Evolution of Sarcomeric Myosin Heavy Chain Genes: Evidence from Fish

Katrina McGuigan,† Patrick C. Phillips,† and John H. Postlethwait**

*Institute of Neurosciences and †Center for Ecology and Evolutionary Biology, University of Oregon

Myosin heavy chain (MYH) is a major structural protein, integral to the function of sarcomeric muscles. We investigated both exon-intron organization and amino acid sequence of sarcomeric MYH genes to infer their evolutionary history in vertebrates. Our results were consistent with the hypothesis that a multigene family encoded MYH proteins in the ancestral chordate, one gene ancestral to human *MYH16* and its homologues and another ancestral to all other vertebrate sarcomeric MYH genes. We identified teleost homologues of mammalian skeletal and cardiac MYH genes, indicating that the ancestors of those genes were present before the divergence of actinopterygians and sarcopterygians. Indeed, the ancestral skeletal genes probably duplicated at least once before the divergence of teleosts and tetrapods. Fish homologues of mammalian skeletal MYH are expressed in skeletal tissue and homologues of mammalian cardiac genes are expressed in the heart but, unlike mammals, there is overlap between these expression domains. Our analyses inferred two other ancestral vertebrate MYH genes, giving rise to human *MYH14* and *MYH15* and their homologues. Relative to the skeletal and cardiac genes, *MYH14* and *MYH15* homologues are characterized by evolution of intron position, differences in evolutionary rate between the functionally differentiated head and rod of the myosin protein, and possible evolution of function among vertebrate classes. Tandem duplication and gene conversion appear to have played major roles in the evolution of at least cardiac and skeletal MYH genes in fish. One outcome of this high level of concerted evolution is that different fish taxa have different suites of MYH genes, i.e., true orthologs do not exist.

Introduction

Myosins are a family of protein motors that effect many types of eukaryotic motility, from organelles, to single cells and whole organisms. Perhaps the best studied myosins are Class II (conventional) myosins, which include nonmuscle as well as smooth muscle and sarcomeric muscle myosins. The contractile and bioenergetic requirements of sarcomeric muscles (vertebrate skeletal and cardiac) vary among organisms as well as spatiotemporally within organisms. These variable demands are met by using different myosin proteins (reviewed by Schiaffino and Reggiani 1996). Sarcomeric myosin is an asymmetrically hexameric protein, consisting of four myosin light chains (MLC) and two myosin heavy chains (MYH). Variation in regulation and amino acid sequence among members of the sarcomeric MYH multigene family contributes strongly to the variation in muscle properties that is important to performance (Schiaffino and Reggiani 1996).

Class II myosin genes are ancient, represented in all eukaryotic kingdoms except plants (Berg, Powell, and Cheney 2001). MYH II genes appear to have evolved at a relatively homogeneous rate across bilaterians and have a strong phylogenetic signal (Ruiz-Trillo et al. 2002), suggesting it is possible to accurately infer the evolutionary history of these functionally important proteins. The ancestral deuterostome probably had two sarcomeric MYH genes, one of which might have been lost in the vertebrate lineage (Chiba et al. 2003). Following divergence of vertebrates from urochordates, further duplication led to at least five sarcomeric MYH genes in the last common ancestor of vertebrates (Desjardins et al. 2002). It is unclear whether these all descended from a single gene, as suggested by Chiba et al. (2003). Three of the putative five ancestral vertebrate sarcomeric MYH genes were inferred from humans, and no orthologs have yet been identified from other vertebrates (Desjardins et al. 2002). Much of our understanding of myosin function comes from studies of skeletal (fast) and cardiac (slow) MYHs. The last common ancestor of these two genes probably duplicated prior to the diversification of the vertebrate classes (e.g., Moore et al. 1993; Desjardins et al. 2002). The ancestral cardiac and skeletal genes appear to have duplicated and diversified independently within each vertebrate class (Moore et al. 1993; Gauthier et al. 2000; Desjardins et al. 2002).

Although vertebrates diversified independently, the same putative process (unequal crossing over) appears to have generated cardiac and skeletal duplicates in different vertebrates. Genes of both classes are tandemly arrayed in mammals (Leinwand et al. 1983; Mahdavi, Chambers, and Nadal-Ginard 1984; Matsuoka et al. 1989; Qin et al. 1990; Weiss et al. 1999) and chickens (Gulick, Kropp, and Robbins 1987; Chen et al. 1997). Recombination has occurred between nonallelic genes of a tandem array (i.e., gene conversion), although not at a uniform rate across genes or taxa (Moore et al. 1992; Epp et al. 1995; Desjardins et al. 2002). Generally, gene conversion decreases apparent time, since duplication within groups of recombining tandem duplicates and increases apparent divergence time between groups of nonrecombining genes (see Moore et al. 1992). Inference of cardiac and skeletal gene evolution is therefore confounded by concerted evolution (Moore et al. 1992; Desjardins et al. 2002).

In comparison to other vertebrates, our knowledge of the MYH gene family in teleosts is limited. MYH genes have been sequenced from fish transcriptomes, although available phylogenetic information suggests all are skeletal genes (e.g., Kikuchi et al. 1999; Gauthier et al. 2000; Gauvry et al. 2000). Because of their basal divergence within vertebrates, knowledge of the evolutionary origin of MYH genes in fish will contribute both to our reconstruction of the vertebrate ancestral state and to our understanding of the processes driving evolution of the gene family.

Key words: gene conversion, gene family, myosin, muscle, Teleost, *Gasterosteus aculeatus*.

E-mail: mcguigan@darkwing.uoregon.edu

Mol. Biol. Evol. 21(6):1042–1056. 2004
DOI:10.1093/molbev/msh103
Advance Access publication March 10, 2004

 © Society for Molecular Biology and Evolution 2004; all rights reserved.

Table 1
Primers Used to Screen the *G. aculeatus* Fosmid Library and cDNA Pools

Primer[a]	Sequence (5′ to 3′)	Exon[b]	Position[b]
F1	TCCGGAAGCCAGAGAAGGAGAG	3	50
R1	AGTGAAGCCCAGGATATCAATAGCA	12	1002
F3	GAAGCAGCGTGAGGAGCAGGC	12	1101
R3	CCCTCCAGCACACCATTACACCT	19	2083
F5	ATCTGYAGRAARGGNTTYCC	19	2113
R5	TCMAGCCKCTCYCTGATCTCCTC	27	3433
F4	TGCTCGGGCCAAGGTTGAGAA	27	3384
R4	GGTTGGCAGCCAGAGAATTAG	31	4307
F2	GGGTCAGGCAGAGCTGGAAGGA	32	4389
R2	CACGTCTCTGCTCAGCTTCAAC	38	5479

[a] Primers designed from *F. rubripes* (v.1. T007448 and v.1. T004025), except F5 and R5 (from Gac 1 and Gac 2: table 2).

[b] Primer start position (bp and Exon) determined from human *MYH3* cDNA (X13988), with the first nucleotide of the translation start codon (Exon 3) designated position 1.

In addition to their potential contribution to understanding the evolution of the MYH gene family, several unique features of sarcomeric muscle in teleosts make them good systems for studying muscle functional evolution. First, in contrast to other vertebrates, muscle fibers with different contractile properties are spatially separated in teleosts. White, fast-twitch muscles make up the bulk of the fish's body whereas red, slow-twitch fibers are found in a narrow midlateral band just under the skin (reviewed by Sänger and Stoiber 2001). When fish swim by oscillation of the pectoral fin, rather than body-caudal oscillation, the pectoral fin also tends to be primarily red muscle (e.g., Walesby and Johnston 1980). Different fiber types express different MYH (e.g., Karasinski and Kilarski 1989; Martinez et al. 1989, 1991), although this interpretation is based on electrophoretic or histoimmunological evidence, not sequence data (but see Rescan et al. 2001). Second, fish muscles grow throughout life by hyperplasia, not just hypertrophy as in tetrapods (reviewed by Sänger and Stoiber 2001). Third, fish muscle is highly plastic, altering MYH isoform in response to many environmental factors, including activity level and temperature (reviewed by Sänger and Stoiber 2001). Fourth, relative to tetrapods, teleosts are thought to have undergone an additional round of genome duplication (Amores et al. 1998; Postlethwait et al. 1998; Taylor et al. 2001) and might therefore have more MYH genes than tetrapods. For example, only seven skeletal MYH genes have been identified in humans (see Desjardins et al. 2002), but 29 have been described from carp (*Cyprinus carpio*) (Gerlach et al. 1990; Kikuchi et al. 1999). Natural variation in expression of duplicate MYH genes is a powerful tool for exploring the relationship between bioenergetic/contractile properties of muscles and amino acid sequence of MYH. Developing a phylogenetic framework within which researchers can explore those relationships will expand the utility of fish systems as models for the evolution of muscle function.

In this article, we address two broad questions. (1) What is the phylogenetic origin of MYH genes in teleosts? Specifically, does evidence from fish support the inference of six MYH genes in the ancestral vertebrate? Does evidence from fish support the inference of ancestral cardiac and skeletal genes that diversified independently within vertebrate classes? Is the cardiac-versus-skeletal expression of these genes ancestral or derived? (2) Have teleost MYH genes evolved through the same processes as MYH genes in other vertebrates? Specifically, have teleost genes been tandemly duplicated by unequal crossing over and, if so, has gene conversion played a role in the evolution of duplicates? We address these questions by analyzing the exon-intron organization of MYH genes and by phylogenetic reconstruction based on protein sequence. In addition to using published cDNA sequence of MYH genes, we sequenced MYH genes from threespine stickleback (*Gasterosteus aculeatus*) and identified MYH genes from the genome databases of pufferfish (*Fugu rubripes*) (from the Medical Research Council's Human Genome Mapping Project Resource Center) and zebrafish (*Danio rerio*) (from the *D. rerio* Sequencing Group at the Sanger Institute).

Methods

Data Collection

A Fosmid library from *G. aculeatus* (Cresko *et al.* 2003) was screened for MYH genes using several primers (table 1). Two unique clones (Gac 1 and Gac 2: table 2) contained MYH genes, and their DNA was sheared and subcloned for sequencing (Cresko *et al.* 2003). Sub-clones were arrayed into 96-well plates as single clones and screened via PCR, using the same primers as for the original screen (table 1). All positive clones were sequenced. Clones overlapped, such that the final sequence for a Fosmid clone was the consensus of multiple overlapping subclone sequences.

cDNA pools were constructed from heart, pectoral (red, slow-twitch) and axial (white, fast-twitch) muscles of *G. aculeatus*. Pectoral (red) muscle was collected from fish from the Mill Race, Eugene, Ore. Heart and axial (white) muscle was from two fish from each of three sites in south central Alaska (Bear Paw Lake, Boot Lake, and Rabbit Slough). RNA was extracted using TRIZOL (Life Technologies, Cat. No. 10296) and messenger purified using Oligotex (Qiagen Cat. No. 70022). cDNA pools were amplified using Smart cDNA Library Construction Kit (Clontech Laboratories Inc., Palo Alto, Calif., protocol number PT300-2, Cat. No. K1052-1) and probed with MYH primers (table 1). PCR products were cloned (Invitrogen TOPO TA Cloning Kit Cat. No. K450001) and both strands were sequenced. cDNA sequences were identified with a letter code corresponding to the tissue from which they were amplified exclusively, or were the most common PCR product: white muscle (W), red muscle (R), and cardiac muscle (C) (table 3).

The *F. rubripes* genome (Medical Research Council's Human Genome Mapping Project Resource Center, London, Releases 2 and 3: http://fugu.hgmp.mrc.ac.uk/blast) and the *D. rerio* genome (the *D. rerio* Sequencing Group at the Sanger Institute: www.sanger.ac.uk/Projects/D_rerio) were searched for putative MYH genes using the Basic Local Alignment Search Tool (BLAST), available on each genome's website. BLAST was initiated with human MYH genes (Desjardins et al. 2002), and each MYH

Table 2
MYH Genes Identified from Genomic Sequence[a]

Accession No.	Taxon	Group[b]	Clade[c]	Code[d]	Missing[e]	Length[f]
AL954717	*D. rerio*	A	C	Dre A1	—	15721 (37)
AL954717	*D. rerio*	A	C	Dre A2	—	15673 (37)
AL954717	*D. rerio*	A	C	Dre A3	—	12459 (37)
AL954717	*D. rerio*	A	C	Dre A4	—	18775 (37)
AL954717	*D. rerio*	A	C	Dre A5	—	11210 (36)
ctg27	*D. rerio*	A	C	—	3–6, 41	—
NA20616	*D. rerio*	A	C	—	3–19, 32–41	—
ctg10439	*D. rerio*	A	C	—	3–26	—
ctg11069.2	*D. rerio*	B	S	Dre B1	29–33	—
ctg11069.2	*D. rerio*	B	S	Dre B2	33–41	—
ctg11069.2	*D. rerio*	B	S	Dre B3	16–24	—
ctg654	*D. rerio*	C	C	Dre C1	23–41	—
ctg654	*D. rerio*	C	C	Dre C2	—	12912 (37)
NA13091	*D. rerio*	D	S	Dre D1	—	12769 (39)
NA10563	*D. rerio*	D	S	Dre D2	28–41	—
ctg24830	*D. rerio*	—	S	Dre Q	—	9955 (39)
ctg12961	*D. rerio*	—	C*	Dre X	23–25	—
ctg9337	*D. rerio*	—	U	Dre Y	15–16, 24–29	—
ctg26810	*D. rerio*	—	U	Dre Z	14, 26–30	—
v.3. M001034	*F. rubripes*	E	S	Fru E1	—	9731 (39)
v.3. M002528[g]	*F. rubripes*	E	S	Fru E2	3–16	—
v.3. M002528	*F. rubripes*	E	S	Fru E3	—	9727 (39)
v.2. S004378	*F. rubripes*	F	S	Fru F1	3–7	—
v.2. S004378[g]	*F. rubripes*	F	S	Fru F2	3–21	—
v.2. S004378	*F. rubripes*	F	S	—	15–41	—
v.3. M006536	*F. rubripes*	F	S	—	17–41	—
v.3. M000939	*F. rubripes*	F	S	—	12–41	—
v.2. S006033	*F. rubripes*	G	C	Fru G1	18–19, 32–34, 37–41	—
v.3. M002126	*F. rubripes*	G	C	Fru G2	3	—
v.3. M002126	*F. rubripes*	G	C	—	24–41	—
v.3. M000454	*F. rubripes*	—	S	Fru V	—	9465 (39)
v.3. M000743	*F. rubripes*	—	S	Fru W	—	9466 (39)
v.3. M000880	*F. rubripes*	—	C*	Fru X	—	5799 (1)
v.3. M003383	*F. rubripes*	—	U*	Fru Y	31–41	—
v.3. M000005	*F. rubripes*	—	U	Fru Z	—	11893 (42)
AY525070, AY525071	*G. aculeatus*	—	S	Gac W	24–27	—
AY525072, AY525073, AY525074	*G. aculeatus*	—	S	Gac R	3–11, 22–23, 28–29	—
Scaffold_878	*C. intestinalis*	—	—	Cin 2	25–41	—
Scaffold_869	*C. intestinalis*	—	—	Cin 3	—	17327 (41)
Scaffold_539	*C. intestinalis*	—	—	Cin 4	23–28	—
Scaffold_240	*C. intestinalis*	—	—	Cin 5	—	8313 (36)
Scaffold_230	*C. intestinalis*	—	—	Cin 6	—	8607 (36)

[a] Summary of supplementary table 1.
[b] hypothesized tandemly arranged genes, based on presence in a single clone or ≥95% sequence similarity across at least one, but not all, introns.
[c] Genes with similar exon organization and in same clade in figure 1.
[d] Code used to represent the sequence in figures (sequences without a code were not included in phylogenetic analyses).
[e] Exons for which gene was missing sequence data.
[f] Number of nucleotides (and exons) in gene, calculated from translation start to stop (complete genes only).
[g] In rod phylogeny only.

sequence identified was BLASTed against each database until no novel sequences resulted. Several sequences did not represent complete MYH genes (table 2); they were included here solely for the purpose of cataloguing MYH genes (sequences without a code in table 2).

A primary aim of this article was to determine the evolutionary relationships among teleost MYH genes as well as the relationships of teleost genes to MYH genes of other vertebrates. There was insufficient taxonomic coverage to achieve this aim solely by analysis of genomic sequence. Therefore, we also considered cDNA sequence. By BLASTing (blastx) against The National Center for Biotechnology Information (NCBI: http://www.ncbi.nlm.nih.gov/BLAST/) hosted databases we obtained sarcomeric MYH cDNA sequences for mammals (*Homo sapiens* and *Felis catus*), a bird (*Gallus gallus*), an amphibian (*Rana catesbeiana*), and a mollusc (*Loligo pealei*). We also considered inferred cDNA sequence of sarcomeric MYH from the ascidian *Ciona intestinalis* (Chiba *et al*. 2003; table 2). Twenty-one teleost MYH genes were identified by the BLAST search (table 3). Although selected serendipitously based on availability of data, the teleost taxa included in our analyses are from a range of evolutionary lineages.

D. rerio genomic sequences were BLASTed against Expressed Sequence Tag (EST) databases (http://www.ncbi.nlm.nih.gov/genome/seq/DrBlast.html) to identify tissue and stage-specific cDNA libraries that would

Table 3
Teleost cDNA MYH Sequences

Taxa	Tissue[a]	L[b]	Accession	Code[c]
Cyprinus carpio	White skeletal (10°C)^#1	F	D89990	Cca W10
C. carpio	White skeletal (20°C)^#1	F	D89991	Cca W20
C. carpio	White skeletal (30°C)^#1	F	D89992	Cca W30
C. carpio	White skeletal (adult growth)^*2	P	Z37108	Cca W
D. rerio	White skeletal (embryonic)^*3	P	AF180893	Dre W1
Danio rerio	White skeletal (embryonic)^*4	F	AF165817	Dre W2
D. rerio	Ventrical^*5	P	AF114427	Dre V
Gasterosteus aculeatus	White skeletal; pectoral; heart^6	P	AY525077	Gac W
G. aculeatus	Pectoral; heart^6	P	AY525076	Gac R
G. aculeatus	Heart^6	P	AY525075	Gac C
Notothenia coriiceps	White skeletal^#*7	F	AJ243767	Nco W
N. coriiceps	Red skeletal^#*7	F	AJ243769	Nco R
Oncohrynchus keta	White skeletal^8	F	AB076182	Oke W
O. mykiss	Red skeletal^*9	P	AF211172	Omy R1
O. mykiss	White and red skeletal; ventrical^#10,*9	P	Z48794	Omy W
O. mykiss	White and red skeletal^#11	P	AY009125	Omy R2
O. mykiss	Ventricle; white and red skeletal^#11	P	AY009126	Omy V
Paracirrhites fosteri	White skeletal^#*7	F	AJ243770	Pfo W
Pennahia argentata	White skeletal^12	F	AB039672	Par W
Seriola dumerili	White skeletal^13	F	AB032020	Sdu W
Theragra chalcogramma	White skeletal^14	F	AB017819	Tch W

[a] Tissue from which the gene was amplified (^), shown to be expressed by Northern blot or RT-PCR (#), or by *in situ* hybridization (*): [1] Hirayama and Watabe 1997; [2] Ennion et al. 1995; [3] Peng et al. 2002 (*myhz1*); [4] Xu et al. 2000 (*myhz2*); [5] Yelon, Horne, and Stainier 1999; [6] This study (AY525077 appears to be the product of Gac 1, AY525076 appears to be the product of Gac 2); [7] Gauvry et al. 2000; [8] Iwami et al. 2002; [9] Rescan et al. 2001; [10] Gauvry and Fauconneau 1996; [11] Weaver, Stauffer, and Coughlin 2001; [12] Yoon et al. 2000; [13] Kawabata et al. 2000; [14] Ojima et al. 1998.

[b] Genes were either full length (F) and included in main phylogenetic analysis (fig. 1*a*, *b*), or only partial (P) and included only in the phylogeny of genes of known expression (fig. 1*c*).

[c] Abbreviation used for these sequences in figure 1.

indicate spatiotemporal expression characteristics of genes identified from the genomic screen. The coding domain of MYH genes is highly conserved, so we conservatively proposed the identities of EST and the gene if the 5′ or 3′ UTR, or the hypervariable Loops 1 and 2 (defined as in Goodson, Warrick, and Spudich 1999), were >95% similar. ESTs were sought with >95% similarity to other regions of the gene only when no matches were found using the foregoing criteria.

Phylogenetic Analyses and Gene Expression

All phylogenetic analyses were conducted on amino acid sequences. MYH proteins are functionally and structurally divided into a carboxy-terminal alpha-helical rod and an amino-terminal globular head (Lowey et al. 1969; Weeds and Pope 1977). The two regions are delineated by a conserved Proline residue (838 of human *MYH3*). Rod and head play different roles in myosin function (e.g., Szentkiralyi 1984; Mitchell, Jakes, and Kendrick-Jones 1986; Toyoshima et al. 1987; Cripps, Suggs, and Bernstein 1999), differ in amino acid composition (Strehler et al. 1986), and do not give identical phylogenies (e.g., Korn 2000; Schachat and Briggs 2002). Within the head there are two hypervariable regions (Loop1 and Loop2: delineated as in Goodson, Warrick, and Spudich 1999), which were difficult to align and were not included in our analyses. The 3′ end of the gene (delineated as in Desjardins et al. 2002) was also hypervariable, difficult to align, and was not included in analyses.

We conducted separate parsimony analyses on the head (up to Proline, without Loop1 or 2) and rod (from Proline to 1091). Heuristic searches were conducted using PAUP* (Swofford 1998) following removal of uninformative characters. Robustness of the topology was assessed using 1000 bootstrap replicates with 10 replicates of random addition of sequence. Branch rearrangements were by tree-bisection rejoining. The molluscan sequence (Lpe) was used to outgroup root both phylogenies. *R. catesbeiana* sequences and two *F. rubripes* sequences (Fru E2 and Fru F2: table 2) were excluded from the reconstruction based on the head due to insufficient sequence.

Several papers report expression of specific MYH genes in teleosts (table 3). However, sequences associated with these publications are usually short (several hundred base pairs of a possible 6 kb), so we did not included them in the above analyses. Conversely, although we have complete sequence information for the genes included in head and rod phylogenies, we do not know where/when they are expressed. To generate hypotheses about the phylogenetic organization of function, we need to determine the relationships among genes whose phylogenetic position we can confidently infer, and genes of whose function is known. To this end, we conducted another phylogenetic analysis of rod sequence, with 1,000 bootstrap replicates and 10 random addition replicates. We constructed the phylogeny using all genes for which expression information was available (and sequence data for at least the last three exons: table 3), as well as *G. aculeatus*, *D. rerio*, and *F. rubripes* genes identified in the head or rod phylogenies as skeletal or cardiac (table 2).

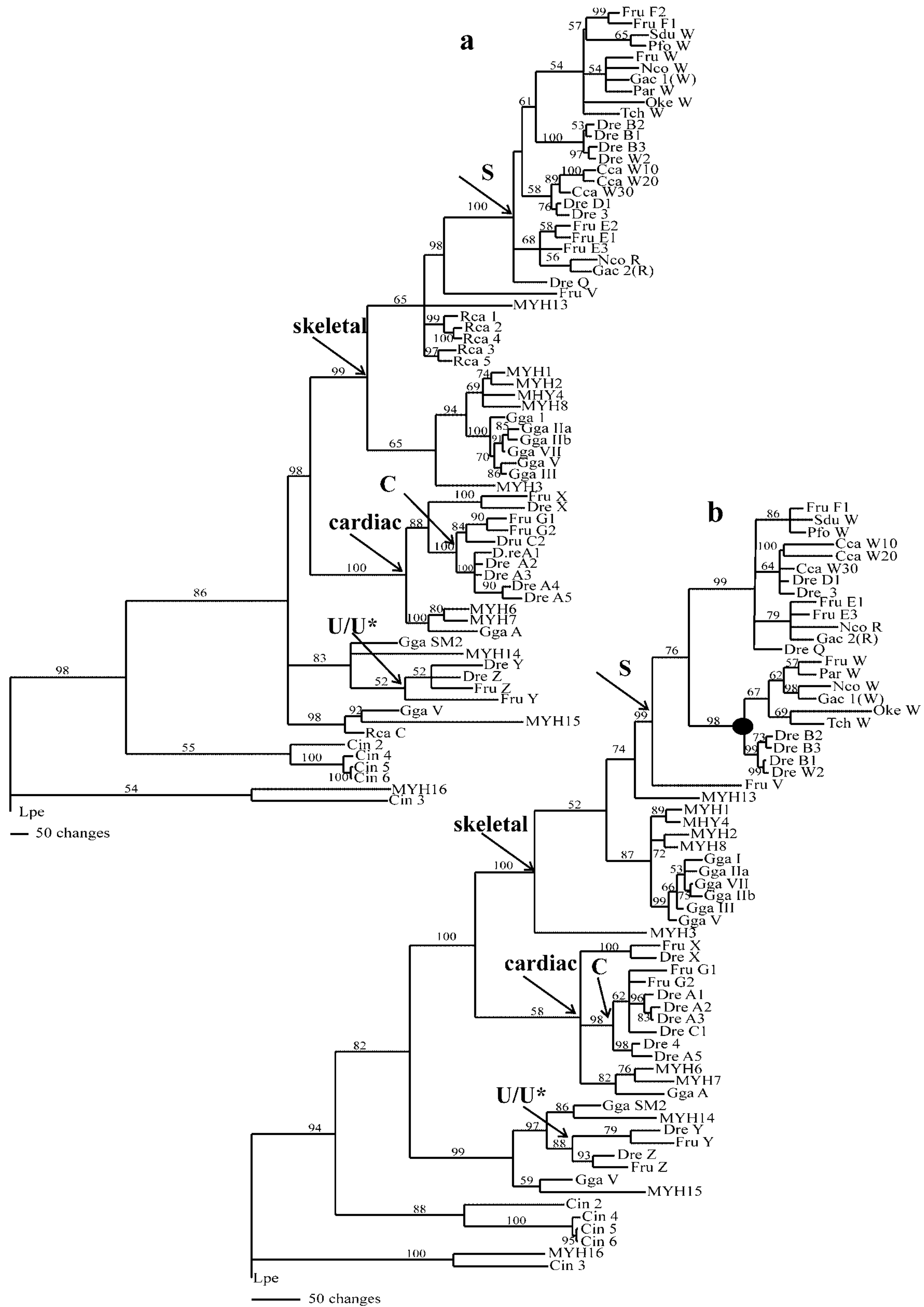

FIG. 1.—Bootstrapped maximum parsimony phylogeny of MYH genes. *H. sapiens*: *MYH1* (adult fast) (AF111785); *MYH2* (adult fast) (AF111784); *MYH3* (embryonic) (X13988); *MYH4* (adult fast) (AF111783); *MYH6* (α-cardiac) (M21664); *MYH7* (β-cardiac) (NM_000257); *MYH8* (perinatal) (Z38133); *MYH13* (extraocular) (AF111782); *MYH14* (AB040945); *MYH15* (AB023217); *F. catus*: *MYH16* (U51472); *G. gallus*: Gga III (fast) (AY116218); Gga V (neonatal) (AB021180); Gga IIa (fast embryonic) (AF272033); Gga VII (fast embryonic) (AF272034); Gga IIb (fast embryonic) (AY116217); Gga I (adult) (U87231); Gga V (ventricular) (AB032197); Gga SM2 (slow) (AB057661); Gga A (atrial) (AB004801); *R. catesbeiana: Rca C (cardiac) (AF265355)*; Rca 1 (fast skeletal) (AF097904); Rca 2 (fast skeletal) (AF097905); Rca 3 (fast skeletal) (AF097906);

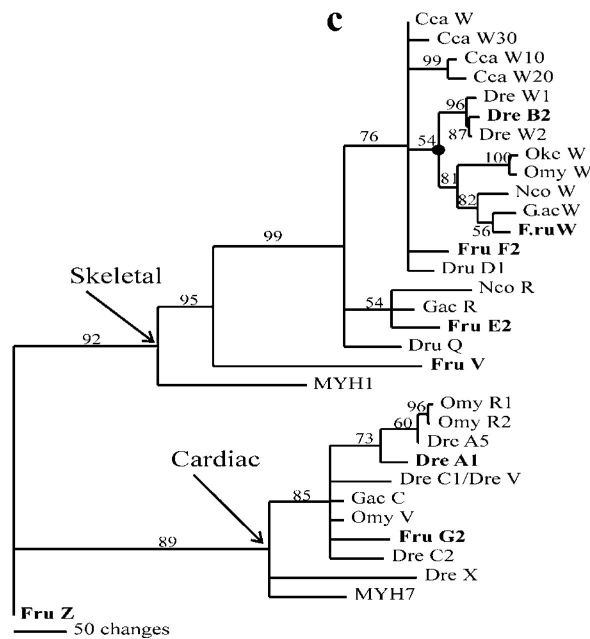

FIG. 1 *(Continued)*

Two human genes (*MYH*1 and *MYH7*) were included to root internal nodes, and Fru Z (table 2) was used to outgroup root the tree.

Exon-Intron Organization and Gene Conversion

We characterized the exon-intron organization of teleost MYH genes to provide evidence of the evolutionary history of MYH genes additional to the data from phylogenetic reconstructions. Sequences were aligned manually, or using Clustal W (Thompson, Higgins, and Gibson 1994) followed by visual assessment of the alignment. Human *MYH3* genomic sequence (AC002347) was used as the reference for comparing intron/exon boundary position, and all exons were numbered with reference to this gene (Strehler et al. 1986). Human sarcomeric MYH genes have two untranslated 5′ exons, with the translation initiation site in the third exon (Strehler et al. 1986; Desjardins et al. 2002). We began alignment at the translation start site and designated the stop codon as the end of the gene. We used Genscan (Burge and Karlin 1997: http://genes.mit.edu/GENSCAN.html/) to identify exon/intron boundaries in fish MYH genes and confirmed these results by visually screening the sequence alignments for the intron splice consensus in regions flanking the exons. We used all available cDNA evidence (including ESTs) to test our predictions of exon position. Gene length was estimated from the first position of the start codon to the third position of the stop codon, and intergene distances of tandemly arranged genes were estimated from the first nucleotide following the stop codon of one gene to the last nucleotide before the start codon of the adjacent gene. Thus, gene length is underestimated, but intergene distance is overestimated due to our treatment of untranslated regions (UTRs).

The exon-intron organization has been described for human MYH genes (Desjardins et al. 2002; Schachat and Briggs 2002) and inferred for ascidian MYH genes (see Chiba et al. 2003). We qualitatively compared the exon-intron organization of fish genes to that of human and ascidian genes. We also constructed a maximum parsimony phylogeny (PAUP* (Swofford 1998) 500 bootstraps with 10 random addition replicates) to infer the evolution of intron position.

In this study, we were specifically interested in whether processes driving evolution of teleost MYH genes were the same as operating in other taxa. Specifically, are teleost MYH genes arrayed tandemly and has gene conversion played a role in the evolution of teleost MYH genes? These questions are interrelated because gene conversion occurs via nonreciprocal genetic exchange among tandemly arrayed paralogs. In contrast to other processes through which sequence similarity across taxa or genes might be generated (e.g., selection or mutation), gene conversion has a similar effect on introns and exons but a heterogeneous effect across the gene, resulting in regions of high similarity coupled with regions of high divergence. We did not attempt a comprehensive analysis of gene conversion across all genes identified. Instead, we focused on one *D. rerio* clone (AL954717) with five MYH genes (Dre A1 to Dre A5: table 2). We followed the approach of Sawyer (1999; http://www.math.wustl.edu/~sawyer/geneconv/) to statistically test for evidence of gene conversion in exons. We conservatively identified converted fragments using the Bonferroni-corrected Karlin-Altschul *P* values, and did not permit any nucleotide mismatches in the putative converted region (see also Drouin 2002).

Results

Skeletal Genes

Based on amino acid characteristics, the majority of fish sequences formed a monophyletic clade most closely related to skeletal MYH genes from frogs, humans and chickens (fig. 1*a, b*). Genes in this clade are expressed in skeletal muscle, but cardiac expression is also observed (tables 3 and 4). Sequences from red (slow) muscle are in the teleost skeletal clade (Gac 2/Gac R and Nco R: table 3; fig. 1), which is in contrast to mammalian pattern of cardiac MYH gene expression in slow-twitch fibers.

←

Rca 3 (fast skeletal) (AF097907); Rca 5 (fast skeletal) (AF097908) *C. intestinalis*: Cin2 (adult heart) (grail.878.1.1); Cin3 (adult) (grail. 869.1.1); Cin4 (embryonic/larval) (grail.539.12.1); Cin5 (embryonic/larval) (grail.240.5.1); Cin6 (embryonic/larval) (grail.230.11.1); *Loligo pealei*: Lpe (AF042349); teleost sequences as in tables 2 and 3, plus *D. rerio* D.re3 (from whole adult library) (BC046881). *a*, reconstruction based on rod sequence. *b*, reconstruction based on head sequence. The closed circle ● demarcates the clade we hypothesize to be the fastest contracting of fish skeletal myosins. *c*, relationships among teleost skeletal and cardiac MYH genes (reconstruction based on rod sequence). Genes with known expression patterns (tables 3 and 4) in plain text and sequences for which expression is unknown in bold.

Six teleost sequences (full-length sequences in Clade S: table 2) each had 38 exons after the translation start site (Exons 3–41: fig. 2*a*), the same pattern of exon-intron organization as human skeletal genes *MYH3* and *MYH13*. A further 12 partial sequences had an exon organization consistent with this pattern (partial sequences in Clade S: table 2). *MYH3* and *MYH13* differ from other human skeletal genes (*MYH1*, *MYH2*, *MYH4*, and *MYH8*) only in having an intron interrupting Exon 40 (Schachat and Briggs 2002). All fish genes had Intron 40, suggesting its absence in the subset of human skeletal genes is a derived condition. The intron phase of fish skeletal sequences was identical to human skeletal genes (compare fig. 2 to Schachat and Briggs 2002). Fish MYH genes had shorter introns, and were therefore shorter overall than human skeletal genes (compare 9,465–12,769 bp (table 2) with 23,371 (*MYH4*)–63,580 bp (*MYH13*) (Desjardins et al. 2002)), and chicken embryonic MYH (~23 kb: Molina et al. 1987). There was slight variation in the length of Exons 3, 7, 16, 17, 40, and 41 in both human and fish skeletal genes. This variation was the result of gain/loss of codons, usually on the 3′ end of the exon; no change of intron phase was observed.

Overall, data from fish was consistent with the hypothesis that the common ancestor of vertebrates had a differentiated skeletal MYH gene and that the descendants of this gene (including MYH genes in fish) have retained the same exon-intron organization and (non-exclusive) skeletal expression. However, our phylogenetic analyses did not support the hypothesis of no duplication of the skeletal MYH gene prior to divergence of vertebrate classes. Human *MYH13* was most closely related to fish and frog genes (fig. 1*a* and *b*), suggesting it arose through a gene duplication in the last common ancestor of actinopterygians and sarcopterygians (see also Desjardins et al. 2002). In addition, *MYH3* appeared older than the divergence of Mammalia and Aves (fig. 1*a, b*: see also Moore et al. 1993), suggesting duplication prior to the divergence of mammals and birds. Frog skeletal MYH genes are a sister clade to teleost genes, rather than to a combined Mammalian and Aves clade (fig. 1*a*) as predicted from relationships among vertebrate classes. These patterns are consistent with several alternative evolutionary hypotheses, including convergent selection, concerted evolution (gene conversion) and lineage sorting (extinction of particular ancestral genes in one descendant taxa but not another).

Cardiac Genes

The second largest clade of teleost sequences is sister to the cardiac genes of humans and chickens (Clade C in fig. 1 and table 2). Human cardiac MYH genes are expressed in the heart and in slow twitch skeletal fibers, as were fish cardiac genes (although expression in fast-twitch muscle could not be discounted because some EST matches were for whole fish) (fig. 1*c*; tables 3 and 4). The six full-length fish cardiac sequences had 36 or 37 of the 39 exons observed in skeletal genes; a further eight sequences were consistent with this pattern over their available length (table 2; fig. 2*b*). Fish cardiac genes are missing Introns 18 and 36 and Dre A5 is also missing Intron 20 (fig. 2*b*). Mammalian cardiac genes also have fewer exons than their skeletal counterparts (*MYH7* lacks Intron 37, whereas *MYH6* lacks Introns 13 and 37: Liew et al. 1990; Epp et al. 1993) but missing introns are not the same in mammals and teleosts. Introns 13, 18, 36, and 37 are present in skeletal genes and in all teleost and mammalian unaffiliated genes. This suggests the ancestor of the cardiac and skeletal genes had the exon-intron organization observed in contemporary teleost skeletal MYH genes (Exons 3–41) and teleost and mammalian cardiac genes have independently lost introns. Overall, these data support the hypothesis that the vertebrate ancestor had a differentiated cardiac gene, which was independently duplicated in the mammal and fish lineages but retained (ancestral) cardiac function.

There is variation among cardiac genes in the length of Exons 7, 16, 17, and 41, but the intron phase is the same in all cardiac and skeletal genes (fig. 2*a, b*). Cardiac genes are slightly longer than skeletal genes, ranging from 11,210 to 18,775 bp (table 2) but are still shorter than described mammalian MYH genes (Liew et al. 1990; Epp et al. 1993; Desjardins et al. 2002; Schachat and Briggs 2002). Variation in exon length among both skeletal and cardiac genes suggests that insertions/deletions occur commonly over short evolutionary timescales.

Within the cardiac clade were a further two teleost genes (Fru X and Dre X: C* of table 2) that were orthologous to one another but distinct from the other teleost cardiac genes (fig. 1). These genes were both intronless. We infer that sometime after the differentiation of cardiac and skeletal genes, but prior to the divergence of the ancestors of *F. rubripes* and *D. rerio* (>150 mya: Cantatore et al. 1994), the X gene lost 37 introns. A plausible mechanism for loss of all introns is retrotransposition (Vanin 1985). Retrotransposed genes usually rapidly become pseudogenes because of the isolation of the gene from its regulatory elements. High similarity (>98%) of Dre X to ESTs from adult and embryonic heart (table 4) suggested that either (1) Dre X is expressed, or (2) Dre X is >98% similar to an unidentified, expressed gene. If, as we interpreted from the phylogeny, Dre X is an ortholog of Fru X and intron loss occurred once prior to the divergence of the ancestors of these two taxa, it is implausible that Dre X is >98% similar to any extant gene. Thus, we infer that Dre X is a functional gene that has retained the ancestral cardiac function.

Unaffiliated Genes

In contrast to the consistency of amino acid and intron evolution of the skeletal and cardiac clades, there appeared to be independent evolution of intron position and head versus rod amino acid sequence among the remaining genes in our analyses. Two genes are outside all other urochordate and vertebrate MYH genes in both phylogenies (*MYH16* and Cin 3: fig. 1*a, b*). This suggested urochordates and vertebrates both retained the two MYH lineages identified by Chiba et al. (2003). However, Cin 3 has an exon-intron organization more similar to other *C. intestinalis* genes than to *MYH16* (fig. 3). We speculate that these two genes have

Table 4
ESTs matching *D. rerio* MYH Genes

Gene[a]	Clade[a]	EST[b]	Tissue
A1	C	—	—
A2	C	—	—
A3	C	—	—
A4	C	AL924233 (99%)	Whole fish (embryo and adult)
A5	C	CB356582 (100%); CB354777 (99%); AW777840 (100%); AI522354.1 (96%)	Whole fish (0–72 hpf); 26 somite embryos, adult livers, shield stage embryos
B1	S	BC044194 (99%)	Whole fish (adult)
B2	S	CB353600 (100%)	Whole fish (0–72 hpf)
B3	S	AF180893 (98%); AI641219 (99%)	Embryonic fast muscle (10 $\geq$ 22 somites); 26 somite embryos, adult livers, shield-stage embryos
Other B[c]		NM_152982; BC044194; AF165817; AI957444; AI618577	Embryonic cranial muscles, pectoral fin, and tail muscle; 26 somite embryos, adult livers, shield-stage embryos and; embryonic heart 3 dpf
C1	C	AF114427 (99%); AI105857 (99%)	Embryonic ventricle; adult myocardium/endocardium/vessel
C2	C	BE693126 (99%); BQ262137 (98%)	Embryonic heart 3 dpf; whole fish (adult male)
D1	S	AL914053 (99%); AI882980 (99%); BM573824 (96%)	Whole fish (embryo and adult); whole fish (adult); adult pectoral fin
Q	S	AL923098 (100%)[1]; AW343155 (98%)[2]	Whole fish (adult and mixed stages)
X	C*	AI878213 (100%)[3]; BM280081 (100%)[4]; AW454123 (99%)[5]	26 Somite embryos, adult livers, shield stage embryos; adult heart; embryonic heart 3 dpf
Y	U	—	—
Z	U	—	—

[a] Gene codes and clades defined in table 2.
[b] EST matches for 3′ or 5′ UTR, Loop 1 or Loop 2, except Dre Q and Dre X: [1] Exons 27–30; [2] Exons 30–34; [3] Exons 27–29; [4] Exons 26–27; [5] Exons 36–37. Across the region matched by ESTs Dre Q was $<$81% similar to other *D. rerio* genes and Dre X was $<$73% similar to other *D. rerio* genes.
[c] Other genes on ctg11069.2.

different functions (i.e., contractile properties). Cin 3 is expressed in adults, possibly in the body wall (Chiba et al. 2003), which has characteristics of smooth, rather than sarcomeric muscle (reviewed by Meedel and Hastings 1993). *MYH16* is expressed in cat masticator muscle and has been considered to be superfast (Qin et al. 2002). If the two genes do have different functional properties, amino acid similarity is less likely to be the result of convergent evolution.

The remaining vertebrate genes formed two monophyletic clades, either of a similar age as the cardiac and skeletal lineages (head: fig. 1*b*) or much older, arising at the same time as the ancestor of the cardiac/skeletal lineage (rod: fig. 1*a*). We refer to these two clades of genes as unaffiliated as they are phylogenetically distinct from the well known skeletal and cardiac genes. One human and one chicken gene was placed in each of the two clades, but fish and frog genes were restricted to one clade or the other (fig. 1*a, b*). All four fish genes formed a monophyletic clade based on amino acid sequences (fig. 1*a, b*), but these genes have different exon-intron organization from one another (U versus U*: table 2; fig. 2*c* and *d*) and from their putative ortholog, *MYH14* (fig. 7 of Desjardins et al. 2002). Maximum parsimony analysis of the presence/absence of introns was unable to resolve the relationship of U genes to U* genes nor to *MYH14* and *MYH15* (fig. 3), which was not surprising given how few parsimony informative characters were considered. The function of these genes is unclear in fish and humans (Desjardins et al. 2002), but a cardiac function has been inferred for both chickens and frogs. The lack of resolution in this region of the phylogeny may reflect either an inadequate sampling of genes or a genuinely complicated evolutionary history of variable evolutionary rates and lineage sorting associated with functional shifts among taxa.

The remaining four *C. intestinalis* genes were placed in a monophyletic clade (fig. 1*a* and *b*). Consistent with Chiba et al. (2003), the three embryonic/larval genes (Cin 4, Cin 5, and Cin 6) were more closely related to one another than to the cardiac gene Cin 2 (fig. 1*a* and *b*). These data further support the hypothesis that duplication and differentiation giving rise to cardiac and skeletal genes occurred following divergence of the urochordate and vertebrate ancestors.

Phylogenetic and Functional Relationships Among Fish Skeletal MYH Genes

Two aspects of the phylogenetic relationships within the fish skeletal clade have implications for this study. First, relationships among sequences do not resemble the relationships among teleost taxa (e.g., Ostariophysi (Salmoniformes (Acanthopterygii))) (e.g., Miya et al. 2003) (fig. 1), indicating skeletal MYH genes underwent some number of duplications prior to the diversification of teleosts. Second, there is no a simple relationship between expression and phylogenetic position (fig. 1; tables 3 and 4). Specifically, genes amplified from white muscle were paraphyletic with genes amplified from slow-twitch

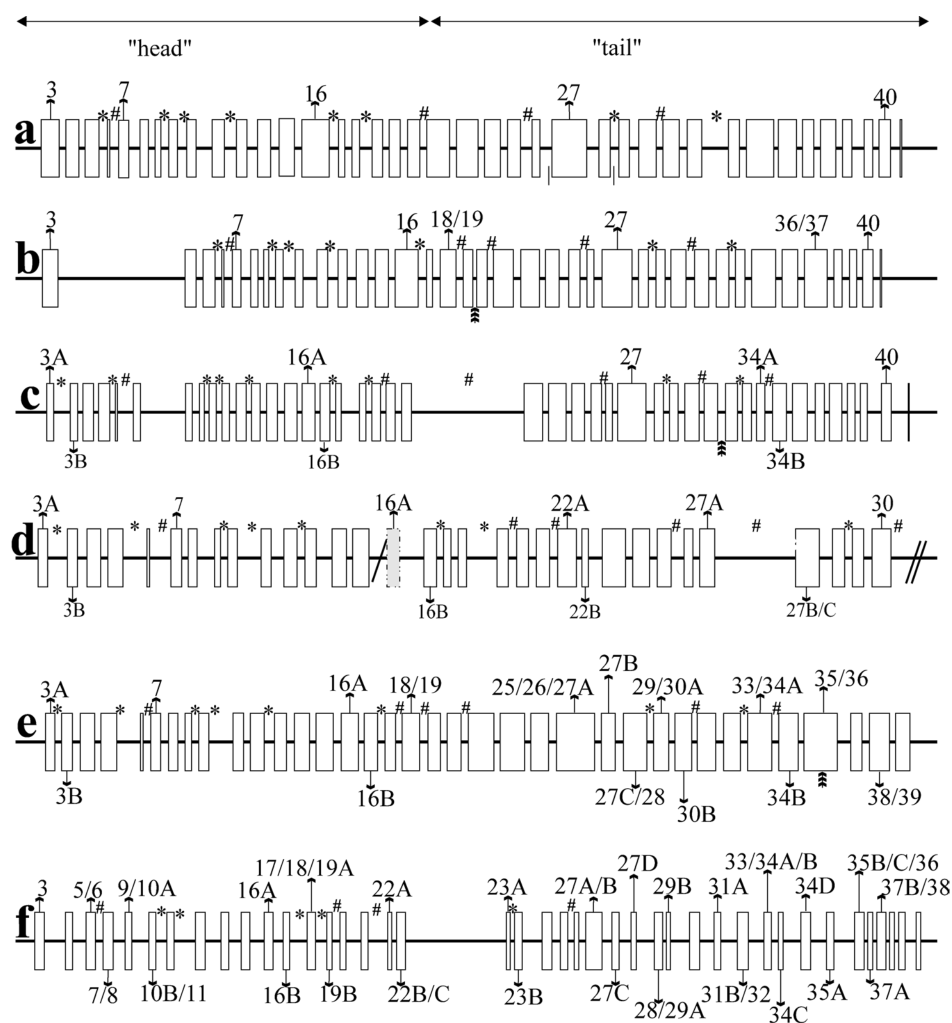

FIG. 2.—Schematic of exon/intron organization of MYH genes in this study, scaled by total gene length. Head and rod regions are indicated (boundary in Exon 22). Exons are represented by boxes and introns by the space between exons. Exon number was based on human *MYH3* (Strehler et al. 1986). Several exons of *MYH3* were interrupted by introns in other genes, and these were labeled with letters after the exon number (A for the new exon comprising the 5′ end of the *MYH3* exon, and so on). When introns were absent, both *MYH3* exon numbers are indicated. Unless otherwise noted, all exons were phase 0. Phase 1 exons were indicated by an asterisk (*) and phase 2 exons by a hash mark (#). (*a*) "S" genes (table 2) (schematic based on Fru E1). (*b*) "C" genes (table 2) (based on Dre A5). Dre A5 lacked Intron 20, as indicated by the position of the triple arrow. (*c*) "U" genes (table 2). Dre Z lacked Intron 31, indicated by the position of the triple arrow. (*d*) Fru Y ("U*" table 2). No sequence was available for intron 15 (/), and only partial sequence was available for Exon 16A. Uncertainty about the length of 16A was indicated by the broken outline. No sequence was available beyond Exon 30, as indicated (//). (*e*) Organization shared by Cin 4, Cin 5, and Cin 6. Cin 4 has an intron interrupting Exon 35, as indicated by the position of the triple arrow. (*f*) Cin 3 organization. Figure based on information summarized in supplementary table 1.

muscle, and slow-twitch genes were localized to both the skeletal and cardiac clades (fig. 1*c*). The localization of slow-twitch MYH genes to both the skeletal and cardiac clades suggests the role of phylogenetically skeletal MYH genes might differ subtly between fish and mammals, the latter of which express only cardiac MYH genes in slow-twitch fibers. Based on the functional differences between MYH head and rod, we suggest phylogenetic relationships based on the head are more likely to reflect functional relationships. There are two major clades of fish skeletal MYH genes in the head phylogeny (fig. 1*b*). One of these clades (indicated with a solid circle in fig. 1*b*) is characterized by a single synapamorphy, an additional amino acid at the 3′ end of Exon 3 (glutamic acid in all sequences except Dre B3 (threonine)). cDNA sequences included in the clade were obtained from axial (white) muscle, including that of *G. aculeatus* and *Notothenia coriiceps* (fig. 1*b*; table 3). The second clade contained some genes amplified from white muscle, in addition to Gac 2(R) and Nco R, which were amplified from red muscle (fig. 1; table 3). We hypothesize that genes in the glutamic acid clade encode MYH proteins with a faster contraction rate than other skeletal MYH genes. Thus, we suggest genes in the clade with Gac 2(R) and Nco R will have slower contraction than genes in the clade with Gac 1(W) and Nco W, regardless of the tissue from which they were amplified.

Processes Involved in the Evolution of Fish MYH Genes

Several clones from both *D. rerio* and *F. rubripes* contained sequence of more than one MYH gene (table 2).

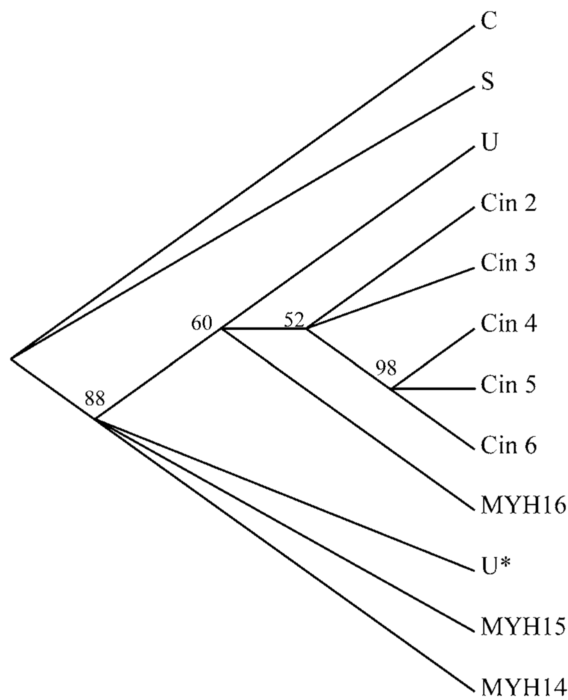

FIG. 3.—Bootstrapped parsimony phylogeny based on presence/absence of introns. Groups S, C, and U are defined in table 2, which also gives accession information for the other genes.

Thus, at least some teleost MYH genes are tandemly arrayed, as in mammals and chickens. Tandemly arrayed genes are either all skeletal or all cardiac, and no tandemly arrayed genes were observed or hypothesized for the unaffiliated group (table 2). *D. rerio* AL954717 (table 2) contained five genes, all oriented in the same direction along 114Kb of a chromosome (fig. 4*a*). Intergenic regions, crudely calculated from stop codon to start codon, ranged from 3.4 up to 14 kb (fig. 4*a*). This is similar to the intergenic distances of 7.5 kb observed between neonatal and embryonic MYH genes in chickens (Gulick, Kropp, and Robbins 1987), as well as the 4.5 to 60 kb among the six mammalian skeletal genes (Weiss et al. 1999).

Sequences inferred to be tandemly arrayed based on their presence in a single clone were assigned the same letter code (table 2) and were observed to be $\geq$95% similar across at least one, but not all, introns. Several sequences from different clones were also $\geq$95% similar across at least one, but not all, introns and these were also assigned a shared letter code (table 2). The pattern of high-sequence similarity in some regions of the gene (exon and intron), but dissimilarity in other regions, was consistent with gene conversion through nonhomologous recombination of tandemly arranged paralogs. Thus, we hypothesize that genes with the same letter code (table 2) are tandemly arrayed. These hypotheses of genome organization can be tested by mapping the genes or by further sequencing.

There is statistical evidence of a history of gene conversion for the MYH genes of the *D. rerio* clone AL954717 ($P < 0.0001$). The hypothesis of gene conversion was supported (no mismatches allowed) for 19 fragments, which represent all gene pairs except Dre A1 with Dre A5 (the genes on the ends of the array: fig. 4*a*). Converted fragments of coding region ranged in length from 45 to 515 bp (avg. 229 bp) (table 5). Several fragments spanned introns over which there was also evidence of gene conversion (e.g., fig. 4*b*), suggesting the reported converted fragment length is a minimum. As we allowed no mismatches, the results suggest gene conversion events were recent. Converted fragments are patchily distributed through the gene, but occur in both the head (<Exon 22) and rod (>Exon 22) (table 5). Dre A5 also had several fragments of unique sequence, which were inferred to be the result of conversion with a gene not included in the analysis (table 5). We identified several other genes that were $\geq$95% similar to AL954717 sequences across several but not all introns (table 2). There might be more than five genes tandemly arrayed along this section of the chromosome.

Discussion

Origin of Vertebrate MYH Genes

Chiba et al. (2003) detected two independent lineages of MYH genes in the urochordate *C. intestinalis*, only one of which was represented in vertebrates. Our phylogenetic analyses identified vertebrate orthologs of both ascidian lineages. This discrepancy is due to the inclusion of the mammalian *MYH16* gene in our study, but not in Chiba et al.'s (2003). *MYH16* was the only vertebrate ortholog of the ascidian Cin 3 gene we identified, probably because of incomplete sampling of vertebrate genomes. The descendants of this ancestral chordate gene do not appear to have duplicated in either ascidians or vertebrates. We cannot postulate whether the single copy of the *MYH16*/Cin 3 gene is due to no duplication, perhaps attributable to some aspect genomic organization, or due to loss/degradation of duplicates, perhaps a result of functional constraints on the protein product. Our results were consistent with the conclusion of Desjardins et al. (2002) that *MYH16* is an ancient gene, evolving at a similar rate to other MYH genes. Identification and functional characterization of other orthologs of this ancient chordate gene may assist in our interpretation of its evolution.

The second MYH gene present in the ancestral chordate underwent independent duplication and divergence in ascidians and vertebrates. Our analyses indicated at least four sarcomeric MYH genes (in addition to the one discussed above) in the last common ancestor of all vertebrates. These four ancestral genes correspond to the contemporary skeletal, cardiac, *MYH14* and *MYH15* lineages. The different functional domains of the MYH protein (head versus rod) inferred different timing of the origin of the latter two lineages (late versus early). The different scenarios inferred from the two functional domains reflect differences in evolutionary rate, possibly resulting from functional shifts in the encoded proteins. We note that, in contrast to the conservation across vertebrates of exon organization and expression patterns in cardiac and skeletal genes, the unaffiliated genes (*MYH14, MYH15*, and their orthologs) vary in exon organization, and their function appears to have evolved among vertebrate classes. Gga SM2, Gga Vent ,and Rca C were identified from the transcriptome of heart or slow-twitch muscle (Oana et al.

a

A1 A2 A3 A4 A5

10.7 Kb 3.4 Kb 14.0 Kb 12.3 Kb

b

Intron 28 (81bp)

```
A4 gtaaggtctcaatagtaatcaactaaaatgtgttgctctaaaatgtaaatttgtcattaaatttattttctggtatttcag
A5 gtagggtctcaatagtaatcaactaaaatgtgtcgctctaaaatgtaaatttgtcattaaatttattttctggtatttcag
```

Intron 31 (82bp)

```
A1 gtgaacaaaaaactgagtgttgaataaattaccaattgagaaccatcagcagatatttaaagcttctttaaataacaaacag
A3 gtgaacaaaaaactgagtgttgaataaattaccaattgagaaccatcagcagatatttaaagcttctttaaataacaaacag
```

Intron 4 (79bp)

```
A1 gtgagtggagtcttaaaacactctcacactcactttatgatggaggagcagatctaaagctgctgtcgatctcttccag
A2 gtgagtggagtcttaaaacactctcacactcactttatgatggaggagcagatctaaagctgctgtcgatctcttccag
A3 gtgagtggagtcttaaaacactctcacactcactttatgatggaggagcagatctaaagctgctgtcgatctcttccag
A4 gtgagtggagtcttaaaacactctcacactcactttatgatggaggagcagatctaaagctgctgtcgatctcttccag
```

FIG. 4.—(*a*) Genomic organization of MYH genes in the *D. rerio* AL954717 clone, depicting the order along the chromosome, the direction of transcription, and intergenic space (estimated from stop codon to start codon). (*b*) Nucleotide alignments for several introns across which there was evidence of gene conversion (table 5).

1998; Gauthier et al. 2000; Machida et al. 2002), but their mammalian and teleost orthologs were identified from genomic sequence. We did not identify any ESTs matching either of the *D. rerio* unaffiliated genes (Dre Y or Dre Z). Desjardins et al. (2002) noted several EST matches for *MYH14* and *MYH15* in humans, but these were not from cardiac or known slow contracting muscles. Desjardins et al. (2002) were unable to detect these two genes using reverse transcriptase-polymerase chain reaction (RT-PCR) on RNA populations from various tissues. Thus, although genes in the unaffiliated group play a prominent role in the heart of some vertebrates, they do not appear to do so in others. Resolving the expression pattern of the unaffiliated genes in mammals and teleosts and determining the contractile properties of protein products may lead us to understand their evolutionary history.

Our results are consistent with the hypothesis that the duplication and differentiation that gave rise to a distinct skeletal versus cardiac MYH gene occurred after the divergence of vertebrates from urochordates, but before the divergence of sarcopterygians and actinopterygians. Both genes have descendants in fish, mammals, and birds. Amino acid phylogenies and exon organization are consistent with the hypothesis that cardiac MYH genes duplicated independently within each vertebrate class. However, our data are not consistent with the hypothesis that duplication and divergence of skeletal MYH genes was class specific. Amphibian MYH genes were sister to teleost genes, rather than to the Mammalia/Aves clade as expected from organismal phylogenies. These relationships were inferred only from the rod of MYH, as the head sequence was not available for the frogs. Protein sequence in the rod influences assembly of MYH proteins into filaments (Lowey et al. 1969; Cripps et al. 1999) and filament stability is temperature dependent (Kawabata et al. 2000). Is the similarity of frog MYH rod to teleost MYH rod due to convergent evolution in poikilotherms rather than to shared ancestry? This question might be addressed by obtaining the sequence for the head of the MYH protein for amphibian taxa and determining if relationships from head and rod are concordant. In addition, phylogenetic analysis of reptile MYH genes might also help to determine whether physiological similarity has driven convergent evolution of MYH genes.

Table 5
Gene Conversion Among *D. rerio* Group A Genes

Genes	*P* Value[a]	Fragment Length (bp)	Exon Position
A1:A4	0.010	107	4–5
A2:A4	0.001	195	4–5
A3:A4	0.001	164	4–5
A1:A3	0.007	272	5–8
A1:A2	0.018	170	9–11
A3:A4	<0.001	149	11–13
A1:A4	0.029	61	11–12
A1:A2	0.006	161	16
A2:A5	0.030	45	21
A2:A3	0.002	368	24–27
A1:A4	<0.001	327	27
A2:A4	0.042	140	27
A4:A5	<0.000	194	28–29
A1:A3	<0.001	414	31–33
A1:A2	0.008	317	31–33
A3:A4	0.016	120	34
A2:A4	0.042	120	34
A3:A5	0.014	506	34–37
A2:A5	0.030	515	34–37
A5	0.001	107	4–5
A5	0.002	61	11–12
A5	0.007	90	27
A5	0.007	53	32

[a] Entries for A5 alone indicate unique fragments: evidence of conversion with a gene not included in the analyses.

[b] Bonferroni-corrected Karlin-Altshul values; fragments ordered by physical position, not *P* value.

If the similarity of teleost and amphibian MYH genes is not due to convergence an alternative explanation is that there were at least two MYH genes in the vertebrate ancestor, one of which was retained in fish and frogs while the other was retained in birds and mammals retained both. Retention of a single ancestral gene in most vertebrate classes could have occurred through lineage sorting or directional gene conversion.

Mammalian skeletal MYH genes are not monophyletic: the ancestors of *MYH3* and *MYH13* were not mammal-specific genes. As these relationships were inferred from the functionally independent head and rod, they were probably not the result of convergent evolution. Moore et al. (1992, 1993) hypothesized that *MYH3* was the product of a duplication event that predated the divergence of Mammalia and Aves, but conversion (especially among chick MYH genes) obscured this relationship. Class-specific gene conversion could generate the same evolutionary pattern as class-specific gene duplication. Consideration of other vertebrate taxa, particularly basal teleosts, reptiles, and other birds might resolve the true history of skeletal MYH genes. However, if skeletal MYH genes evolved through lineage sorting, convergence, and/or gene conversion, it might be impossible to reconstruct their history.

Processes Driving Evolution of MYH genes

Based on our observation of several MYH genes sequenced from a single clone, we inferred at least some teleost skeletal and cardiac MYH genes are tandemly arrayed, as in mammals and chickens. This physical arrangement implicates tandem duplication as a major process in the evolution of the MYH gene family. There was no evidence that unaffiliated teleost MYH genes were tandemly arrayed, consistent with the observation that *MYH14* and *MYH15* occur independently on different chromosomes in humans (Desjardins et al. 2002). We identified retrotransposition as a rare mechanism of gene duplication in the teleost MYH gene family. Based on the number of genes we identified from the *D. rerio* and *F. rubripes* genomes, there has been greater duplication (or gene retention) in teleosts (table 2) than mammals (six skeletal, two cardiac, and two unaffiliated: Desjardins et al. 2002) or chickens (seven skeletal, one cardiac and two unaffiliated: Bandman and Rosser 2000, fig. 2). We detected several sarcomeric MYH ESTs in *D. rerio* that did not match genes identified from the genomic screen, indicating we have not yet identified all genes from zebrafish. Of the 29 MYH genes in carp (Gerlach et al. 1990; Kikuchi et al. 1999) three were skeletal (Cca W10, Cca W20, and Cca W30) and shared intron sequence with most of the others (see fig. 1 of Kikuchi et al. 1999), a level of similarity that suggests all 29 genes are probably skeletal. The number of carp MYH genes may be unrepresentatively and transiently high due to the recent tetraploidization of carp (Larhammar and Risinger 1994; David et al. 2003). Thus, although we predict that there are more than seven skeletal MYH genes in most fish species, we expect the number to be fewer than in carp.

We proposed (see *Introduction*) whole genome duplications in the teleost lineage (Amores *et al.* 1998; Postlethwait *et al.* 1998; Taylor *et al.* 2001) would have generated more MYH genes in fish than other vertebrates. The relationships of MYH genes from different fish taxa do not appear to be consistent with proliferation of MYH genes through whole genome duplication. However, gene conversion would make it impossible to distinguish orthologs from paralogs based solely on amino acid phylogenies (see below), masking any signature of whole genome duplication. Mapping the genomic position of sarcomeric MYH genes and subsequent synteny analyses should determine whether MYH whole genome duplication played a role in generating more MYH genes in teleosts than tetrapods. The unaffiliated genes in teleosts exhibited no evidence of gene conversion and may therefore represent the best genes with which to begin this task.

Our analysis indicated widespread and recent gene conversion among five tandemly arrayed cardiac genes in *D. rerio*. We suggest that gene conversion has played a dominant role in shaping the diversity of MYH genes observed in modern vertebrates (see also Moore et al. 1992; Epp et al. 1995; Desjardins et al. 2002). This may make it impossible to identify orthologs across taxa; in fact, rampant gene conversion would mean true orthologs no longer exist. Different fish taxa could have different suites of MYH genes, either as a result of chance (different conversion events in different populations) or selection (the same gene conversion has a detrimental effect in some populations but not in others). Thus, extensive gene conversion can restrict the divergent evolution of genes

within a taxon while accelerating divergence among taxa. There are two practical implications of high levels of gene conversion obscuring ortholog relationships. First, it is uninformative to ask whether different taxa utilize the same genes to perform a particular function as there is no such thing as the "same" gene. However, a single cluster of tandemly arrayed genes might fulfill the same function in a range of taxa. Second, there can be no informative system for naming MYH genes in fish, such as is available for mammals. Sequence from more MYH genes across a greater range of teleost taxa will allow us to determine how much of an issue gene conversion is, as well as how comparative studies of gene function should proceed. Meanwhile, researchers should be aware of the difficulty in inferring orthology.

For the most part, the evolutionary patterns of amino acid sequences were concordant with the relationships inferred from intron position. There was evidence of intron loss, and slight shifts in intron position (through addition/deletion of codons at the 3′ end of exons) among skeletal and cardiac genes, but more extensive evolution of intron position across all vertebrate sarcomeric MYH genes. For the most part, we have not analyzed intron position within a broad enough phylogenetic framework to infer direction (i.e., gain or loss), but both events appear to have occurred. Although amino acid similarity strongly supports the orthology of *MYH16* and Cin 3, position of introns does not. We interpret this result to indicate taxa-specific intron evolution. In support of this, we note that urochordates have a different distribution of intron phases (across all annotated genes) (urochordata: 30% Phase 0, 35% Phase 1, 35% Phase 2; gnathostomata 41% Phase 0, 30% Phase 1, 29% Phase 2: from the Intron Server website, introns.abc.hu/; Barta, Kaján, and Pongor 2003). *C. intestinalis* might have a different distribution of intron insertion (protosplice) sites. Also, *C. intestinalis* has a highly AT-rich genome compared to humans (Dehal et al. 2002), which holds for the MYH gene itself (e.g., Cin 3 has 61% AT; *MYH16* has 49% AT).

Evolution of MYH Gene Expression

It is well accepted that fish express fast MYH isoforms in white, fast-twitch muscles and a slow-twitch MYH in red muscle (e.g., Sänger and Stoiber 2001). However, this has not robustly investigated from a molecular genetic perspective, with coordinated information on gene sequence, protein contractile properties, and bioenergetic properties of muscle lacking. We expected to observe MYH genes amplified from white, fast-twitch muscle in a different phylogenetic clade to MYH genes from red, slow-twitch muscle. Specifically, based on the mammalian condition of skeletal gene expression in fast-twitch fibers versus cardiac gene expression in slow-twitch fibers, we expected the distinct phylogenetic clades of fast and slow fish MYH to be sister to the mammalian skeletal and cardiac clades respectively. For trout (*Oncorhynchus mykiss*) this pattern held. However, white and red muscle MYH genes from *G. aculeatus* and *N. coriiceps* (the only other taxa with sequence from both muscle types) were placed in different subclades within the skeletal clade. Do fish vary in whether they express cardiac or skeletal genes in red muscle? Expression and sequence data from a broader taxonomic range of teleosts are required before this question can be addressed. We note that both *G. aculeatus* and *N. coriiceps* are labriform (oscillation of paired fins) swimmers, whereas *O. mykiss* is a subcarangiform (oscillation of body and caudal fin) swimmer. This difference in locomotor style (and biology) might be correlated with differences in the phylogenetic origin of the MYH genes expressed in red muscle. Based in part on the phylogenetic placement of the red versus white muscle skeletal genes of *G. aculeatus* and *N. coriiceps*, we propose that a subset of skeletal genes (the glutamic acid clade) encode proteins with the fastest contraction of skeletal proteins. We require more coordinated information on amino acid sequence and contractile properties of the proteins before we can begin to assess this hypothesis.

Contradicting our expectation (based on protein motility analyses: e.g., Karasinski and Kilarski 1989; Martinez et al. 1989, 1991) that red, white, and cardiac muscle would express different MYH genes, we observed multiple MYH genes in the RNA population of a single tissue type. Specifically, in *G. aculeatus*, Gac W was amplified from white, red, and cardiac muscle, and Gac R was amplified from red muscle and the heart. This was also the distribution of expression patterns in *O. mykiss* (Gauvry and Fauconneau 1996; Weaver, Stauffer, and Coughlin et al. 2001). Coexpression of several genes might reflect developmental transitions, although as this seems unlikely in the case of *G. aculeatus*, as all RNA was extracted from sexually mature adults. Is the MYH isoform determined at the level of translation rather than transcription? We need both more detailed investigation of the RNA and proteins in muscles, as well as data from a broader array of taxa to address these questions.

Generally, our results support the history of MYH genes inferred from analyses of other vertebrates. However, the inclusion of teleost sequences in a taxonomically broad phylogenetic analysis of MYH genes highlighted several issues. Perhaps most importantly, gene conversion appears to have had a major influence both on early (i.e., phylogenetically deep) and contemporary evolution of MYH genes. Expression patterns of MYH genes appears to be more complex in fish than mammals, but this might be an artifact of mammalian data coming from model systems, whereas fish data are available from more varied systems. Our results provide a phylogenetic framework for future studies of fish MYH gene evolution and function. As predicted, fish appear to have ample natural variation to make good systems for the analysis of the relationship between gene sequence and contractile properties.

Acknowledgments

This manuscript is a product of the University of Oregon's Evolution, Development and Genomics training program, funded by the National Science Foundation. This material is based on work supported by the National Science Foundation under grant. 0236239. The authors thank W. Cresko, C. Everroad, and J. Tolman-Thompson for comments on the manuscript.

Literature Cited

Amores, A., A. Force, Y. L. Yan et al. (10 co-authors). 1998. Zebrafish hox clusters and vertebrate genome evolution. Science **282**:1711–1714.

Bandman, E., and B. W. C. Rosser. 2000. Evolutionary significance of myosin heavy chain heterogeneity in birds. Microsc. Res. Tech. **50**:473–491.

Barta, E., L. Kájan, and S. Pongor. 2003. IS: a web-site for intron statistics. Bioinformatics **19**:543.

Berg, J. S., B. C. Powell, and R. E. Cheney. 2001. A millennial myosin census. Mol. Biol. Cell **12**:780–794.

Burge, C., and S. Karlin. 1997. Prediction of complete gene structures in human genomic DNA. J. Mol. Biol. **268**:78–94.

Cantatore, P., M. Roberti, G. Pesole, A. Ludovico, F. Milella, M. N. Gadaleta, and C. Saccone. 1994. Evolutionary analysis of cytochrome b sequences in some Perciformes: Evidence for a slower rate of evolution than in mammals. J. Mol. Evol. **39**:589–597.

Chen, Q., L. A. Moore, M. Wick, and E. Bandman. 1997. Identification of a genomic locus containing three slow myosin heavy chain genes in the chicken. Biochem. Biophys. Acta. **1353**:148–156.

Chiba, S., S. Awazu, M. Itoh, S. T. Chin-Bow, N. Satoh, Y. Satou, and K. E. M. Hastings. 2003. A genomewide survey of developmentally relevant genes in *Ciona intestinalis* IX. Genes for muscle structural proteins. Dev. Genes Evol. **213**:291–302.

Cresko, W., Y.-L. Yan, D. Baltrus, A. Singer, A. Rodrigues-Mari, and J. H. Postlethwait. 2003. Genome duplication, subfunction partitioning, and lineage divergence: *Sox9* in stickleback and zebrafish. Dev. Dynam. **228**:480–489.

Cripps, R. M., J. Suggs, and S. Bernstein. 1999. Assembly of thick filaments and myofibrils occurs in the absence of the myosin head. EMBO J. **18**:1793–1804.

David, L., S. Blum, M. W. Feldman, U. Lavi, and J. Hillel. 2003. Recent duplication of the common carp (*Cyprinus carpio* L.) genome as revealed by analyses of microsatellite loci. Mol. Biol. Evol. **20**:1425–1434.

Dehal, P., Y. Satou, R. K. Campbell et. al. (84 coauthors). 2002. The draft genome of *Ciona intestinalis*: Insights into chordate and vertebrate origins. Science. **298**:2157–2167.

Desjardins, P. R., J. M. Burkman, J. B. Shrager, L. A. Allmond, and H. H. Stedman. 2002. Evolutionary implications of three novel members of the human sarcomeric myosin heavy chain gene family. Mol. Biol. Evol. **19**:375–393.

Drouin, G. 2002. Testing claims of gene conversion between multigene family members: Examples from echinoderm actin genes. J. Mol. Evol. **54**:138–139.

Ennion, S., L. Gauvry, P. Butterworth, and G. Goldspink. 1995. Small-diameter white myotomal muscle fibers associated with growth hyperplasia in the carp (*Cyprinus carpio*) express a distinct myosin heavy chain gene. J. Exp. Biol. **198**:1603–1611.

Epp, T. A., I. M. Dixon, H. Y. Wang, M. J. Sole, and C. C. Liew. 1993. Structural organization of the human cardiac alpha-myosin heavy chain gene (MYH6). Genomics. **18**:505–509.

Epp, T. A., R. Wang, M. J. Sole, and C. C. Liew. 1995. Concerted evolution of mammalian cardiac myosin heavy chain genes. J. Mol. Evol. **41**:284–292.

Gauthier, F. V., M. A. Qadir, P. A. Merrifield, and B. G. Atkinson. 2000. Postembryonic expression of the myosin heavy chain genes in the limb, tail, and heart muscles of metamorphosing amphibian tadpoles. Microsc. Res. Tech. **50**:458–472.

Gauvry, L., and B. Fauconneau. 1996. Cloning of a trout fast skeletal myosin heavy chain expressed both in embryo and adult muscles and in myotubes neoformed *in vitro*. Comp. Biochem. Physiol. B Biochem. Mol. Biol. **115**:183–190.

Gauvry, L., S. Ennion, C. Ettelaie, and G. Goldspink. 2000. Characterization of red and white muscle myosin heavy chain gene coding sequences from Antarctic and tropical fish. Comp. Biochem. Physiol. B Biochem. Mol. Biol. **127**:575–588.

Gerlach, G. F., L. Turay, K. T. A. Malik, J. Lida, A. Scutt, and G. Goldspink. 1990. Mechanisms of temperature acclimation in the carp: A molecular biology approach. Am. J. Physiol. **259**:R237–R244.

Goodson, H. V., H. M. Warrick, and J. A. Spudich. 1999. Specialized conservation of surface loops of myosin: Evidence that loops are involved in determining functional characteristics. J. Mol. Biol. **287**:173–185.

Gulick, J., K. Kropp, and J. Robbins. 1987. The developmentally regulated expression of two linked myosin heavy-chain genes. Eur. J. Biochem. **169**:79–84.

Hirayama, Y., and S. Watabe. 1997. Structural differences in the crossbridge head of temperature-associated myosin subfragment-1 isoforms from carp fast skeletal muscle. Eur. J. Biochem. **246**:380–387.

Iwami, Y., T. Ojima, A. Inoue, and K. Nishita. 2002. Primary structure of myosin heavy chain from fast skeletal muscle of Chum salmon *Oncorhynchus keta*. Comp. Biochem. Physiol. B. **133**:257–267.

Karasinksi, J., and W. Kilarski. 1989. Polymorphism of myosin isoenzymes and myosin heavy chains in histochemically typed skeletal muscles of the roach (*Rutilus rutilus* L., Cyprindae, fish). Comp. Biochem. Physiol. **92B**:727–731.

Kawabata, R., N. Kanzawa, M. Ogawa, and T. Tsuchiya. 2000. Determination of primary structure of amberjack myosin heavy chain and its relationship with structural stability of various fish myosin rods. Fish Physiol. Biochem. **23**:283–294.

Kikuchi, K., M. Muramatsu, Y. Hirayama, and S. Watabe. 1999. Characterization of the carp myosin heavy chain multigene family. Gene **228**:189–196.

Korn, E. 2000. Coevolution of head, neck, and tail domains of myosin heavy chains. Proc. Natl. Acad. Sci. U.S.A. **97**: 12559–12564.

Larhammar, D., and C. Risinger. 1994. Molecular genetic aspects of tetraploidy in the common carp *Cyprinus carpio*. Mol. Phylogenet. Evol. **3**:59–68.

Leinwand, L. A., R. E. Fournier, B. Nadal-Ginard, and T. B. Shows. 1983. Multigene family for sarcomeric myosin heavy chain in mouse and human DNA. Science. **221**:766–799.

Liew, C. C., M. J. Sole, K. Yamauchi-Takihara, B. Kellam, D. H. Anderson, L. P. Lin, and J. C. Liew. 1990. Complete sequence and organization of the human cardiac beta-myosin heavy chain gene. Nucl. Acids Res. **18**:3647–3651.

Lowey, S., H. S. Slayter, A. Weeds, and H. Baker. 1969. Substructure of the myosin molecule I: Subfragments of myosin by enzymatic degradation. J. Mol. Biol. **42**:1–29.

Machida, S., S. Noda, A. Takao, M. Nakazawa, and R. Matsuoka. 2002. Expression of slow skeletal myosin heavy chain 2 gene in Purkinje fiber cells in chick heart. Biol. Cell **94**:389–399.

Mahdavi, V., A. P. Chambers, and B. Nadal-Ginard. 1984. Cardiac alpha and beta myosin heavy chain genes are organized in tandem. Proc. Natl. Acad. Sci. U.S.A. **81**: 2626–2630.

Martinez, I., R. L. Olsen, R. Ofstad, C. Janmot, and A. d'Albis. 1989. Myosin isoforms in mackerel (*Scomber scombrus*) red and white muscles. FEBS Lett. **252**:69–72.

Martinez, I., J. S. Christiansen, R. Ofstad, and R. L. Olsen. 1991. Comparison of myosin isoenzymes present in skeletal and cardiac muscles of the Arctic charr *Salvelinus alpinus* (L.). Eur. J. Biochem. **195**:743–753.

Matsuoka, R., M. C. Yoshida, N. Kanda, M. Kimura, H. Ozasa, and A. Takao. 1989. Human cardiac myosin heavy chain gene mapped within chromosome region 14q11.2-q13. Am. J. Med. Genet. **32**:279–284.

Meedel, T. H., and K. E. M. Hastings. 1993. Striated muscle-type tropomyosin in a chordate smooth muscle, ascidian body-wall muscle. J. Biol. Chem. **268**:6755–6764.

Mitchell, E. J., R. Jakes, and J. Kendrick-Jones. 1986. Localization of light chain and actin binding sites on myosin. Eur. J. Biochem. **161**:25–35.

Miya, M., H. Takeshima, H. Endo et al. (12 coauthors). 2003. Major patterns of higher teleostean phylogenies: a new perspective based on 100 complete mitochondrial DNA sequences. Mol. Phylogenet. Evol. **26**:121–138.

Molina, M. I., K. E. Kropp, J. Gulick, and J. Robbins. 1987. The sequence of an embryonic myosin heavy chain gene and isolation of its corresponding cDNA. J. Biol. Chem. **262**:6478–6488.

Moore, L. A., W. E. Tidyman, M. J. Arrizubieta, and E. Bandman. 1992. Gene conversions within the skeletal myosin multigene family. J. Mol. Biol. **223**:383–387.

Moore, L. A., W. E. Tidyman, M. J. Arrizubieta, and E. Bandman. 1993. The evolutionary relationships of avian and mammalian myosin heavy-chain genes. J. Mol. Evol. **36**: 21–30.

Oana, S., S. Machida, E. Hiratsuka, Y. Furutani, K. Momma, A. Takao, and R. Matsuoka. 1998. The complete sequence and expression patterns of the atrial myosin heavy chain in the developing chick. Biol. Cell **90**:605–613.

Ojima, T., N. Kawashima, A. Inoue, A. Amauchi, M. Togashi, S. Watabe, and K. Nishita. 1998. Determination of primary structure of heavy meromyosin region of walleye pollack myosin heavy chain by cDNA cloning. Fisheries Sci. **64**:812–819.

Peng, M. Y., H. J. Wen, L. J. Shih, C. M. Kuo, and S. P. Hwang. 2002. Myosin heavy chain expression in cranial, pectoral fin, and tail muscle regions of zebrafish embryos. Mol. Reprod. Dev. **63**:422–429.

Postlethwait, J. H., Y.-L. Yan, M. Gates et al. (26 coauthors). 1998. Vertebrate genome evolution and the zebrafish gene map. Nat. Genet. **18**:345–349.

Qin, K., J. Kemp, M.-Y. Yip, P. R. L. Lam-Po-Tang, J. F. Y. Hoh, and B. J. Morris. 1990. Localization of human cardiac B-myosin heavy chain gene (MYH7) to chromosome 14q12 by *in situ* hybridization. Cytogenet. Cell Genet. **54**:74–76.

Qin, K., M. K. H. Hsu, B. J. Morris, and J. F. Y. Hoh. 2002. A distinct subclass of mammalian striated myosins: structure and molecular evolution of "superfast" or masticatory myosin heavy chain. J. Mol. Evol. **55**:544–552.

Rescan, P., B. Collet, C. Ralliere, C. Cauty, J. Delalande, G. Goldspink, and B. Fauconneau. 2001. Red and white muscle development in the trout (*Oncorhynchus mykiss*) as shown by *in situ* hybridization of fast and slow myosin heavy chain transcripts. J. Exp. Biol. **204**:2097–2101.

Ruiz-Trillo, I., J. Paps, M. Loukota, C. Ribera, U. Jondelius, J. Baguna, and M. Riutort. 2002. A phylogenetic analysis of myosin heavy chain type II sequences corroborates that Acoela and Nemertodermatida are basal bilaterians. Proc. Nat. Acad. Sci. **99**:11246–11251.

Sänger, A. M., and W. Stoiber. 2001. Muscle fiber diversity and plasticity. Pp. 187–250 *in* I. Johnston, ed. Muscle development and growth. Academic Press, San Diego.

Sawyer, S. 1999. Statistical tests for detecting gene conversion. Mol. Biol. Evol. **6**:526–538.

Schachat, F., and M. M. Briggs. 2002. Phylogenetic implications of the superfast myosin in extaocular muscles. J. Exp. Biol. **205**:2189–2201.

Schiaffino, S., and C. Reggiani. 1996. Molecular diversity of myofibrillar proteins: Gene regulation and functional significance. Physiol. Rev. **76**:371–423.

Strehler, E. E., M. Strehler-Page, J. Periard, M. Periasamy, and B. Nadal-Ginard. 1986. Complete nucleotide and encoded amino acid sequence of a mammalian myosin heavy chain gene: Evidence against intron-dependent evolution of the rod. J. Mol. Biol. **190**:291–317.

Swofford, D. 1998. PAUP*: Phylogenetic Analysis Using Parsimony (*and other methods). Version 4.0. Sinauer Associates Inc., Sunderland, Conn.

Szentkiralyi, E. 1984. Tryptic digestion of scallop S1: evidence for a complex between the two light-chains and a heavy-chain peptide. J. Muscle Res. Cell M. **5**:147–164.

Taylor, J. S., Y. Van de Peer, I. Braasch, and A. Meyer. 2001. Comparative genomics provides evidence for an ancient genome duplication event in fish. Phil. Trans. R. Soc. Lond. B. **356**:1661–1679.

Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. Clustal-W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. Nucl. Acids Res. **22**: 4673–4680.

Toyoshima, Y., S. Kron, E. McNally, K. Niebling, C. Toyoshima, and J. A. Spudich. 1987. Myosin subfragment-1 is sufficient to move actin filaments *in vitro*. Nature **328**:536–539.

Vanin, E. F. 1985. Processed pseudogenes: Characteristics and evolution. Ann. Rev. Genet. **19**:253–272.

Walesby, N. J., and I. A. Johnston. 1980. Fiber types in the locomotory muscles of an Antarctic teleost, *Notothenia rossii*: A histochemical ultrastructural and biochemical study. Cell Tissue Res. **208**:143–164.

Weaver, F. E., K. A. Stauffer, and D. J. Coughlin. 2001. Myosin heavy chain expression in the red, white, and ventricular muscle of juvenile stages of rainbow trout. J. Exp. Zool. **290**:751–758.

Weeds, A., and B. J. Pope. 1977. Studies on the chymotryptic digestion of myosin: Effects of divalent cations on proteolytic susceptibility. J. Mol. Biol. **111**:129–157.

Weiss, A., D. McDonough, B. Wertman, L. Acakpo-Satchivi, K. Montgomery, R. Kucherlapati, L. Leinwand, and K. Krauter. 1999. Organization of human and mouse skeletal myosin heavy chain gene clusters is highly conserved. Proc. Natl. Acad. Sci. U.S.A. **96**:2958–2963.

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. Dynam. **219**:201–215.

Yelon, D., S. A. Horne, and D. Y. Stainier. 1999. Restricted expression of cardiac myosin genes reveals regulated aspects of heart tube assembly in zebrafish. Dev. Biol. **214**:23–37.

Yoon, S. H., M. Kakinuma, Y. Hirayama Y, (??) T. Yamamoto, and S. Watabe. 2000. cDNA cloning of myosin heavy chain from white croaker fast skeletal muscle and characterization of its complete primary structure. Fisheries Sci. **66**:1163–1171.

William R. Jeffery; Associate Editor

Accepted January 14, 2004