J Mol Evol (1997) 44:289–298

JOURNAL OF MOLECULAR EVOLUTION

Evolution of Chordate Actin Genes: Evidence from Genomic Organization and Amino Acid Sequences

Takehiro Kusakabe,[1,2,3,*] Isato Araki,[3] Noriyuki Satoh,[3] William R. Jeffery[2]

[1] Division of Biological Sciences, Graduate School of Science, Hokkaido University, Sapporo 060, Japan
[2] Bodega Marine Laboratory, University of California, P.O. Box 247, Bodega Bay, CA 94923, USA
[3] Department of Zoology, Graduate School of Science, Kyoto University, Kyoto 606-01, Japan

Received: 20 June 1996 / Accepted: 16 October 1996

Abstract. The origin and evolutionary relationship of actin isoforms was investigated in chordates by isolating and characterizing two new ascidian cytoplasmic and muscle actin genes. The exon–intron organization and sequences of these genes were compared with those of other invertebrate and vertebrate actin genes. The gene *HrCA1* encodes a cytoplasmic (nonmuscle)-type actin, whereas the *MocuMA2* gene encodes an adult muscle-type actin. Our analysis of these genes showed that intron positions are conserved among the deuterostome actin genes. This suggests that actin gene families evolved from a single actin gene in the ancestral deuterostome. Sequence comparisons and molecular phylogenetic analyses also suggested a close relationship between the ascidian and vertebrate actin isoforms. It was also found that there are two distinct lineages of muscle actin isoforms in ascidians: the larval muscle and adult body-wall isoforms. The four muscle isoforms in vertebrates show a closer relationship to each other than to the ascidian muscle isoforms. Similarly, the two cytoplasmic isoforms in vertebrates show a closer relationship to each other than to the ascidian and echinoderm cytoplasmic isoforms. In contrast, the two types of ascidian muscle actin diverge from each other. The close relationship between the ascidian larval muscle actin and the vertebrate muscle isoforms was supported by both neighbor-joining and maximum parsimony analyses. These results suggest that the chordate ancestor had at least two muscle actin isoforms and that the vertebrate actin isoforms evolved after the separation of the vertebrates and urochordates.

Key words: Actin — Ascidians — Vertebrates — Multigene family — Muscle actin — Cytoplasmic actin — Chordates — Introns

Introduction

Most animals exhibit multiple actin isoforms which are encoded by a small gene family. For example, there are four muscle isoforms (α-skeletal, α-cardiac, α-vascular, and γ-enteric) and two nonmuscle isoforms (β- and γ-cytoplasmic) in mammals (Vandekerckhove and Weber 1979). The evolution of vertebrate actin genes has been discussed previously based on the amino acid sequences and gene structure (Vandekerckhove and Weber 1984; Alonso 1987; Miwa et al. 1991; Kovilur et al. 1993; Kusakabe 1995). To understand the origin and evolution of vertebrate actin genes, however, it is necessary to study actin genes in other chordates, including the urochordates (ascidians, salps, and larvaceans) and the cephalochordates (amphioxus).

In previous studies, we isolated ascidian larval-muscle actin genes and determined their genomic structure (Kusakabe et al. 1992, 1995, 1996). In the ascidian *Halocynthia roretzi*, at least two clusters of actin genes

*Present address: Division of Biological Sciences, Graduate School of Science, Hokkaido University, Sapporo 060, Japan; e-mail tgkusakabe@bio.hokudai.ac.jp
Correspondence to: T. Kusakabe

are expressed in the larval muscle cells. The *HrMA2/4* cluster contains at least five actin genes (Kusakabe et al. 1992) and the *HrMA1* cluster contains a pair of actin genes whose expression is regulated by a single bidirectional promoter (Kusakabe et al. 1995). The comparison of amino acid sequences among ascidian muscle actins revealed that they possess at least two distinct types of muscle actins, one expressed in larval muscle and the other in adult body-wall (Kusakabe 1995). However, only cDNA clones have been characterized for the ascidian body-wall muscle and cytoplasmic actins, and little is known about the genomic organization of the corresponding genes.

The origin and evolution of the chordates have been the subject of considerable discussion and speculation for more than a century (Haeckel 1868; Garstang 1928; Berrill 1955; Jefferies 1986; Wada and Satoh 1994; Holland and Garcia-Fernàndez 1996). The ascidian larva exhibits the hallmarks of a chordate, including a motile tail containing a notochord, dorsal nerve cord, and striated muscle cells. In contrast, the adult ascidian is a sessile organism with little resemblance to other chordates, except for the presence of pharyngeal gill slits (Satoh 1994). Since ascidians have distinct muscle tissues in both larval and adult phases, the evolutionary relationships between ascidian larval and adult actin isoforms and vertebrate muscle actins are of particular interest with respect to the chordate ancestor.

In this paper, we describe the exon–intron organization and nucleotide sequences of the muscle actin gene *MocuMA2* and cytoplasmic actin gene *HrCA1*. The sequence of these ascidian genes suggests that *MocuMA2* encodes an adult muscle actin and *HrCA1* a cytoplasmic actin. To infer the evolution of chordate actin genes, we compared the organization and sequences of these actin genes to those of other organisms and performed a phylogenetic analysis. The results suggest a monophyletic origin of the chordate muscle actin genes and suggest that vertebrate actin evolved after the separation of the vertebrate from the urochordate lineage.

Materials and Methods

Isolation and Characterization of Ascidian Actin Genes. A λFIXII clone containing the ascidian adult muscle-actin gene *MocuMA2* was isolated by screening a *Molgula oculata* genomic library (Kusakabe et al. 1996) with the ascidian muscle-actin probe HrcMA4 (Kusakabe et al. 1991). A λFIXII clone containing the ascidian cytoplasmic actin gene *HrCA1* was isolated by screening a *Halocynthia roretzi* genomic library (Kusakabe et al. 1995) with the ascidian cytoplasmic-actin cDNA clone *HrCA1* (Araki et al. 1996). The ^{32}P-labeled DNA probes were synthesized using the random primer labeling kit (United States Biochemical, Cleveland, OH) and [α-^{32}P]dCTP. Isolated genomic clones were digested with restriction enzymes and the digested fragments containing actin genes were subcloned into pBluescriptII SK(+). The subcloned genomic fragments were sequenced by the dideoxy chain termination procedure (Sanger et al. 1977) using Sequenase version 2.0 (USB) and [α-^{35}S]dATP. Oligonucleotide primers were synthesized on a Pharmacia LKB Gene Assembler Plus (Pharmacia Biosystems, Inc., Piscataway, NJ) or an Applied Biosystems DNA synthesizer (Applied Biosystems Japan, Tokyo, Japan). The sequencing reactions were loaded on 6% or 8% polyacrylamide gels. The *MocuMA2* and *HrCA1* sequences were deposited into DDBJ/EMBL/GenBank databases under the accession numbers D85743 and D45164, respectively.

Phylogenetic Analysis of Actin Sequences. The amino acid sequences of MocuMA2 and HrCA1 actins were aligned with those of six mammalian actin isoforms, seven ascidian actins, the starfish cytoplasmic and muscle actin isoforms, and *Arabidopsis thaliana* actin with the aid of the sequence editor SeqPup (D. Gilbert, Indiana University). GenBank/EMBL/DDBJ accession numbers for the actin sequences are: M20543, human α-skeletal muscle (Taylor et al. 1988); J00073, human α-cardiac muscle (Hamada et al. 1982); X13839, human α-aortic smooth muscle (Kamada and Kakunaga 1989); X16940, human γ-enteric smooth muscle (Miwa and Kamada 1990); M10277, human β-cytoplasmic (Nakajima-Iijima et al. 1985); M19283, human γ-cytoplasmic (Erba et al. 1988); X61042, *Styela plicata* adult muscle SpMA1 (Kovilur et al. 1993); L21915, *Molgula citrina* adult muscle McMA1 (Swalla et al. 1994); D10887, *Halocynthia roretzi* larval muscle HrMA4 (Kusakabe et al. 1992); D29014, *H. roretzi* larval muscle HrMA1 (Kusakabe et al. 1995); D78190, *Molgula oculata* larval muscle MocuMA1 (Kusakabe et al. 1996); X61040, *Styela clava* larval muscle ScTb1 (Beach and Jeffery 1992); X61041, *S. plicata* SpCA8 (Kovilur et al. 1993); M26500, *Pisaster ochraceus* muscle (Kowbel and Smith 1989); M26501, *P. ochraceus* cytoplasmic (Kowbel and Smith 1989); M20016, *Arabidopsis thaliana* AAc1 (Nairn et al. 1988). Phylogenetic trees were constructed with the aligned sequences by the maximum parsimony and neighbor-joining (Saitou and Nei 1987) algorithms in the PROTPARS program of PHYLIP (version 3.572; Felsenstein 1989) and the ClustalW program (Thompson et al. 1994), respectively. For the neighbor-joining analysis, evolutionary distances were estimated using Kimura's empirical method for protein distances (Kimura 1983). One thousand bootstrap replicates were performed for each phylogenetic analysis (Felsenstein 1985).

Results

Structure of Ascidian Adult Muscle and Cytoplasmic Actin Genes

We have isolated ascidian muscle and cytoplasmic actin genes to infer the evolutionary relationship of chordate actin genes. The *MocuMA2* and *HrCA1* genes were isolated from *M. oculata* and *H. roretzi* genomic libraries, respectively. The nucleotide and deduced amino acid sequences of the *MocuMA2* and *HrCA1* genes are shown in Fig. 1. The coding region of *MocuMA2* contains four introns (Fig. 1A), whereas the coding region of *HrCA1* is interrupted by six introns (Fig. 1B). Comparison of the genomic sequence of the *HrCA1* gene with the *HrCA1* cDNA clone (Araki et al. 1996) suggests that there are no introns in the 5′- and 3′-noncoding regions of the gene. Because a cDNA clone of *MocuMA2* is unavailable, we did not determine whether introns are present in the untranslated regions of the gene. Amino acid identities at diagnostic positions indicate that the *MocuMA2* and *HrCA1* genes encode muscle and cytoplasmic actins, respectively (Table 1). The predicted amino acid sequence

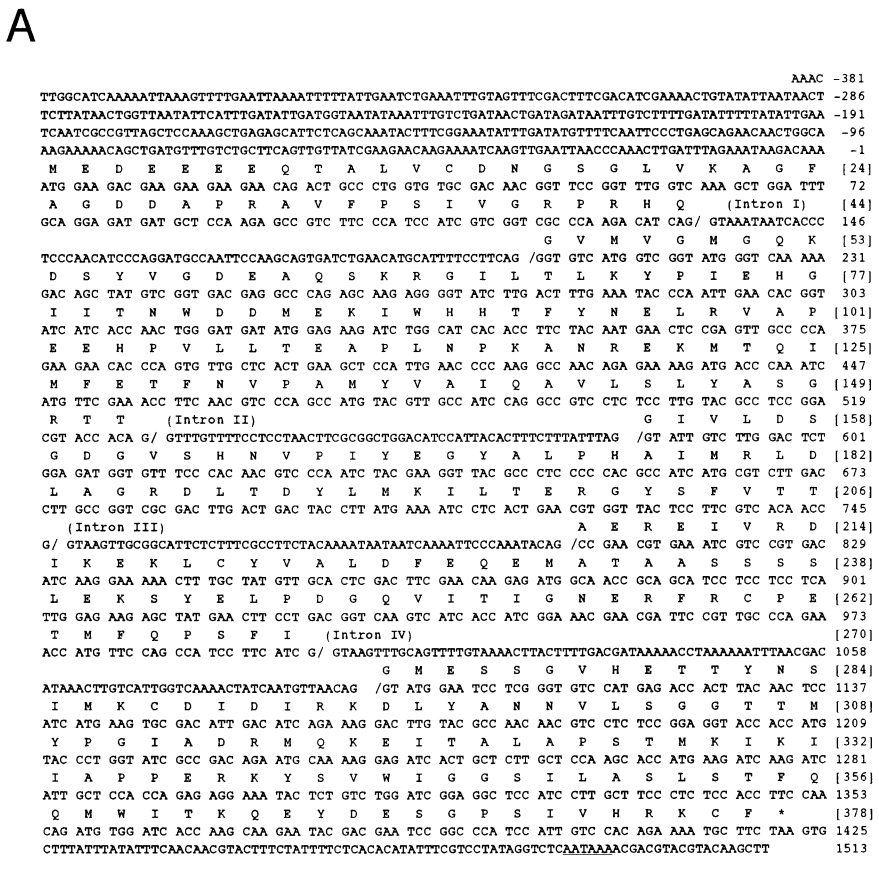

Fig. 1. The exon–intron organization of the *MocuMA2* **(A)** and *HrCA1* **(B)** genes. The nucleotide and deduced amino acid sequences are shown. The amino acid residues are represented by *one-letter symbols.* The translation termination codon is indicated by an *asterisk.* Nucleotide and amino acid positions are indicated relative to a putative translation start site. The putative polyadenylation signals are *underlined.*

Table 1. Comparison of the diagnostic amino acid positions in actins of various organisms[a]

Amino acid position[b]	Human α-skeletal	Ascidian HrMA4	Ascidian MocuMA2	Human β-cyto-plasmic	Ascidian HrCA1	Starfish muscle	Starfish cyto-plasmic	*Drosophila* muscle	*Drosophila* cytoplasmic	*C. elegans* muscle	*Arabidopsis*
5	Thr	Thr	Gln	Ile	Val	Val	Val	Ala	Val	Val	Ile
6	Thr	Thr	Thr	Ala	Ala	Ala	Ala	Ser	Ala	Ala	Gln
10	Cys	Cys	Cys	Val	Val	Val	Val	Val	Val	Val	Cys
16	Leu	Leu	Leu	Met	Met	Met	Met	Met	Met	Met	Met
17	Val	Val	Val	Cys	Cys	Cys	Cys	Cys	Cys	Cys	Val
76	Ile	Ile	Ile	Val	Val	Val	Val	Ile	Val	Val	Val
103	Thr	Thr	Val	Val	Val	Val	Val	Val	Val	Val	Ile
129	Val	Val	Val	Thr	Thr	Ser	Thr	Ser	Thr	Thr	Ala
153	Leu	Leu	Leu	Met	Phe	Phe	Phe	Leu	Leu	Leu	Leu
162	Asn	Asn	Asn	Thr	Thr	Thr	Thr	Thr	Thr	Thr	Thr
176	Met	Ala	Met	Leu	Leu	Leu	Leu	Leu	Leu	Leu	Leu
201	Val	Val	Val	Thr	Thr	Thr	Thr	Thr	Thr	Thr	Thr
225	Asn	Gln	Gln	Gln	Thr	Gln	Gln	Gln	Gln	Gln	Gln
260	Thr	Thr	Thr	Ala	Ala	Thr	Ala	Ala	Ala	Ala	Val
267	Ile	Ile	Ile	Leu	Leu	Ile	Leu	Leu	Leu	Leu	Ile
272	Ala	Ala	Ser	Cys	Ala	Ala	Ala	Cys	Cys	Ala	Ala
279	Tyr	Tyr	Tyr	Phe	Tyr	Tyr	Tyr	Tyr	Tyr	Tyr	Tyr
287	Ile	Ile	Ile	Val	Val	Ile	Val	Val	Val	Ile	Val
297	Asn	Asn	Asn	Thr	Thr	Thr	Thr	Asn	Thr	Thr	Ile
365	Ala	Ala	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser

[a] Amino acid residues of actins from various organisms are compared at positions that differentiate the mammalian α-striated muscle actin from the mammalian β-cytoplasmic actin. Amino acids idential to those of the rat α-skeletal muscle actin are shaded
EMBL/GenBank/DDBJ accession number of each actin is as follows: human α-skeletal muscle actin M20543; ascidian larval muscle actin D10887; human β-cytoplasmic actin M10277; starfish muscle actin M26500; starfish cytoplasmic actin M26501; *Drosophila melanogaster* muscle actin (79B) M18829; *Drosophila melanogaster* cytoplasmic actin (5C) K00667; *Caenorhabditis elegans* muscle actin X16796; *Arabidopsis thaliana* actin M20016

[b] The position numbers of the amino acid residues in the mammalian α-actin based on Vandekerckhove and Weber (1984)

of *MocuMA2* is similar to the *M. citrina* adult muscle actin gene *McMA1* (Swalla et al. 1994), suggesting that *MocuMA2* is an *M. oculata* adult muscle actin gene.

Comparison of Exon–Intron Organization of Various Actin Genes

The exon–intron organization of the *MocuMA2* and *HrCA1* genes was compared with that of other ascidian and selected nonascidian actin genes (Fig. 2). Six intron positions (41/42, 121/122, 150, 204, 268, 328/329) are present in both muscle and nonmuscle actin genes in deuterostomes, suggesting that deuterostome actin gene families evolved from an ancestral actin gene. Each intron position of the ascidian muscle actin genes is shared with the vertebrate and echinoderm muscle actin genes, although the number and length of introns vary among these genes. The ascidian adult muscle actin gene *MocuMA2* has an intron at position 41/42. This intron position is common to vertebrate and echinoderm actin genes but absent in ascidian larval muscle actin genes.

Three intron positions (41/42, 150, 204) of *HrCA1* are the same as those of the vertebrate and echinoderm actin genes. It is interesting that *HrCA1* contains an intron at position 150, because an intron at this position has not been reported in vertebrate and echinoderm cytoplasmic actin genes (Erba et al. 1988; Kowbel and Smith 1989). Three other introns (113/114, 246/247, 308) of *HrCA1* are located at unique positions with respect to other deuterostome and most protostome actins. An intron at position 308 is also present in the *Drosophila melanogaster* muscle actin genes 79B and 88F (Fyrberg et al. 1981).

Comparison of Amino Acid Sequences of Various Actins

Table 1 shows a comparison of the amino acid residues of various actins at positions that distinguish the mammalian α-striated muscle actin from the β-cytoplasmic actin (Vandekerckhove and Weber 1978, 1979). The HrMA4 and mammalian α-skeletal muscle actins share 18 of 20 diagnostic amino acid positions (Kusakabe et al. 1992). Similarly, MocuMA2 actin shares 15 of 20 diagnostic amino acids with the vertebrate muscle actin. The identity of diagnostic positions suggests that the ascidian muscle actins are more closely related to vertebrate muscle actin than to vertebrate cytoplasmic actins (Tomlinson et al. 1987; Kusakabe et al. 1992; Kovilur et al. 1993). Similarly, HrCA1 and the mammalian cytoplasmic actins share 15 of diagnostic amino acids, suggesting a close relationship between ascidian and vertebrate cytoplasmic actins.

The length and sequence of the amino-terminal regions are highly variable among actin isoforms and dif-

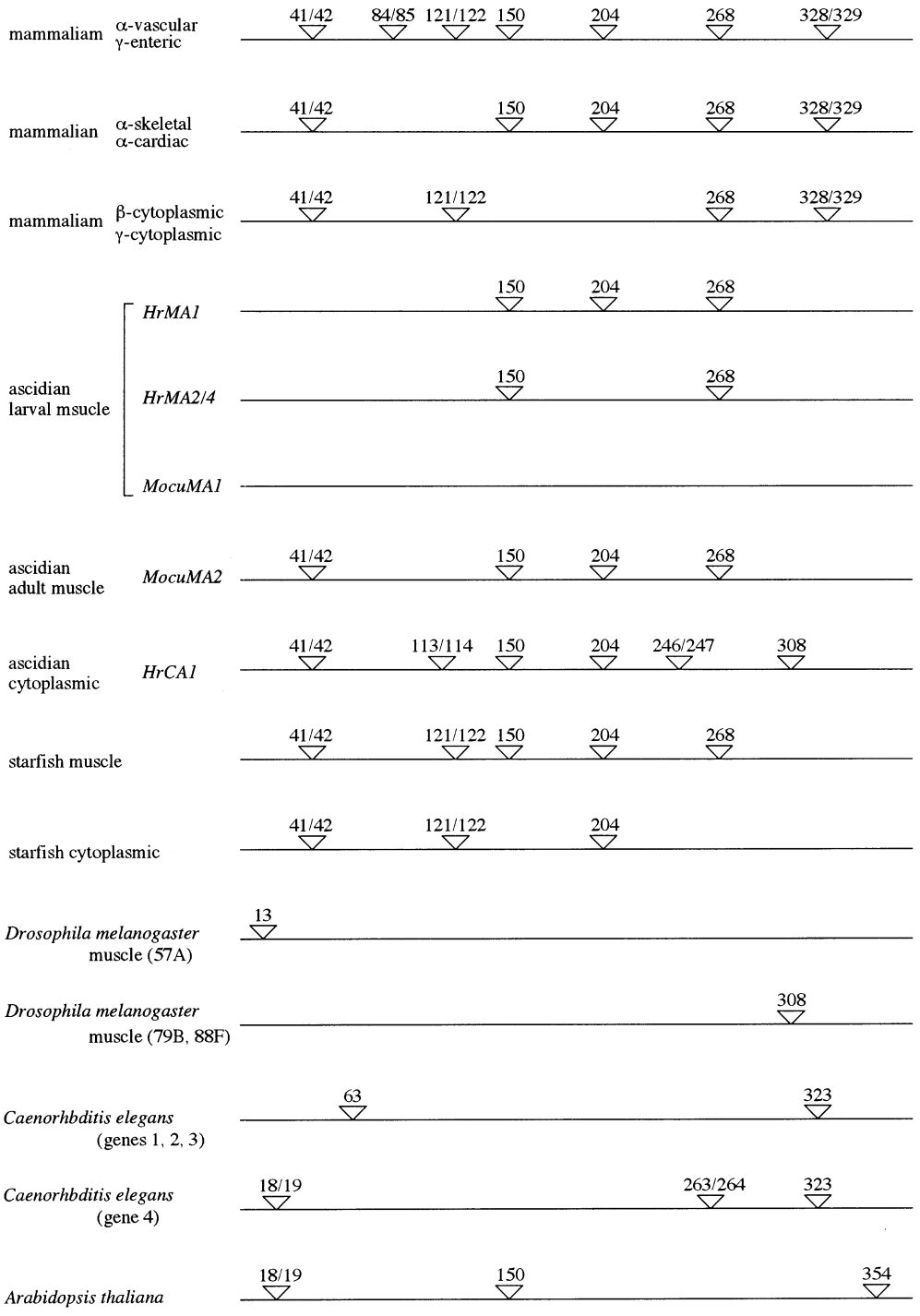

Fig. 2. Comparison of positions of introns in actin genes of various organisms. The *horizontal bars* represent actin proteins with the amino terminus at the left and the carboxy terminus at the right. The positions of introns in the genes are indicated by *open triangles* relative to the amino acid sequence of the protein. *Numbers above the triangles* indicate the position numbers of the amino acid residues at which introns interrupt the coding sequence. The numbering system of amino acid positions is based on Vandekerckhove and Weber (1984). Mammalian actin genes have an additional intron in the 5′-untranslated region (not shown; Miwa et al. 1991). Data on these gene structures were taken from the following sources: mammalian α-vascular actin, Ueyama et al. (1984); mammalian γ-enteric actin, Miwa et al. (1991); mammalian α-skeletal actin, Zakut et al. (1982); mammalian α-cardiac actin, Hamada et al. (1982); mammalian β-cytoplasmic actin, Nakajima-Iijima et al. (1985); mammalian γ-cytoplasmic actin, Erba et al. (1988); acidian larval muscle actins, Kusakabe et al. (1992, 1995, 1996); starfish actins, Kowbel and Smith (1989); *Drosophila melanogaster* actins, Fyrberg et al. (1981); *Caenorhabditis elegans* actins, Krause et al. (1989); *Arabidopsis thaliana* actin, Nairn et al. (1988).

```
HrMA4           MSDG-EEDTTAIVCDNGSGLVKSGFAGDDAPRAVFPSIVGRPRHQGVMVGMGQKDSYVGDEAQSKRGILTLKYPIEHGIITNWDDMEKIWHHTFY 91
HrMA1           ....-...Q.........................T.............................L...G.................
ScTb1           ....-...Q.....................................................................M.
MocuMA1         ..SE-...Q............A.....................................................................
MocuMA2         .E.E-..EQ..L.........A.....................................................................
McMA1           -----..EQ..L.........A.....................................................................
SpMA1           .E.DQD.EQ..L.........A..P..AP......LT.....................................................N.........
starfish-m      .C.---..VA.L.V.....MC.A.....................................................V.............
starfish-c      .C.---..VA.L.V.....MC.A.....................................................V.............
HrCA1           .C.---..VA.L.V.....MC.A.....................................................V.............
SpCA8           .----DDEVA.L.V.....MC.A...E.....................V.....................................V.............
β-cytoplasmic   .----DD.IA.L.V.....MC.A.....................................................V.............
γ-cytoplasmic   .----..EIA.L.I.....MC.A.....................................................V.............
α-skeletal      .C.--.DE...L.........A.....................................................................
α-cardiac       .C.--D.E...L.........A.....................................................................
α-vascular      .CE--...S..L........C.A.....................................................................S..
γ-enteric       .C---..E...L........C.A.....................................................................S..
A.thaliana      .A..--..IQPL....T.M..A.....................T.........A.....................VN.............

HrMA4           NELRVAPEEHPTLLTEAPLNPKANREKMTQIMFETFNVPAMYVAIQAVLSLYASGRTTGIVLDAGDGVSHNVPIYEGYALPHAIARLDLAGRDLT 186
HrMA1           ..........................................................S..........................
ScTb1           ................H.........................................M..........................
MocuMA1         ....I.....................................................S...I.................M........
MocuMA2         ..........V...............................................S...................M.........
McMA1           ..........V...............................................S...................M.........
SpMA1           ..........V...............................................S...................M.........
starfish-m      ..........V.....................S.....................F.T.....T............L.........
starfish-c      ..........V.....................T.....................F.S.....T............L.........
HrCA1           ..........V.....................T.....................F.S.....T............L.........
SpCA8           ..........V.....................TR....N.............L.M.S.....T............F.........
β-cytoplasmic   ..........V.....................T.....................M.S....T.T............L.........
γ-cytoplasmic   ..........V.....................T.....................M.S....T.T............L.........
α-skeletal      ..........................................................S....T............M.........
α-cardiac       ..........................................................S....T............M.........
α-vascular      ..........................................................S....T............M.........
γ-enteric       ..........................................................S....T............M.........
A.thaliana      ..........I.....................A.....................S.....T............L.........

HrMA4           DYLMKILTERGYSFVTTAEREIVRDIKEKLCYVALDFEQEMATAASSTSLEKSYELPDGQVITIGNERFRCPETLFQPSFIGMESAGIHETTYNS 281
HrMA1           ..........................................................................................
ScTb1           ..........................................................................................
MocuMA1         E..T.............K.........................................................................
MocuMA2         ..........................................S.....................M.........S.V.......
McMA1           ..........................................S.....................G...M.........S.V.......
SpMA1           ..........................................S.....................................S.V.......
starfish-m      ............T.............T........Q..S..S.........................A.................
starfish-c      ............T.....................Q.....S.....................A.....L.............
HrCA1           ............T...........A.....T..Q...T.S.I.............V.........A.....L.............
SpCA8           ............T..............G....S..G..S..SV.....................AM.....L.............
β-cytoplasmic   ............T.........................S.....................A.....L....C......F..
γ-cytoplasmic   ............T.........................S.....................A.....L....C......F..
α-skeletal      .....................................N........S.....................................
α-cardiac       .....................................N........S.....................................
α-vascular      .....................................N........S.....................................
γ-enteric       .....................................N........S.....................................
A.thaliana      .A..........T.............I...Y...LE..KT.S.V..N...........S.......V.Y...M....N.........

HrMA4           IMKCDIDIRKDLYANNVLSGGTTMYPGIADRMQKEITALAPSTMKIKIIAPPERKYSVWIGGSILASLSTFQQMWISKQEYDEAGPSIVHRKCF  375
HrMA1           ...................S..........................................................T..............
ScTb1           ..........................................................................................
MocuMA1         ..................................................................S.....N...T.A.............
MocuMA2         ..........................................................................T......S.........
McMA1           ..........................................................................T......S.........
SpMA1           ..........................................S...............A.............T......S.........
starfish-m      ............T......S............Q....P.....................................S.........
starfish-c      .....V.........T.....S..F..........V.....P.....................................S.........
HrCA1           .....V.........T.....S..F..........V.....P.....................................S.........
SpCA8           .....V.........T.....................S....P.....................................S.........
β-cytoplasmic   .....V.........T.....................................................................S.........
γ-cytoplasmic   .....V.........T.....................................................................S.........
α-skeletal      ................M.....................................................................T.........
α-cardiac       ..........................................................................................
α-vascular      ..........................................................................................
γ-enteric       ..........................................................................P...............
A.thaliana      .....V.......G.I........F.......S.........S....VV.....................L.....A.A....S.........
```

Fig. 3. Alignment of the amino acid sequences of various deuterostome actins and a plant actin. Amino acid sequences of nine ascidian actins (HrMA4, HrMA1, ScTb1, MocuMA1, MocuMA2, McMA1, SpMA1, HrCA1, and SpCA8), two echinoderm actins (starfish-m, *Pisaster ochraceus* muscle actin; starfish-c, *Pisaster ochraceus* cytoplasmic actin), six human actins, and an *Arabidopsis thaliana* actin are compared. Amino acids are indicated with *one-letter codes.* The entire sequence of HrMA4 is shown. The *dots* represent amino acids identical to those of HrMA4 and the *letters* represent variable positions in other actins. *Dashes* indicate gaps introduced in the sequence to optimize the alignment. The first eight positions (−4 to 4) are not used for the phylogenetic analyses shown in Fig. 4. The amino-terminal amino acids of McMA1 have not been determined (Swalla et al. 1994). The *numbering* of the amino acid residues is according to Vandekerckhove and Weber (1984). Sources, references, and accession numbers for the actin sequences are described in Materials and Methods.

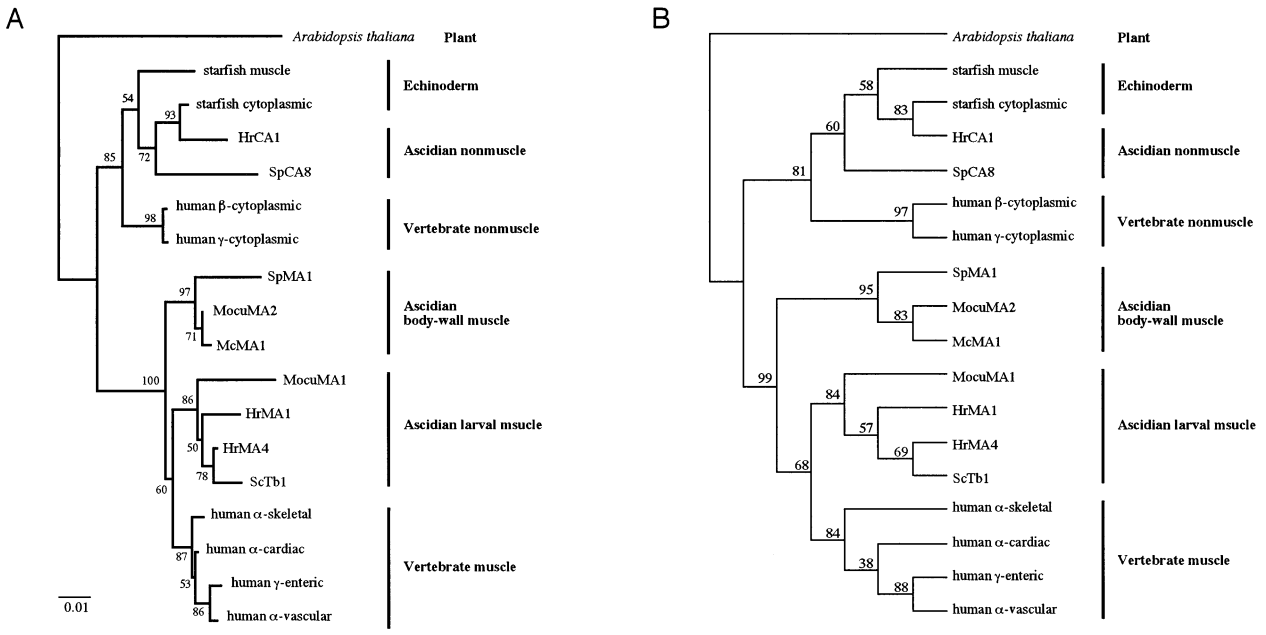

Fig. 4. Molecular phylogenetic analysis of deuterostome actins. The carboxy-terminal 371 amino acids were subjected to phylogenetic analysis. A plant actin (*Arabidopsis thaliana* AAc1) was included as the outgroup. **A** Phylogenetic tree inferred by the neighbor-joining method. Branch lengths are proportional to evolutionary distances. *Scale bar* indicates an evolutionary distance of 0.01 amino acid substitution per position in the sequence. A tree with similar topology was inferred in an analysis using the entire actin coding sequence (see text). **B** Phylogenetic tree inferred by the maximum parsimony method. *Numbers* shown in both trees are percentages of 1,000 bootstrap replicates in which the same internal branch was recovered. Sources, references, and accession numbers for the actin sequences are described in Materials and Methods.

ferent species (Fig. 3). While vertebrate muscle actins and most invertebrate actins have a Met-Cys sequence followed by a cluster of acidic amino acids (Glu and/or Asp), the vertebrate cytoplasmic actins lack a Cys residue next to the first Met. The entire amino acid sequence of ascidian muscle actins is similar to that of the vertebrate muscle isoforms as mentioned above (Table 1), whereas the amino-terminal sequence of the ascidian muscle isoforms is unique and lacks a Cys residue next to the first Met (Fig. 3). HrCA1 has an amino-terminal sequence with a Met-Cys sequence and thus resembles the starfish actins. In contrast, the amino-terminal sequence of SpCA8 lacks a Cys residue and is similar to the vertebrate cytoplasmic actins (Fig. 3).

Molecular Phylogenetic Analysis of Deuterostome Actins

To examine the evolutionary relationships of deuterostome actin isoforms, we performed a molecular phylogenetic analysis of actin amino acid sequences. Phylogenetic trees were constructed using the neighbor-joining (NJ) (Saitou and Nei 1987) (Fig. 4A) and maximum parsimony (MP) methods (Fig. 4B). The *Arabidopsis thaliana* actin (AAc1; Nairn et al. 1988) was used as the outgroup in both analyses. Since the lengths of the amino-terminal region of actins are highly variable, we excluded the sequence from the first Met to the end of the cluster of acidic amino acids (from −4 to 4; Fig. 3) from the analyses. To evaluate the effect of excluding this amino-terminal sequence from the analysis, we also constructed a molecular phylogenetic tree based on the entire amino acid sequences (Fig. 3) using the NJ method. The topology of the tree obtained was the same except that the mammalian cardiac actin was grouped with the mammalian skeletal muscle actin.

The NJ and MP analyses gave essentially identical results with respect to the major clusters identified (Fig. 4). The MP method generated one minimal tree that was identical to the consensus tree obtained by bootstrap resampling (Fig. 4B). The grouping of all ascidian muscle actins with the vertebrate muscle actins was strongly supported in 99% or more of the bootstrap replicates. The echinoderm muscle actin grouped with all deuterostome cytoplasmic actins and was separate from other deuterostome muscle actins (supported by more than 80% of bootstraps). This is consistent with the view obtained by comparing diagnostic amino acids (Table 1). The chordate muscle actin clade consisted of three branches, each supported by relatively high bootstrap values (84% or more). One of the branches contained the vertebrate α-skeletal, α-cardiac, α-vascular, and γ-enteric muscle actin isoforms. The other two branches contained the ascidian larval and ascidian body-wall actins. The presence of two distinct lineages of ascidian muscle actins is consistent with our previous study showing that the larval and adult muscle actins are distinguished by diagnostic amino acids (Kusakabe 1995). Among the three clades of the chordate muscle actins, a closer relationship of vertebrate muscle and ascidian larval muscle

actins was supported in 60% (NJ) or 68% (MP) of the bootstrap replicates, suggesting an earlier divergence of acidian adult body-wall isoforms.

The vertebrate β- and γ-cytoplasmic actins grouped together in 97% or more of the bootstrap replicates. The formation of a clade including the ascidian cytoplasmic actin HrCA1 and the echinoderm cytoplasmic actin with the ascidian cytoplasmic actin SpCA8 as the outgroup was supported by 93% (NJ) or 83% (MP) of bootstrap replicates. The position of the echinoderm muscle actin was different in the NJ and MP trees, but bootstrap values supporting the different topologies were relatively small in both analyses (54% in NJ and 58% in MP).

Discussion

In the present study, we have compared the exon–intron organization and the deduced amino acid sequences of various actin genes. These actin sequences were used in molecular phylogenetic analyses to gain new insights into the evolution of the actin gene family in chordates.

The intron positions of ascidian muscle actin genes were shown to be identical to those of vertebrate muscle actin genes. This is consistent with our molecular phylogenetic analyses showing a closer relationship of ascidian muscle actin genes to the vertebrate muscle actin genes. However, the number of introns in the ascidian larval muscle actin genes is smaller than that in other deuterostome actin genes. An extreme case is the *Molgula oculata* muscle actin gene *MocuMA1*, which contains no introns (Kusakabe et al. 1996). The primitive situation in deuterostomes, however, seems to be muscle actin genes with introns (see below). The ascidian larval muscle actin genes may have lost their introns to expedite the processing and cytoplasmic accumulation of transcripts during the relatively short interval of muscle cell differentiation during larval development (Kusakabe et al. 1996).

Six intron positions (41/42, 121/122, 150, 204, 268, 328/329) are conserved between muscle actin genes and nonmuscle actin genes in deuterostomes. The ancestral deuterostome may have had a single prototypic actin gene that contained seven or more introns. Since the number of introns varies from zero to seven in the extant deuterostome actin genes, different introns seem to have been lost during the evolution of each lineage. The conservation of intron positions in both the deuterostome cytoplasmic and muscle actin genes suggests that the ancestral vertebrate-type muscle actin gene appeared during chordate evolution, and that its characteristic amino acid sequence was established in a relatively short time. Since the intron at position 328/329 is only present in the vertebrate actin genes, however, this intron might have been acquired during vertebrate evolution and have been shared between muscle and cytoplasmic actin genes via gene conversion.

The intron at position 308 in the *HrCA1* cytoplasmic actin gene is unique among deuterostome actin genes but is also present in *Drosophila melanogaster* muscle actin genes 79B and 88F (Fyrberg et al. 1981). Common intron positions in the actin genes of distantly related species are known for plant and vertebrate muscle actin genes (position 150) and for plant actin genes and a *Caenorhabditis elegans* actin gene (position 18/19). The presence of these conserved intron positions supports the hypothesis that the ancient eukaryote actin gene had a large number of introns (Doolittle 1978). Although it is uncertain at present whether conserved intron positions in the actin genes of distantly related organisms had the same origin, further analysis of exon–intron organization among these genes would provide important information about the origin of introns in actin genes.

Comparison of the diagnostic amino acids and phylogenetic analysis revealed that both ascidian larval and adult muscle actins are more similar to the vertebrate muscle actin than to the vertebrate cytoplasmic actin. In contrast, nonchordate muscle actins, including an echinoderm muscle actin, are more closely related to the vertebrate cytoplasmic actins than to the vertebrate muscle actin (Vandekerckhove and Weber 1984). In addition, the starfish and *Drosophila* muscle and nonmuscle isoforms are more similar to each other than to the actins of other animals (Table 1). Thus, it is likely that muscle isoforms emerged several times independently during metazoan evolution, as suggested for arthropod actins (Mounier et al. 1992). Gene conversion has maintained homogeneity between the sea urchin muscle and nonmuscle isoforms (Crain et al. 1987). In many invertebrates, however, muscle actins show more amino acids characteristic of the vertebrate muscle actin than do nonmuscle actins. For example, a starfish muscle actin has amino acids characteristic of vertebrate muscle actin at five positions (260, 267, 272, 279, and 287), while starfish cytoplasmic actin has two (positions 272 and 279). Similarly, four muscle-type amino acids are present in *Drosophila* muscle actin (positions 76, 153, 279, and 297), while *Drosophila* cytoplasmic actin has two muscle-type amino acids (positions 153 and 279). This feature may be a consequence of convergent evolution related to muscle contractile properties. It is uncertain, however, whether the chordate muscle-like features of echinoderm muscle actin are representative of a transition from a nonmuscle-type to a muscle-type actin. The chordate muscle-type actins probably diverged from a nonmuscle-like actin before the divergence of urochordates and vertebrates.

Kovilur et al. (1993) proposed that the divergence of skeletal and cardiac isoforms of vertebrate muscle actin occurred before the emergence of urochordates. However, a detailed comparison of amino acid sequences of ascidian actins suggested that the divergence of two sar-

comeric actin in vertebrates occurred after urochordates separated from the vertebrate lineage (Kusakabe 1995). Our molecular phylogenetic analyses showed that the four muscle actin isoforms in vertebrates (α-skeletal, α-cardiac, α-vascular, and γ-enteric) are more closely related to each other than to the ascidian muscle isoforms. Similarly, the β- and γ-cytoplasmic actins of vertebrates show a closer relationship to each other than to the ascidian and echinoderm nonmuscle isoforms. These results suggest that the vertebrate actin gene family was established by the duplication of one ancestral muscle actin gene and one ancestral cytoplasmic actin gene after the divergence of the vertebrate and urochordate lineages. The diversification of multigene families is thought to have played an important role during vertebrate evolution (Miyata et al. 1994; Iwabe et al. 1995) and may have coincided with the evolution of the complex vertebrate body plan.

The comparison of the amino-terminal sequences and the molecular phylogenetic analyses showed that the HrCA1 cytoplasmic actin is closely related to the echinoderm cytoplasmic actin. In contrast the SpCA8 cytoplasmic actin is more distantly related to the echinoderm cytoplasmic actin, and its amino-terminal sequence is similar to that of the vertebrate cytoplasmic actins. These results suggest that at least two types of nonmuscle actins are present in ascidians. Multiple nonmuscle actin genes in ascidians were suggested by genomic Southern hybridization (Beach and Jeffery 1990) and the expression pattern of a cytoplasmic actin gene (Araki et al. 1996). Since a vertebrate-type cytoplasmic actin lacking a Cys residue in the amino-terminal region has not been reported in echinoderms, the vertebrate-type cytoplasmic actin genes may have arisen from an ancestral actin gene by losing the Cys residue.

Both the NJ and MP trees indicated that the ascidian larval muscle actin is more closely related to the vertebrate muscle actin than the ascidian adult actin (Fig. 4). This suggests that the chordate ancestor had at least two muscle actin isoforms: the ancestral adult-muscle actin and ancestral vertebrate/larval muscle actin. Since the bootstrap value supporting the vertebrate-muscle/ascidian-larval actin clade is relatively low (60% in NJ tree and 68% in MP tree), however, an alternative possibility is that the adult isoform appeared in the urochordate lineage after the vertebrate lineage diverged and that the ancestral adult actin evolved rapidly. The similarity of amino-terminal sequence in the ascidian larval and adult muscle actins also supports this alternative evolutionary pathway.

Acknowledgments. This work was supported by Grants-in-Aid from the Ministry of Education, Science and Culture, Japan, for Scientific Research (08780694) to T.K. and for Specially Promoted Research (07102012) to N.S.; by NSF grant DCB-9115543 to W.R.J.; and by NIH grant HD-113970 to W.R.J.

References

Alonso S (1987) Coexpression and evolution of the two sarcomeric actin genes in vertebrates. Biochemie 69:1119–1125

Araki I, Tagawa K, Kusakabe T, Satoh N (1996) Predominant expression of a cytoskeletal actin gene in mesenchyme cells during embryogenesis of the ascidian *Halocynthia roretzi.* Dev Growth Differ 38:401–411

Beach RL, Jeffery WR (1990) Temporal and spatial expression of a cytoskeletal actin gene in the ascidian *Styela clava.* Dev Genet 11:2–14

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

Berrill NJ (1955) The origin of vertebrates. Oxford University Press, Oxford

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

Doolittle WF (1978) Genes in pieces: were they ever together? Nature 272:581–582

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

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791

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

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

Garstang W (1928) The morphology of the tunicata and its bearing on the phylogeny of the chordata. Q J Microsc Sci 72:51–187

Haeckel E (1868) Natürliche Schöpfungsgeschichte. Georg Reimer, Berlin

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

Holland PWH, Garcia-Fernàndez J (1996) Hox genes and chordate evolution. Dev Biol 173:382–395

Iwabe N, Kuma K, Miyata T (1995) Evolution of gene families and relationship with organismal evolution: rapid divergence of tissue-specific genes in the early evolution of chordates. Mol Biol Evol 13:483–493

Jefferies RPS (1986) The ancestry of the vertebrates. British Museum (Natural History), London

Kamada S, Kakunaga T (1989) The nucleotide sequence of a human smooth muscle alpha-actin (aortic type) cDNA. Nucleic Acids Res 17:1767

Kimura M (1983) The neutral theory of molecular evolution. Cambridge University Press, Cambridge, England, p 75

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

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

Krause M, Wild M, Rosenzweig B, Hirsh D (1989) Wild-type and mutant actin genes in *Caenorhabditis elegans.* J Mol Biol 208:381–392

Kusakabe T (1995) Expression of larval-type muscle actin-encoding genes in the ascidian *Halocynthia roretzi.* Gene 153:215–218

Kusakabe T, Suzuki J, Saiga H, Jeffery WR, Makabe KW, Satoh N (1991) Temporal and spatial expression of a muscle actin gene

during embryogenesis of the ascidian *Halocynthia roretzi*. Dev Growth Differ 33:227–234

Kusakabe T, Makabe KW, Satoh N (1992) Tunicate muscle actin genes: structure and organization as a gene cluster. J Mol Biol 227:955–960

Kusakabe T, Hikosaka A, Satoh N (1995) Coexpression and promoter function in two muscle actin gene complexes of different structural organization in the ascidian *Halocynthia roretzi*. Dev Biol 169: 461–472

Kusakabe T, Swalla BJ, Satoh N, Jeffery WR (1996) Mechanism of an evolutionary change in muscle cell differentiation in ascidians with different modes of development. Dev Biol 174:379–392

Miwa T, Kamada S (1990) The nucleotide sequence of a human smooth muscle (enteric type) gamma-actin cDNA. Nucleic Acids Res 18: 4263

Miwa T, Manabe Y, Kurokawa K, Kamada S, Kanda N, Bruns G, Ueyama H, Kakunaga T (1991) Structure, chromosome location, and expression of the human smooth muscle (enteric type) gamma-actin gene: evolution of six human actin genes. Mol Cell Biol 11:3296–3306

Miyata T, Kuma K, Iwabe N, Nikoh N (1994) A possible link between molecular evolution and tissue evolution demonstrated by tissue specific genes. Jpn J Genet 69:473–480

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

Nairn CJ, Winesett L, Ferl RJ (1988) Nucleotide sequence of an actin gene from *Arabidopsis thaliana*. Gene 65:247–257

Nakajima-Iijima S, Hamada H, Reddy P, Kakunaga T (1985) Molecular structure of the human cytoplasmic β-actin gene: interspecies homology of sequences in the introns. Proc Natl Acad Sci USA 82:6133–6137

Satoh N (1994) Developmental biology of Ascidians. Cambridge University Press, New York

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

Sanger F, Nichlen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5467

Swalla BJ, White ME, Zhou J, Jeffery WR (1994) Heterochronic expression of an adult muscle actin gene during ascidian larval development. Dev Genet 15:51–63

Taylor A, Erba HP, Muscat GE, Kedes L (1988) Nucleotide sequence and expression of the human skeletal alpha-actin gene: evolution of functional regulatory domains. Genomics 3:323–336

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680

Tomlinson CR, Beach RL, Jeffery WR (1987) Differential expression of a muscle actin gene in muscle cell lineages of ascidian embryos. Development 101:751–765

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 (1978) 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 sequence of actins from bovine aorta, bovine heart, bovine fast skeletal muscle and rabbit slow skeletal muscle. Differentiation 14:123–133

Vandekerckhove J, Weber K (1984) Chordate muscle actins differ distinctly from invertebrate muscle actins. The evolution of the different vertebrate muscle actins. J Mol Biol 179:391–413

Wada H, Satoh N (1994) Details of the evolutionary history from invertebrates to vertebrates, as deduced from the sequences of 18S rDNA. Proc Natl Acad Sci USA 91:1801–1804

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