

Isolation and Molecular Phylogeny of Three Muscle Actin Isoforms of an Endangered Freshwater Fish Species *Hemibarbus mylodon* (Cypriniformes; Cyprinidae)

Keun-Yong Kim and Yoon Kwon Nam*

Department of Aquaculture, Pukyong National University, Busan 608-737, Korea

The Korean doty barbel *Hemibarbus mylodon* (Cypriniformes; Cyprinidae) is a critically endangered freshwater fish species mainly because of its natural habitat degradation. Three full-length complementary DNA (cDNA) clones representing different muscle actin isoforms were isolated and characterized. The three muscle actin isoforms were 1,294-1,601 bp long with the identical open reading frames of 1,134 bp with the deduced amino acid residues of 377. They showed 83.9-87.2% identities in the coding nucleotide level and 96.8-98.1% identities in the amino acid level. Phylogenetic analysis with the coding nucleotide sequences revealed that three muscle actin isoforms of *H. mylodon* formed strongly supported monophyletic groups with one of cypriniform skeletal α -actin (*acta1*), cypriniform aortic α -actins (*acta2*), and uncharacterized *Danio rerio* muscle actin isoform/*Salmo trutta* slow muscle actin (a novel muscle actin type). Our phylogenetic tree further suggested that cypriniform *acta2* only showed the orthologous relationship to tetrapod *acta2*. Other multiple actin isoforms from diverse teleostean taxa were however clustered to no tetrapod orthologs, i.e., *acta1*, cardiac α -actin (*actc1*), *acta2*, and enteric γ -actin (*actg2*). This result strongly suggested that teleostean muscle actins have experienced different and complicated evolutionary history in comparison to mammalian counterparts.

Keywords: *Hemibarbus mylodon*, Muscle actin isoform, cDNA, Phylogeny

Introduction

The Korean doty barbel *Hemibarbus mylodon* (Cypriniformes; Cyprinidae) is an endemic freshwater fish species only found in the Korean peninsula (Kim et al., 2005). During the last decade its population size has been significantly decreased because of various anthropogenic impacts such as dam constructions, water pollution, and dredging to mine sand around natural habitats (Jang et al., 2003). Moreover the genetic diversity of *H. mylodon* populations is reported to be extremely low as assessed by polymorphic microsatellite markers (Kim et al., 2007) and amplified fragment length polymorphism (AFLP) analysis (Lee et al., 2008). Due to such a highly threatened status, conservation and restoration efforts for this species are urgently needed. In line with various activities for *in situ* and/or *ex situ* restoration of such an endangered fish species, the mining of important gene sequences is crucial for getting a deeper insight into its relevant physiology as well as for better understanding its phylogenetic relationship with conspecific members or related species.

Actins are ubiquitous and highly conserved proteins found in eukaryotic cells. The multiple actin family plays key roles in maintaining cytoskeletal structure, cellular mobility, cell division, intracellular movements, and contractile processes, and exhibits distinct tissue- and stage-specific expression patterns (Vandekerckhove and Weber, 1979; Pollard and Cooper, 1986). Mammals have six actin isoforms: they consist of two striated muscle actins [skeletal α -actin (*acta1*) and cardiac α -actin (*actc1*)], two smooth muscle actins [aortic α -actin (*acta2*) and enteric γ -actin (*actg2*)], and two cytoplasmic actins [β -actin (*actb*) and γ -actin (*actg1*)] (Vandekerckhove and Weber, 1979; Miwa et al., 1991). Meanwhile nine actin isoforms were isolated from the pufferfish *Takifugu rubripes*, and they include six muscle actins [three cardiac, two skeletal, and one anomalous (testis type) α -actins] and three cytoplasmic actins [β -cytoplasmic actins 1 and 2, and β -cytoplasmic (vascular) actin] (Venkatesh et al., 1996). Genetic determinants of muscle actins have been isolated from diverse fish species (Watabe et al., 1995; Kim et al., 2000; Xu et al., 2000; Moutou et al., 2001; Krasnov et al., 2003), and presences of two *acta1* isoforms were reported in

*Corresponding author: yoonknam@pknu.ac.kr

three *Coryphaenoides* species (Morita, 2000, 2003) and *Theragra chalcogramma* (Tanaka et al., 2004).

Unlike mammals where four distinct muscle actin genes have been comprehensively characterized, the multiple muscle actin isoforms from a given teleost species has been little studied. Moreover, the high sequence identity and multiple gene duplications in the teleost lineage make it difficult to perform the orthology prediction of diverse teleost muscle actin genes to the four mammalian counterparts. No molecular phylogenetic study has been conducted to reveal the orthology of muscle actin isoforms of teleosts to date.

Recently, we have reported the molecular structure of two cytoplasmic actin genes isolated from *H. mylodon* and presented phylogenetic relationships among vertebrate cytoplasmic actins (Kim et al., 2008). Along with this previous study, we aimed to extend our knowledge on the actin gene family of *H. mylodon* by further mining of muscle actin isoforms from the expressed sequence tag (EST) database (Bang et al., 2007). In this study, we isolated three muscle actin mRNA species from *H. mylodon* and carried out the phylogenetic analysis within the vertebrate lineage to reveal their orthologous relationships and also to further discuss evolutionary history of teleost muscle actins in comparison to tetrapod counterparts.

Material and Methods

Fish specimen, cDNA library construction, and nucleic acids preparation

Fish specimens used in this study were adult individuals maintained at Soonchunhyang University, Korea. Information on the *H. mylodon* cDNA libraries constructed with diverse tissue types for EST analysis can be referred to Bang et al. (2007). For RT-PCR isolation of full-length open reading frame (ORF) sequence of muscle actin mRNA species, total RNA was extracted from skeletal muscle using TriPure Reagent (Roche Applied Science, Mannheim, Germany) according to the manufacturer's instruction manual. The RNA preparation was purified again using RNeasy Mini Clean-up kit (Qiagen, Hilden, Germany) including the DNase treatment according to the manufacturer's recommendation. Integrity of the extracted total RNA was confirmed by the ratio of 28S and 18S ribosomal RNA (rRNA) bands from MOPS-formaldehyde agarose gel electrophoresis.

Isolation of muscle actin cDNA sequences

From the survey of the *H. mylodon* ESTs (Bang et al.,

2007), 28 ESTs from the muscle and intestine sources matched significantly with previously known teleost muscle actin sequences based on the homology search against NCBI GenBank (<http://www.ncbi.nlm.nih.gov/>) using the BLASTx option. They were assembled into three contigs encoding muscle actins in Sequencher™ (Gene Codes, Ann Arbor, MI, USA). With the contig sequences full-length ORF sequences were isolated by RT-PCR using One-Step RT-PCR kit (Roche Applied Science, Mannheim, Germany). Three oligonucleotide primer pairs used for RT-PCR isolation were HMACT01 1F/1R (5'-AAACCAACCATGTGCGACGA-3' and 5'-ATGTACGGCTGAGACTGAGAGA-3'), HMACT02 1F/1R (5'-CTCCTGTTGGAGTAAGGAAG-3' and 5'-ATGTGGGCAGTGCAAAGACA-3') and HMACT03 1F/1R (5'-AAGCTCTCCTGGCTGTCTAA-3' and 5'-AACGCTCGCGTACAAACAC-3'). Total RNA (200 ng) from the muscle or intestine was reverse transcribed into cDNAs at 42°C for 60 min, followed by 30 cycles of 94°C for 45 s, 58°C for 45 s, and 72°C for 1 min with an initial denaturation step at 94°C for 4 min. RT-PCR product was cloned into pGEM®-T Easy Vector (Promega, Madison, WI, USA) according to the manufacturer's instruction. The insert DNA was read at both directions using ABI 3700 Automatic Sequence Analyzer (Applied Biosystems, Foster City, CA, USA) in order to confirm the sequence of each isoform. The muscle actin isoforms isolated in this study were deposited in GenBank under accession numbers, FJ713567-FJ713569.

Phylogenetic analysis

Full-length cDNA sequences of muscle actin isoforms for teleosts, and representative bird (*Gallus gallus*) and mammals (*Mus musculus* and *Homo sapiens*) were retrieved from GenBank (Table 1). Muscle actins of two ascidians (*Molgula oculata* and *Halocynthia roretzi*), which are the closest relatives of vertebrate muscle actins, were used as outgroups. The sequence data were aligned using ClustalW in BioEdit (Hall, 1999), and the output was used to reconstruct a phylogenetic tree. Redundant actin sequences belonging to the same species and clustered into the same clade or group were excluded in the phylogenetic analysis to reduce computational time.

The aligned data matrix was subjected to neighbor-joining (NJ) analysis with the coding nucleotide sequences with the Kimura 2-parameter model in PAUP* (Swofford, 2002). Three nucleotide codon positions in the reading frame were differentially weighted in a ratio of 2:3:1, because of satura-

Table 1. List of muscle actin sequences used for molecular phylogenetic analysis

Species	Gene name	Gene ^a	GenBank acc. no.
<i>Carassius auratus</i>	skeletal alpha-actin	<i>acta1</i>	D50029
<i>Coryphaenoides acrolepis</i>	alpha-actin type-2	<i>acta1.2</i>	AB021650
<i>Coryphaenoides acrolepis</i>	skeletal alpha-actin type-1	<i>acta1.1</i>	AB021649
<i>Coryphaenoides armatus</i>	skeletal alpha-actin type-2a	<i>acta1.2a</i>	AB086240
<i>Coryphaenoides armatus</i>	skeletal alpha-actin type-2b	<i>acta1.2b</i>	AB086241
<i>Coryphaenoides cinereus</i>	skeletal alpha-actin type-1	<i>acta1.1</i>	AB021651
<i>Coryphaenoides cinereus</i>	skeletal alpha-actin type-2	<i>acta1.2</i>	AB021652
<i>Coryphaenoides yaquinae</i>	skeletal alpha-actin type-2a	<i>acta1.2a</i>	AB086242
<i>Coryphaenoides yaquinae</i>	skeletal alpha-actin type-2b	<i>acta1.2b</i>	AB086243
<i>Cyprinus carpio</i>	skeletal alpha-actin	<i>acta1</i>	D50028
<i>Cyprinus carpio</i>	skeletal muscle actin	<i>acta1</i>	AY309091
<i>Danio rerio</i>	actin, alpha 1, skeletal muscle	<i>acta1</i>	NM_131591
<i>Danio rerio</i>	actin, alpha 2, smooth muscle, aorta	<i>acta2</i>	NM_212620
<i>Danio rerio</i>	actin, alpha, cardiac muscle 1	<i>actc1</i>	NM_214784
<i>Danio rerio</i>	actin, alpha, cardiac muscle 1 like	<i>actc1l</i>	NM_001001409
<i>Danio rerio</i>	zgc:112098	<i>zgc:112098</i>	NM_001017750
<i>Danio rerio</i>	zgc:86709	<i>zgc:86709</i>	NM_001002066
<i>Danio rerio</i>	zgc:86725	<i>zgc:86725</i>	NM_001002074
<i>Gadus morhua</i>	skeletal muscle alpha-actin	<i>acta1</i>	AF500273
<i>Gallus gallus</i>	actin, alpha 1, skeletal muscle	<i>ACTA1</i>	NM_001031063
<i>Gallus gallus</i>	actin, alpha 2, smooth muscle, aorta	<i>ACTA2</i>	NM_001031229
<i>Gallus gallus</i>	actin, alpha, cardiac muscle 1	<i>ACTC1</i>	NM_001079481
<i>Gallus gallus</i>	actin, gamma 2, smooth muscle, enteric muscle actin	<i>ACTG2</i>	NM_205172
<i>Halocynthia roretzi</i>	muscle actin	<i>HrMA4a</i>	D10887
<i>Hemibarbus mylodon</i>	skeletal alpha-actin	<i>acta1</i>	FJ713567
<i>Hemibarbus mylodon</i>	aortic alpha-actin	<i>acta2</i>	FJ713568
<i>Hemibarbus mylodon</i>	novel muscle actin type 1	–	FJ713569
<i>Homo sapiens</i>	actin, alpha 1, skeletal muscle	<i>ACTA1</i>	NM_001100
<i>Homo sapiens</i>	actin, alpha 2, smooth muscle, aorta	<i>ACTA2</i>	NM_001141945
<i>Homo sapiens</i>	actin, alpha, cardiac muscle 1	<i>ACTC1</i>	NM_005159
<i>Homo sapiens</i>	actin, gamma 2, smooth muscle, enteric	<i>ACTG2</i>	NM_001615
<i>Lampanyctus regalis</i>	alpha actin	–	AF503592
<i>Molgula oculata</i>	adult muscle-type actin	<i>MocuMA2</i>	D85743
<i>Mus musculus</i>	actin, alpha 1, skeletal muscle	<i>Acta1</i>	NM_009606
<i>Mus musculus</i>	actin, alpha 2, smooth muscle, aorta	<i>Acta2</i>	NM_007392
<i>Mus musculus</i>	actin, alpha, cardiac muscle 1	<i>Actc1</i>	NM_009608
<i>Mus musculus</i>	actin, gamma 2, smooth muscle, enteric	<i>Actg2</i>	NM_009610
<i>Notothenia coriiceps</i>	alpha actin	–	AF503590
<i>Oncorhynchus keta</i>	actin	–	AB032464
<i>Oreochromis mossambicus</i>	alpha-actin	–	AB037866
<i>Oryzias latipes</i>	cardiac muscle actin	<i>actc1</i>	AB016259
<i>Oryzias latipes</i>	muscle actin	<i>OIMA1</i>	NM_001104806
<i>Pleurogrammus azonus</i>	alpha skeletal actin	<i>acta1</i>	AB073381
<i>Salmo salar</i>	actin alpha 1-1 (actc1-1)	<i>actc1.1</i>	BT043779
<i>Salmo salar</i>	actin, alpha 1-2 (actc1-2)	<i>actc1.2</i>	BT043780
<i>Salmo salar</i>	actin, alpha 1-3 (actc1-3)	<i>actc1.3</i>	BT043781
<i>Salmo trutta</i>	alpha actin	–	AF267496
<i>Salmo trutta</i>	cardiac muscle actin	<i>actc1</i>	AF303985
<i>Scomber scombrus</i>	alpha actin	–	EF607093
<i>Siniperca chuatsi</i>	skeletal muscle alpha-actin	<i>acta1</i>	AY395872
<i>Sparus aurata</i>	skeletal alpha-actin	–	AF190473
<i>Sphyaena idiaestes</i>	alpha actin	–	AF503593
<i>Takifugu rubripes</i>	alpha-anomalous (testis) actin	<i>actx</i>	U38962
<i>Takifugu rubripes</i>	alpha-cardiac actin 1	<i>actc1.1</i>	U38959
<i>Takifugu rubripes</i>	alpha-cardiac actin 2	<i>actc1.2</i>	U38960
<i>Takifugu rubripes</i>	alpha-cardiac actin 3	<i>actc1.3</i>	U38961
<i>Takifugu rubripes</i>	alpha-skeletal actin 1	<i>acta1.1</i>	U38850
<i>Takifugu rubripes</i>	alpha-skeletal actin 2	<i>acta1.2</i>	U38958
<i>Theragra chalcogramma</i>	alpha skeletal actin-1	<i>acta1.1</i>	AB073379
<i>Theragra chalcogramma</i>	alpha skeletal actin-2	<i>acta1.2</i>	AB073380
<i>Trematomus bernacchii</i>	alpha actin	–	AF503589

^a–, the gene nomenclature not provided

```

acta1                               -44 catcagtcctgtga taacaacttgaccacagaaaaaccaacc
acta2                               -69 cccctcccc ggcccagcactgtcaggtgattgtctcctg ttggagtaaggaaggaatacagagctgaaa
novel                               -43 gttcaggcctcctt gtgcaagctctcctggctgtctaagcaaac

acta1 ATGTGCGACGACGAGGAGACTACCGCCCTT GTGTGCGACAACGGCTCTGGCTTGGTCAAG GCTGGATTGSCCGAGATGACGCTCCCCGT 90
acta2 .....T.....A..A.G..T.....T.....T..A..A..AC..TGT... ..A..C..C.....G.....T.....TA..A
novel .....A..A.....C.....T..G .....T.....T.....A..TC...G... ..C..C.....T..C.....C...A.G

acta1 GCTGTCTTCCCTCCATCGTTGGAAGGCC CGTCACCAGGGTGTGATGGTCGGTATGGGT CAGAAGGACAGCTACGTAGGTGATGAGGCT 180
acta2 ..A..G.....A..G..TC..T... A.A.....G..A.....G .....A..T..T..T..T.....G
novel ..C..G..T.....T..G..CC..A... ..C..T.....C.....C.....T.....A.....T.....A..C.....C

acta1 CAGAGCAAGAGAGGTATCTTGACCCTGAAG TACCCATTGAGCAGGCATCATCACTAAC TGGGACGACATGGAGAAGATCTGGCACCAC 270
acta2 ..A.....A.....A..C.....C... ..A.....C.....A..T .....A.....A.....A.....T...
novel .....C...C..C.....C..A .....T.....C... .....A.....A.....A.....T...

acta1 ACCTTCTACAATGAGCTGCGTGTGGCCCTT GAGGAGCACCAACCCCTGCTCACTGAAGCC CCTCTCAACCCCAAGGCCAACAGGGAGAAG 360
acta2 T.T.....C..A..C.....T..A .....C..T..T..G..A..... ..A.....T.....A.....
novel ..T.....C.....C.....CGTT..CT..G...G..T .....G..T..T.....

acta1 ATGACTCAGATCATGTTTGTAGACCTTCAAC GTCCCTGCCATGTATGTGGCCATCCAGGCT GTGCTGTCCCTGTACGCCTCCGGCCGTACC 450
acta2 .....T.....T.....A..T.....A..G .....C..T..T..T..T..A.....A
novel .....C.....T..G..T.....C.....T.....C.....T.....T.....

acta1 ACCGGTATCGTGTGGACGCTGGTACGGT GTGACCCACAACGTCCTCATGAGGGT TACGCCCTGCCCATGCATCATGCGTCTG 540
acta2 ..A..C..T..C..T..TT.....T... ..C.....T..T..G..CA..T..C.....C .....T..T..A.....T.....T..
novel ..A.....T.....T.....T.....T.....T..G..C..A.....T.....T..A.....A.....

acta1 GATCTGGCCGGTCGCGATCTGACTGACTAT CTCATGAAGATCCTGACTGAGAGAGGCTAC TCTTTCGTTACAACCGCTGAGCGTGAGATC 630
acta2 ...T.....C..T.....C.....C .....A.....A.....A..... ..A.....C.....T.....AA..A..A..T
novel ...T.....T.....CA..A..C.....C ..G.....A.....T.....T.....T.....C..C..T..C..A..G.....T

acta1 GTGCGGACATCAAGGAGAAGCTCTGCTAT GTGGCTCTGGACTTCGAGAACGAGATGGCC ACCGCCCTCCTCCTCCTCCCTGGAGAAG 720
acta2 ..C..T.....T.....T..G.....T..... ..C..G.....T.....T.....T..... ..A.....T.....
novel ..T..T.....T.....T..G.....C .....T..T..T.....T.....T..... ..A..T.....

acta1 AGTACGAGCTGCCGACGGTCAAGGTGATC ACCATCGGTAAACGAGCGTTTCCCGTGCCCC GAGACCCCTCTCCAGCCTTCTTCATTGGT 810
acta2 .....T.....T..T..T..A.....C... ..T..C..T..A..G.....T .....G.....A.....A.....A
novel TC.....T.....T.....C.....C..... ..A.....G.....T.....T..... ..A.....C.....T.....C.....

acta1 ATGGAGTCTGTGGCATCCATGAGACCGCC TACAACAGCATCATGAAGTGCACATTGAC ATCCGTAAGGATCTGTAGCCCAACAACGTC 900
acta2 .....A..G.....A.....A..T .....A.....A.....A..... ..T.....C.....T.....G
novel .....A.....A.....A..T .....G.....T.....T..... ..T.....T.....T.....A

acta1 TTGTCCGGTGTACCACCATGTACCCTGGT ATTGCTGACCGCATGAGAAGGAGTCACT GCTCTGGCCCCAGCACCATGAAGATCAAG 990
acta2 C...T.....T.....T..... ..A..A..A..G..... ..T.....
novel C...T.....T.....T..... ..C..G...A..G.....A.....A .....T.....T.....

acta1 ATCATTGCTCCTCCGAGCGTAAGTACTCC GTCTGGATCGGTGGCTCCATCCTGGCTTCC CTGTCCACCTCCAGCAGATGTGGATCACC 1080
acta2 ...C..C..A.....C.....T..A ..A.....T..C.....C..T .....C.....A.....A..G
novel ..G.....C.....C.....A.....T ..G.....T.....T..... ..C.....

acta1 AAGCAGGAGTACGACGAGGCGGTCCTTCC ATTGTCCACAGGAAGTGTCTTaaaccaac aacaatctctcagtcacgcgtacatttg 1170
acta2 .....A..A..... ..G..TC.C.....taaacggca ccttcaactccctcctggtcccctaattct
novel ..AG.C.....T..G.....A..C..... ..taaacacct cctctctcattctcctctctgctgactc

acta1 tactgtctcttttactgtatatacatgtaa tgttgtataaaaaacagataactgataaca gaaaaaaaaaaaaaaaaaaaaa 1250
acta2 gctattgctctcttagcctgtaaaactgtgaa catacttcctctgttttgcctttgcaactg cccacatctgtttgaatgttttgattaat 1260
novel tccatgtcacgtctcgtgtttgtaccgcag acgtttgttcaagctcccctgtacaacaa ttcaagttcacagccacaataaaattgttca

acta2 gatgtcatgacgatatagtgttatcaatgg gttaaaaagccttgaatgtgctgccaatag aaggatttgggagattctttggggtggcga 1350
novel ataaaattcaaaaaaaaaaaaaaaaaaaaaa 1286

acta2 gactccagagatgaggagaggtagaggaaac tcccttttgtgtaaaacgctacagagtaa cacaaactgcatgtcagctttcaggcagctct 1440

acta2 gtatggacatcatatctccatgaaggaata aagcctctgcaataataaaaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1530
acta2 aa 1532

```

Fig. 1. Nucleotide sequences of cDNAs encoding three *Hemibarbus mylodon* muscle actin isoforms. *Acta1*, skeletal α -actin; *acta2*, aortic α -actin; novel, novel muscle actin type I. Coding sequences are represented by upper case letters and non-coding by lower case. Dots indicate identical nucleotide bases for the coding region. The putative polyadenylation signal (AATAAA) is indicated in boldface.

tion of synonymous substitutions in the third codon positions. Robustness of tree topologies was evaluated by bootstrap analysis with 5,000 replicates (Felsenstein, 1985).

Bayesian inference (BI) analysis was carried out in MrBayes 3.1.2 (Ronquist and Huelsenbeck, 2003). The nucleotide sequence matrix was partitioned into three codon

positions for coding nucleotide sequences. Model selection strategy of Akaike Information Criterion (AIC) implemented in MrModeltest 2.2 (Nylander, 2004) was used to determine the best-fit evolutionary model of nucleotide substitutions for each partition, and the GTR+I+ Γ (Iset nst = 6 rates = invgamma) was selected for all three positions. All model parameters were

unlinked [unlink statefreq = (all) revmat = (all) shape = (all) pinvar = (all)], and all partitions were allowed to have different rates (prset ratepr = variable). Two independent Markov chains were performed with four simultaneous chains (three heated and one cold) with random starting trees for 5,000,000 generations, sampling trees at intervals of 100 generations. A total of 10,000 out of 50,001 resulting trees were discarded as "burn-in." The remaining 40,001 trees were used to construct a 50% majority-rule consensus tree and to estimate statistical supports for tree topologies determined on the basis of posterior probabilities.

Results and Discussion

Sequence characteristics of three muscle actin cDNAs

Three full-length muscle actin clones were identified from the EST database of *H. mylodon* (Fig. 1). Each was designated as *acta1*, *acta2*, or novel muscle actin type 1, of which annotation was based on the highest sequence similarity (see below), high homology in 5'- and 3'-untranslated region (UTR) sequences (data not shown), and reciprocal phylogenetic clustering to one of seven muscle actin isoforms of *Danio rerio* (see below). The cDNA sequences of *acta1*, *acta2*, and novel actin type 1 were 1,294, 1,601, and 1,329 bp long with poly (A⁺) tails of 19, 47, and 18 bp, respectively.

The long nucleotide sequence of *acta2* resulted from much longer 3'-UTR than the other two isoforms. All of the cDNAs shared the canonical poly (A⁺) signal (AATAAA) prior to the poly (A⁺) tail. Two putative poly (A⁺) signals were found in novel actin type 1. The coding nucleotide sequences showed 83.9-87.2% identity each other.

Multiple alignment of amino acid sequences

Three cDNA sequences of *H. mylodon* muscle actin isoforms identically encoded 377 amino acid proteins (Fig. 2) and contained the Met-Cys residues in the N-terminal region, which are known to undergo the posttranscriptional removal (Gunning et al., 1983; Sheff and Rubenstein, 1992). They showed 96.8-98.1% identity each other. Comparison of amino acid sequences revealed that *H. mylodon* *acta1* showed 100% identity to *D. rerio* *acta1*, and 99.7% identities to *Cyprinus carpio* and *Carassius auratus* *acta1*. *H. mylodon* *acta2* showed 99.7% identity to *D. rerio* *acta2* and 98.9-99.2% identities to tetrapod *acta2*. Novel muscle actin type 1 of *H. mylodon* showed 99.7% identities to an uncharacterized actin isoform of *D. rerio* (zgc:86725) and a slow skeletal muscle of *Salmo trutta* (AF267496). The six nonconservative substitutions in the salmonid slow skeletal actin (as defined by a change in polarity, charge, and Gly content) (Mudalige et al., 2007) were also conserved in the novel actin type 1 of



Fig. 2. Alignment of deduced amino acid sequences of three *Hemibarbus mylodon* muscle actin isoforms. Acta1, skeletal α -actin; acta2, aortic α -actin; novel, novel muscle actin type 1. Dots indicate the identical residues, and letters represent amino acids where substitutions occur. Characteristic Met-Cys residues in the N-terminal, which are known to undergo the posttranscriptional removal, are boxed. The six non-conservative substitutions (as defined by a change in polarity, charge, and Gly content) in the novel actin type 1 (Mudalige et al., 2007) are indicated by arrowheads. The numbering system is for the mature protein comprising 375 amino acid residues after excluding the first two codons.

H. mylodon (i.e., 103-Val, 155-Ala, 278-Thr, 281-Gly, 310-Gly, and 360-Asp; Fig. 2). This novel piscine muscle actin type is known to be exclusive in slow skeletal muscle in salmonids, and differs significantly from other muscle actins in terms of various biochemical properties (Mudalige et al., 2007).

Phylogenetic analysis

Phylogenetic tree of teleost muscle actins including three muscle actin isoforms of *H. mylodon* and representative bird and mammals (tetrapods) were reconstructed to understand their phylogenetic relationships using BI and NJ methods with their coding nucleotide sequences (Fig. 3). Our tree

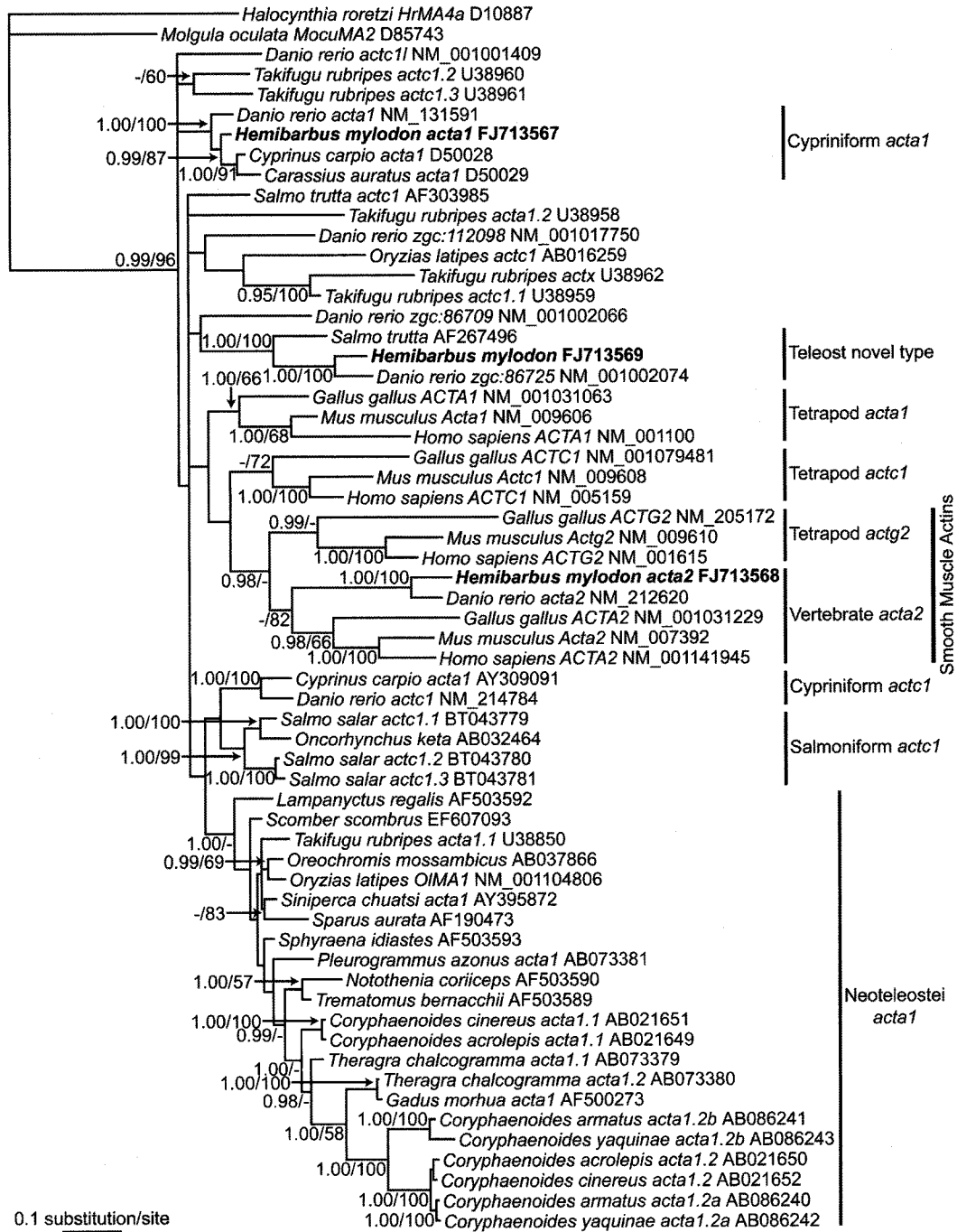


Fig. 3. Bayesian tree inferred from the coding nucleotide sequences of muscle actins of teleosts and representative vertebrate taxa. Two ascidian species were used as outgroups. Numbers above each branch node indicate posterior probability above 0.90 in Bayesian analysis and bootstrap values above 50% in neighbor-joining analysis. Taxonomic placements with actin types were indicated for major actin groups. See Table 1 for more information.

showed the monophyletic status of muscle actins from teleosts and tetrapods with 0.99 posterior probability in BI analysis and 96% bootstrap value in NJ analysis. Three muscle actin isoforms cloned from *H. mylodon* branched off in the clades, in which three of seven muscle actin isoforms from *D. rerio* [i.e., *acta1*, *actc1*, cardiac α -actin like (*actc1l*), *acta2*, and three characterized actin isoforms (*zgc:86709*, *zgc:86725*, and *zgc:112098*)] placed, with highest statistical supports. *H. mylodon acta1* placed in the clade of cypriniform *acta1* from *C. carpio*, *C. auratus*, and *D. rerio*. *H. mylodon acta2* showed the closest phylogenetic affiliation to *D. rerio acta2*, and the two cypriniform actins formed a monophyletic group with tetrapod *acta2*. Finally the third muscle actin isoform of *H. mylodon* formed a strongly supported monophyly to *D. rerio zgc:86725* and *S. trutta* slow muscle actin, which were designated as novel muscle actin type 1 in this study. Given the presence of as much as seven muscle actin isoforms in *D. rerio* (Fig. 3) we would expect the presence of additional muscle actin isoforms in *H. mylodon* when extending EST sampling.

Within the muscle actin lineage of vertebrates, two smooth muscle actins (*acta2* and *actg2*) emerged together with 0.98 posterior probability in BI analysis. The *acta2* clade composed of cypriniforms and tetrapods were phylogenetically separated from the *actg2* clade only composed of tetrapods. Two cypriniform species (*D. rerio* and *H. mylodon*) were further separated from tetrapods within the *acta2* clade. Beside this notable clade of smooth muscle actins of vertebrates, other muscle actins tended to cluster together according to a taxonomic group with a specific actin isoform rather than major actin isoforms. For example tetrapod *acta1* and *actc1*, cypriniform *acta1* and *actc1*, novel actin type 1 of teleosts, salmoniform *actc1*, and Neoteleostei *acta1* formed each monophyletic group. However there were no clear phylogenetic relationships among these actin clades. In addition *T. rubripes actc1.2* and *actc1.3*, and *T. rubripes actc1.1* and *actx* independently branched off among the above main muscle actin clades.

The actin multigene family has evolved through duplication and divergence from a common ancestral gene resulting in functionally distinct actin isoforms. Our phylogenetic analysis also showed that diverse muscle actin isoforms of vertebrates were recovered as the monophyly, implying that their actin isoforms have been derived from a single common ancestor after the emergence of vertebrates. Furthermore teleosts, in which as much as seven muscle actins isoforms

were found, appeared to have experienced more frequent duplication events than tetrapods, in which four muscle actin isoforms were only found (Vandekerckhove and Weber, 1979; Miwa et al., 1991). Our result also showed the different evolutionary history between striated and smooth muscle actins. The monophyletic status of two smooth muscle actin isoforms (*acta2* and *actg2*) was supported in BI analysis. Their monophyly was independently supported by the identical exon/intron organization (Miwa et al., 1991). These results strongly suggest that the two isoforms originated from a common ancestral gene and further diverged into two distinct isoforms after the emergence of vertebrates. In contrast to the *acta2* clade, in which cypriniforms and tetrapods consistently clustered together, the *actg2* clade consisted of only tetrapods, implying either the specific gene duplication event in the tetrapod lineage or the loss of *actg2* orthologs during the teleostean evolution. When considering the apparent monophyly of teleosts and the isolation of the nearly complete set of muscle actin isoforms from both *D. rerio* and *T. rubripes* it is interesting that no other teleostean taxa except cypriniform species showed the phylogenetic affiliation to the *acta2* clade.

Other muscle actin isoforms however appeared to have experienced much complicated evolutionary history by multiple gene duplication events. Two striated muscle actin isoforms of tetrapods (*acta1* and *actc1*) formed each monophyletic group. Most of muscle actin isoforms of teleosts phylogenetically clustered according to specific taxonomic groups, i.e., Cypriniformes, Salmoniformes, or Neoteleostei, and no notable phylogenetic groupings were identified among them. This result is congruent with the phylogenetic analysis of cytoplasmic actins, in which *actb* of cypriniform and percormorph species showed no phylogenetic relationship in spite of their apparent orthology to tetrapod *actb* inferred from the exon/intron organization and high homology in the 5'-upstream and 3'-UTR sequence (Kim et al., 2008). Muscle actins of teleosts also distinct by species- or lineage-specific gene duplications. For example, salmoniform *acta1* was duplicated into *acta1.1* and *acta1.2/acta1.3*, and the latter two isoforms were further duplicated in *Salmo salar*. *Acta1* of some gadiform species (*Coryphaenoides acrolepis*, *Coryphaenoides cinereus*, and *Theragra chalcogramma*) was also duplicated into *acta1.1* and *acta1.2*, and two abyssal *Coryphaenoides* species (*C. armatus* and *C. yaquinae*) further duplicated into *acta1.2a* and *acta1.2b*. These gene duplications were also observed in *T. rubripes* (*actc1.1* and *actx*, and

act1.2 and *act1.3*). Biological significance of the presence of multiple actin isoforms remains to be elucidated, but it appears to be a source of skeletal muscle plasticity (Mudalige et al., 2007) or to be resulted from adaptive modification to the new environment as exemplified in abyssal *C. armatus* and *C. yaquinae* (Morita, 2003). Meanwhile the strong phylogenetic clustering of the uncharacterized novel muscle actin type 1 of teleosts between the two higher taxonomic levels (i.e., Cypriniformes and Salmoniformes) was notable from other phylogenetic groups of teleost actins, implying this novel actin type is evolutionary well conserved and plays some important roles in teleosts.

In this study we isolated three muscle actin isoforms from *H. mylodon*. One of them showed the clear orthologous relationship to mammalian *acta2* in our phylogenetic analysis. However the other two isoforms did not show clear phylogenetic relationships to four mammalian counterparts and were tentatively designated as *acta1* and novel actin type 1 inferred from their phylogenetic cluserings to cypriniform skeletal α -actins and uncharacterized *D. rerio* muscle actin isoform/*S. trutta* slow muscle actin, respectively. Comparisons of the 5'- and 3'-UTR sequences to four orthologous actin genes of mammals did not reveal high degree of homology for the three muscle actins of *H. mylodon* (data not shown). In the future study comparisons of the exon/intron organization or sequences of 5'-flanking region will help to reveal their orthology.

Acknowledgement

Authors thank Dr. I.-C. Bang of Soonchunhyang University for providing the fish specimen. This study was supported by the research fund (Project no. #20088033-1) from the Ministry of Land, Transport and Maritime Affairs, the Republic of Korea.

References

- Bang, I.-C., Y. H. Lim, Y. S. Cho, S. Y. Lee and Y. K. Nam, 2007. Survey of expressed sequence tags from tissue-specific cDNA libraries in *Hemibarbus mylodon*, an endangered fish species. *J. Aquacult.*, 20, 248–254 (in Korean).
- Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, 39, 783–791.
- Gunning, P., P. Ponte, H. Okayama, J. Engel, H. Blau and L. Kedesi, 1983. Isolation and characterization of full-length cDNA clones for human α -, β -, and γ -actin mRNAs: skeletal but not cytoplasmic actins have an amino-terminal cysteine that is subsequently removed. *Mol. Cell. Biol.*, 3, 787–795.
- Hall, T. A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.*, 41, 95–98.
- Jang, M.-H., M. C. Lucas and G.-J. Joo, 2003. The fish fauna of mountain streams in South Korean national parks and its significance to conservation of regional freshwater fish biodiversity. *Biol. Conserv.*, 114, 115–126.
- Kim, I.-S., Y. Choi, C.-L. Lee, Y.-J. Lee, B.-J. Kim and J.-H. Kim, 2005. Illustrated Book of Korean Fishes. Kyo-Hak Publishing Co. Ltd., Seoul, 615 pp. (in Korean).
- Kim, W.-J., Y.-A. Lee and I.-C. Bang, 2007. Isolation and characterization of polymorphic microsatellite markers for the endangered Korean freshwater fish *Hemibarbus mylodon*. *Mol. Ecol. Notes*, 7, 516–518.
- Kim, K.-Y., S. Y. Lee, Y. S. Cho, I.-C. Bang, D. S. Kim and Y. K. Nam, 2008. Characterization and phylogeny of two β -cytoskeletal actins from *Hemibarbus mylodon* (Cyprinidae, Cypriniformes), a threatened fish species in Korea. *DNA Seq.*, 19, 87–97.
- Kim, S., A. Karsi, R. A. Dunham and Z. Liu, 2000. The skeletal muscle α -actin gene of channel catfish (*Ictalurus punctatus*) and its association with piscine specific SINE elements. *Gene*, 252, 173–181.
- Krasnov, A., H. Teerijoki, Y. Gorodilov and H. Mölsä, 2003. Cloning of rainbow trout (*Oncorhynchus mykiss*) α -actin, myosin regulatory light chain genes and the 5'-flanking region of α -tropomyosin: functional assessment of promoters. *J. Exp. Biol.*, 206, 601–608.
- Lee, Y.-A., Y.-E. Yun, Y. K. Nam and I.-C. Bang, 2008. Genetic diversity of endangered fish *Hemibarbus mylodon* (Cyprinidae) assessed by AFLP. *J. Aquacult.*, 21, 196–200.
- Miwa, T., Y. Manabe, K. Kurokawa, S. Kamada, N. Kanda, G. Bruns, H. Ueyama and T. Kakunaga, 1991. Structure, chromosome location, and expression of the human smooth muscle (enteric type) γ -actin gene: evolution of six human actin genes. *Mol. Cell. Biol.*, 11, 3296–3306.
- Morita, T., 2000. Amino acid sequences of α -skeletal muscle actin isoforms in two species of rattail fish, *Coryphaenoides acrolepis* and *Coryphaenoides cinereus*. *Fish. Sci.*, 66, 1150–1157.
- Morita, T., 2003. Structure-based analysis of high pressure adaptation of α -actin. *J. Biol. Chem.*, 278, 28060–28066.
- Moutou, K. A., S. Socorro, D. M. Power, Z. Mamuris and A. V. M. Canario, 2001. Molecular cloning and sequence of gilthead sea bream (*Sparus aurata*) α -skeletal actin: tissue and developmental expression. *Comp. Biochem. Physiol. B*, 130, 13–21.
- Mudalige, W. A. K. A., D. M. Jackman, D. M. Waddleton and D. H. Heeley, 2007. A vertebrate slow skeletal muscle actin isoform. *FEBS J.*, 274, 3452–3461.
- Nylander, J. A. A., 2004. MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.
- Pollard, T. D. and J. A. Cooper, 1986. Actin and actin-binding proteins: a critical evaluation of mechanisms and functions.

- Annu. Rev. Biochem., 55, 987–1035.
- Ronquist, F. and J. P. Huelsenbeck, 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, 19, 1572–1574.
- Sheff, D. R. and P. A. Rubenstein, 1992. Isolation and characterisation of the rat liver actin N-acetylaminopeptidase. *J. Biol. Chem.*, 267, 20217–20224.
- Swofford, D. L., 2002. PAUP: Phylogenetic Analysis Using Parsimony (and Other Methods). Ver 4.0b10, Sinauer Associates, Sunderland.
- Tanaka, H., Y. Maezawa, T. Ojima and K. Nishita, 2004. Cloning and sequencing of cDNAs encoding walleye pollack α -skeletal actin isoforms. *Fish. Sci.*, 70, 198–200.
- Vandekerckhove, J. and K. Weber, 1979. The complete amino acid sequence of actins from bovine aorta, bovine heart, bovine fast skeletal muscle, and rabbit slow skeletal muscle: a protein-chemical analysis of muscle actin differentiation. *Differentiation*, 14, 123–133.
- Venkatesh, B., B. H. Tay, G. Elgar and S. Brenner, 1996. Isolation, characterization and evolution of nine pufferfish (*Fugu rubripes*) actin genes. *J. Mol. Biol.*, 259, 655–665.
- Watabe, S., Y. Hirayama, J. Imai, K. Kikuchi and M. Yamashita, 1995. Sequences of cDNA clones encoding α -actin of carp and goldfish skeletal muscles. *Fish. Sci.*, 61, 998–1003.
- Xu, Y., J. He, X. Wang, T. M. Lim and Z. Gong, 2000. Asynchronous activation of 10 muscle-specific protein (MSP) genes during zebrafish somitogenesis. *Dev. Dyn.*, 219, 201–215.

논문접수 : 2009년 1월 28일

심사완료 : 2009년 2월 6일

수정본 수리 : 2009년 2월 10일