DR, Moore GP (1983) Organization and evolution of the actin gene family in sea urchins. Mol Cell Biol 3:1824–1833

Kovesdi I, Smith MJ (1985) Actin gene number in the sea star *Pisaster ochraceus*. Can J Biochem Cell Biol 63:1145–1151

Kovilur S, Jacobson JW, Beach RL, Jeffery WR, Tomlinson CR (1993) Evolution of the chordate muscle actin gene. J Mol Evol 36:361–368

Kowbel DJ, Smith MJ (1989) The genomic nucleotide sequences of two differentially expressed actin-coding genes from the sea star *Pisaster ochraceus*. Gene 77:297–308

Lee JJ, Calzone FJ, Britten RJ, Angerer RC, Davidson EH (1986) Activation of sea urchin actin genes during embryogenesis: measurement of transcript accumulation from five different genes in *Strongylocentrotus purpuratus*. J Mol Biol 188:173–183

Lee JJ, Shott RJ, Rose SJ III, Thomas TL, Britten RJ, Davidson EH (1984) Sea urchin actin gene subtypes: gene number, linkage and evolution. J Mol Biol 172:149–176

Levison G, Gutman GA (1987) Slipped-strand mispairing: a major mechanism for DNA sequence evolution. Mol Biol Evol 4:203–221

Li W-H, Graur D (1991) Fundamentals of molecular evolution. Sinauer Associates, Sunderland, MA

Li W-H, Wu C-I, Luo C-C (1985) A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. Mol Biol Evol 2:150–174

Lloyd C, Gunning P (1993) Noncoding region of the γ-actin gene influence the impact of the gene on myoblast morphology. J Cell Biol 121:73–82

Lu RC, Elzinga M (1977) Partial amino acid sequence of brain actin and its homology with muscle actin. Biochemistry 16:5801–5806

Minor JE, Lee JJ, Akhurst RJ, Leahy PS, Britten RJ, Davidson EH (1987) Sea urchin actin gene linkages determined by genetic segregation. Dev Biol 122:291–295

Mounier N, Gouy M, Mouchiroud D, Prudhomme JC (1992) Insect muscle actins differ distinctly from invertebrate and vertebrate cytoplasmic actins. J Mol Evol 34:406–415

Ng S, Gunning P, Eddy R, Ponte P, Leavitt J, Shows T, Kedes L (1985) Evolution of the functional human β-actin gene and its multipseudogene family: conservation of non-coding regions and chromosomal dispersion of pseudogenes. Mol Cell Biol 5:2720–2732

Niemeyer CC, Flytzanis CN (1993) Upstream elements involved in the embryonic regulation of the sea urchin CyIIIb actin gene: temporal and spatial specific interactions at a single cis-actin element. Dev Biol 156:293–302

Nisson PE, Dike LE, Crain WR (1989) The *Strongylocentrotus purpuratus* actin genes show correct cell-specific expression in hybrid embryos of *S. purpuratus* and *Lytechinus pictus*. Development 105: 407–413

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Scheller RH, McAllister LB, Crain WR, Durica DS, Posakony JW, Thomas TL, Britten RJ, Davidson EH (1981) Organization and expression of multiple actin genes in the sea urchin. Mol Cell Biol 1:609–628

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

Shott RJ, Lee JJ, Britten RJ, Davidson EH (1984) Differential expression of the actin gene family of *Strongylocentrotus purpuratus*. Dev Biol 101:295–306

Smith AB (1988) Phylogenetic relationship, divergence times, and rates of molecular evolution for camarodont sea urchins. Mol Biol Evol 5:345–365

Vandekerckhove J, Weber K (1984) Chordate muscle actins differ distinctly from invertebrate muscle actins. J Mol Biol 152:413–426

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