The Journal of Experimental Biology 205, 2189–2201 (2002)
Printed in Great Britain © The Company of Biologists Limited 2002
JEB4016

Review

Phylogenetic implications of the superfast myosin in extraocular muscles

Fred Schachat* and Margaret M. Briggs

Department of Cell Biology, Division of Physiology, Duke University Medical School, Durham, NC 27710, USA

*e-mail: f.schachat@cellbio.duke.edu

Accepted 13 May 2002

Summary

Extraocular muscle exhibits higher-velocity and lower-tension contractions than other vertebrate striated muscles. These distinctive physiological properties are associated with the expression of a novel extraocular myosin heavy chain (MYH). Encoded by the MYH13 gene, the extraocular myosin heavy chain is a member of the fast/developmental MYH gene cluster on human chromosome 17 and the syntenic MYH cluster on mouse chromosome 11. Comparison of cDNA sequences reveals that MYH13 also encodes the atypical MYH identified in laryngeal muscles, which have similar fast contractile properties. Comparing the MYH13 sequence with the other members of the fast/developmental cluster, the slow/cardiac MYH genes and two orphan skeletal MYH genes in the human genome provides insights into the origins of specialization in striated muscle myosins. Specifically, these studies indicate (i) that the extraocular myosin is not derived from the adult fast skeletal muscle myosins, but was the first member of the fast/developmental MYH gene cluster to diverge and specialize, (ii) that the motor and rod domains of the MYH13 have evolved under different selective pressures and (iii) that the MYH13 gene has been largely insulated from genomic events that have shaped other members of the fast/developmental cluster. In addition, phylogenetic footprinting suggests that regulation of the extraocular MYH gene is not governed primarily by myogenic factors, but by a hierarchical network of regulatory factors that relate its expression to the development of extraocular muscles.

Key words: extraocular myosin, MYH13, myosin, phylogeny, exon–intron organization, gene structure.

Introduction

Extraocular muscles (EOMs) are of special interest because they differ from other skeletal muscles in ontogeny, innervation and contractile physiology. For these reasons, they are likely to provide novel insights into the origin of specialized vertebrate muscles and into the evolution of striated muscle itself.

EOMs arise from unsegmented head mesenchyme rather than somites, which give rise to trunk and limb muscles (Noden et al., 1999). They are innervated by cranial, rather than spinal, motoneurons and are organized into motor units composed of fibers with different molecular and ultrastructural properties (Goldberg and Shall, 1999). Ultrastructural analyses have identified six fiber types in extraocular muscle (Peachey, 1971; Spencer and Porter, 1988), and the majority of EOM fibers are singly-innervated and phasic, like limb and trunk skeletal muscles, while a minority are multiply-innervated and tonic. But, irrespective of their innervation, almost all are heterogeneous in myosin expression (Jacoby et al., 1990; McLoon et al., 1999; Rubinstein and Hoh, 2000) and none corresponds to the Type I and II fibers found in other skeletal muscles. Most importantly, they exhibit distinctive physiological properties, in particular, superfast contractions (Close and Luff, 1974; Li et al., 2000) that are linked to the expression of a novel superfast myosin heavy chain.

In many ways, the superfast EOM myosin is a molecular reflection of EOM's distinctive ontogenic, structural and physiological features. First, its expression is tissue-restricted: aside from extraocular muscle, it is found only in laryngeal muscles (Briggs and Schachat, 2000; Lucas et al., 1995). Second, in many EOM fibers, it is co-expressed with other myosins (Briggs and Schachat, 2001; Rubinstein and Hoh, 2000). And third, its localization to the central end-plate band region where EOM fibers are innervated links its expression directly to innervation (Briggs and Schachat, 2001; Rubinstein and Hoh, 2000). Here, the MYH13 gene that encodes this distinctive molecular marker is analyzed by genomic mapping, sequence-based phylogenetic techniques and phylogenetic footprinting to gain insights into the evolution of diversity and specialization in striated muscle myosins and the factors that probably account for the tissue-restricted expression of MYH13.

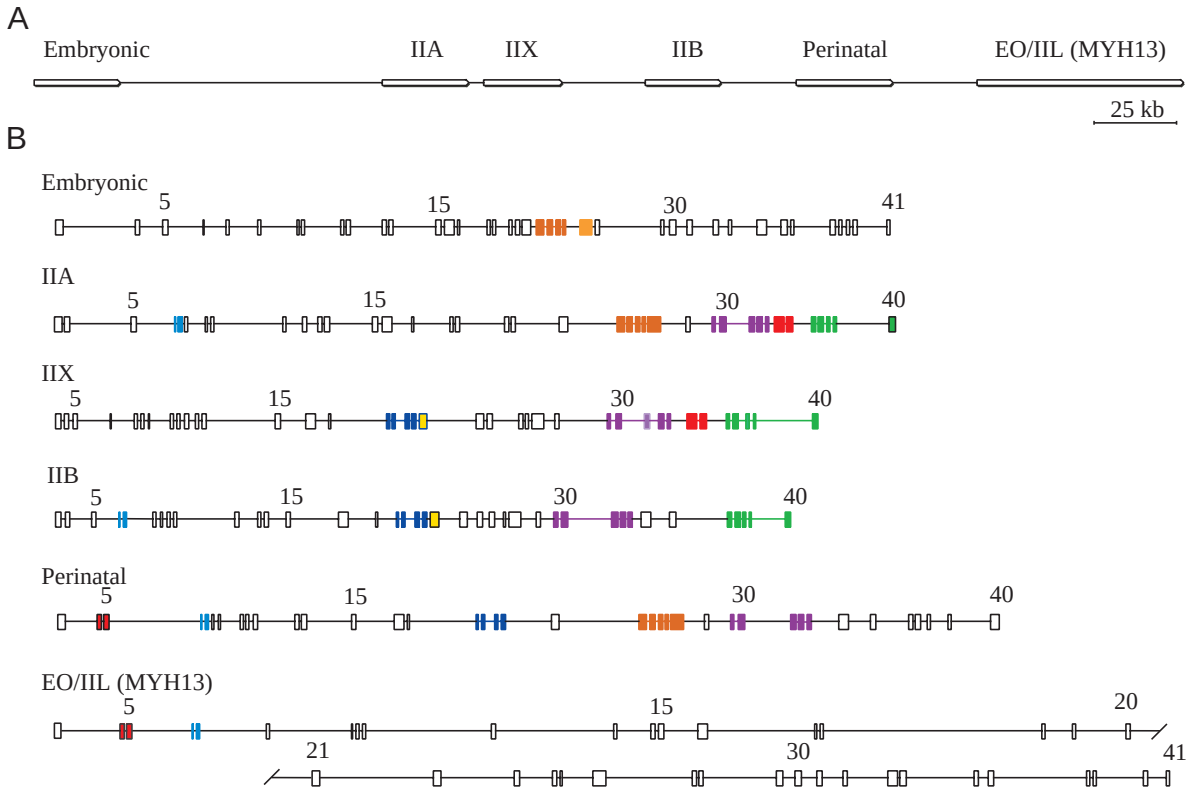


Fig. 1. Physical map and exon organization of the fast/developmental myosin heavy chain genes. (A) The position, orientation and size of the MYH genes that compose the fast/developmental cluster. (B) The exon organization of each of the MYH genes. The position and exon organization of the fast/developmental MYH genes are based on analysis of the MYH region of human chromosome 17 using GeneQuest and the known cDNA sequences of the rat embryonic MYH gene (Strehler et al., 1986) and the rabbit MYH13 gene (Briggs and Schachat, 2000) to define exon boundaries. Only the positions of the coding exons (3–40 for the embryonic and MYH13 gene and 3–41 for the other members of the cluster) are presented. The predicted coding sequences exhibited greater than 99 % identity with the cDNA sequences reported by Weiss et al. (1999b). Regions of homologous organization (exon–intron size and spacing) are indicated by distinctive colors. These structural homologies primarily reflect the order of gene duplication, although some, such as the two conserved patterns in the 5′-region of the MYH13 gene, are more probably the result of gene conversion events. This is readily evident for the segment from the beginning of exon 4 to the end of exon 5, which is identical to the corresponding sequence in the adjacent perinatal MYH gene.

The superfast extraocular myosin heavy chain gene is a member of the fast/developmental MYH gene cluster

Cloning of the 3′-untranslated region of the extraocular (EO) myosin heavy chain mRNA from rabbit and rat EOM enabled us to design primers to amplify the 3′ untranslated region (3′-UTR) of the human and mouse EO MYH genes and localize them to human chromosome 17 on the Stanford G3 radiation hybrid panel and mouse chromosome 11 on the EUCIB (European Backcross) panel (Winters et al., 1998). Neighboring sequence-tagged sites more specifically localized the human EO MYH gene to 17p13.1 to p12, placing it within 4 centimorgans of the fast/developmental MYH gene cluster identified by Yoon et al. (1992), and further screening showed that the EO MYH gene resides on YAC B120C11, which contains other members of the fast/developmental cluster. These chromosomal analyses led to the designation of the human and mouse EO MYH genes as MYH13 and myh13, respectively (Winters et al., 1998).

The structure of the MYH13 gene implies it has been ‘insulated’ from the other members of the fast/developmental cluster

The sequence of a full-length cDNA for the rabbit EOM MYH enabled us to identify the superfast myosin gene MYH from the human genome project sequences and to define its exon organization using GeneQuest (Briggs and Schachat, 1999, 2000; Schachat and Briggs, 1999). The exon boundaries of the other members of the fast/developmental MYH cluster – the embryonic, IIA, IIX, IIB and perinatal MYH genes – were similarly identified (Schachat and Briggs, 1999). The resulting map of the fast/developmental MYH cluster is presented in Fig. 1. These maps reveal that the MYH13 gene differs significantly from its neighbors in two ways. First, it spans almost 64 kb, more than twice the length of any other member of the cluster (Fig. 1A); and second, with the exception of two short 5′ regions, MYH13 lacks the internal similarities in exon organization evident in other members of the cluster (Fig. 1B).

Table 1. *The phase (reading frame) and length of the coding exons of the fast/developmental MYH genes*

Coding exon	Phase	Exon length					
		Embryonic	IIA	IIX	IIB	Perinatal	MYH13
3	(0,0)	204	183	204	204	210	204
4	(0,0)	144	144	144	144	144	144
5	(0,1)	157	157	157	157	157	157
6	(1,2)	28	28	28	28	28	28
7	(2,0)	109	115	115	115	109	112
8	(0,0)	93	93	93	93	93	93
9	(0,1)	64	64	64	64	64	64
10	(1,1)	99	99	99	99	99	99
11	(1,0)	104	104	104	104	104	104
12	(0,1)	139	139	139	139	139	139
13	(1,0)	119	119	119	119	119	119
14	(0,0)	150	150	150	150	150	150
15	(0,0)	171	171	171	171	171	171
16	(0,1)	307	310	310	310	310	310
17	(1,0)	71	68	71	71	68	74
18	(0,1)	88	88	88	88	88	88
19	(1,2)	118	118	118	118	118	118
20	(2,0)	124	124	124	124	124	124
21	(0,2)	137	137	137	137	137	137
22	(2,0)	256	256	256	256	256	256
23	(0,0)	243	243	243	243	243	243
24	(0,0)	177	177	177	177	177	177
25	(0,2)	146	146	146	146	146	146
26	(2,0)	91	91	91	91	91	91
27	(0,0)	390	390	390	390	390	390
28	(0,1)	127	127	127	127	127	127
29	(1,0)	119	119	119	119	119	119
30	(0,2)	197	197	197	197	197	197
31	(2,0)	184	184	184	184	184	184
32	(0,1)	166	166	166	166	166	166
33	(1,0)	125	125	125	125	125	125
34	(0,0)	309	309	309	309	309	309
35	(0,0)	204	204	204	204	204	204
36	(0,0)	126	126	126	126	126	126
37	(0,0)	171	171	171	171	171	171
38	(0,0)	105	105	105	105	105	105
39	(0,0)	96	96	96	96	96	96
40	(0,0)	138	150	150	150	147	135
41	(0,0)	24					12

The exon structure of these genes is highly conserved. The pattern of split codons is identical among all the genes. Exon lengths are also identical, with four exceptions. The exons encoding the N and C termini reflect known differences in amino acid sequence; and the C termini of MYH13 and the embryonic MYH are encoded by two exons, whereas the homologous sequences of the other MYHs are encoded by a single exon.

Additional length variations occur in the exons that encode the variable-length loop 1 and loop 2. The variation in loop 1 appears to be due to exon slippage at the 3′ boundary of exon 7, as does the variation in loop 2, which occurs at the boundary of exons 16 and 17.

The conservation of exon size and phase (reading frame) among members of the fast/developmental MYH cluster (Table 1) enables regions of homologous intron–exon spacing to be identified. In Fig. 1B, these regions are indicated by common color; only two common patterns in the 5′ region are present in MYH13. Because these common spatial patterns probably reflect lineage (e.g. gene duplications) and genetic exchanges between neighboring members of the cluster, both the size and the non-homologous intron–exon organization of MYH13 indicate that it has been largely ‘insulated’ from events that have shaped the current organization of the other members of the fast/developmental MYH gene cluster.

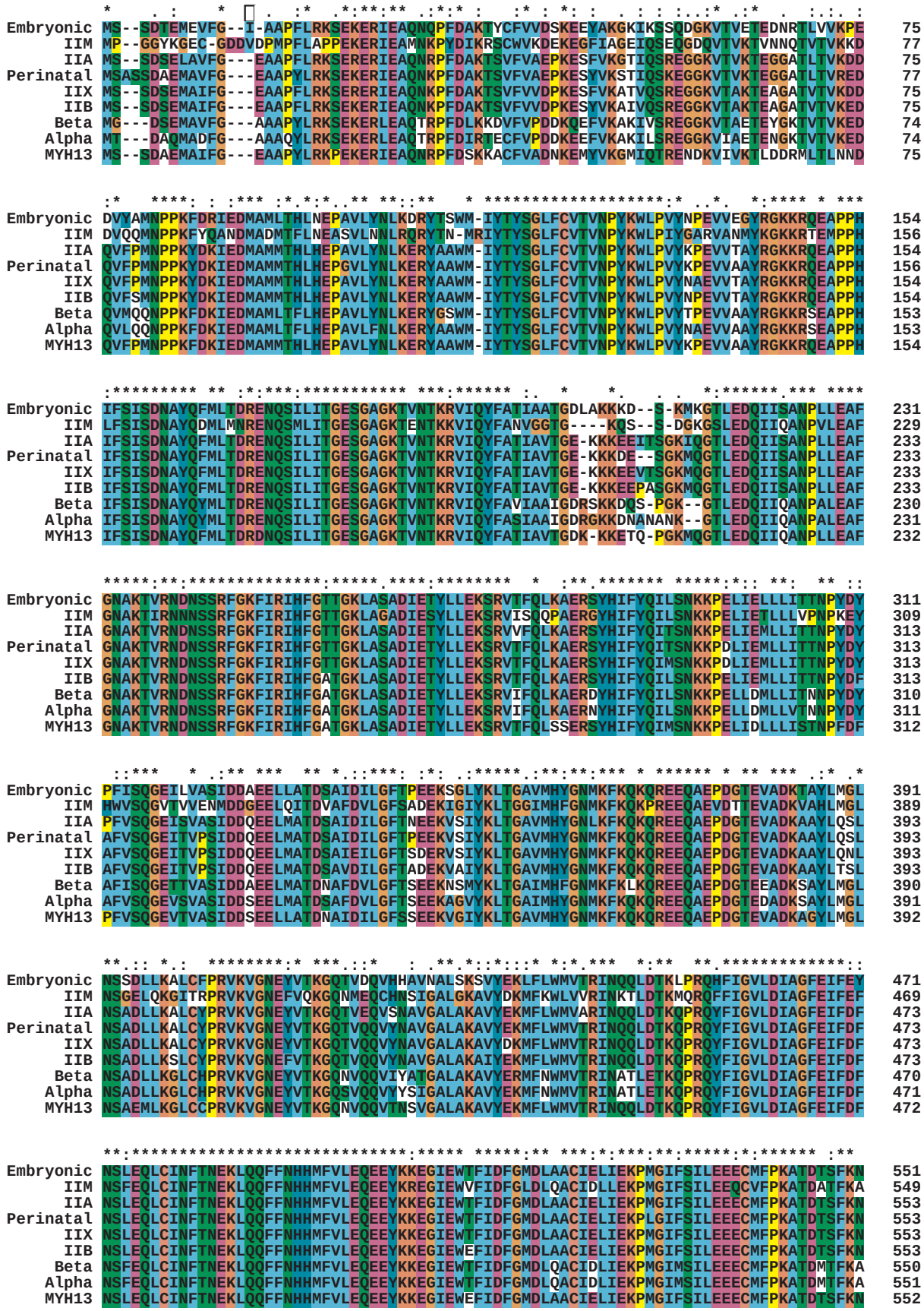


Fig. 2. For legend see p. 2195

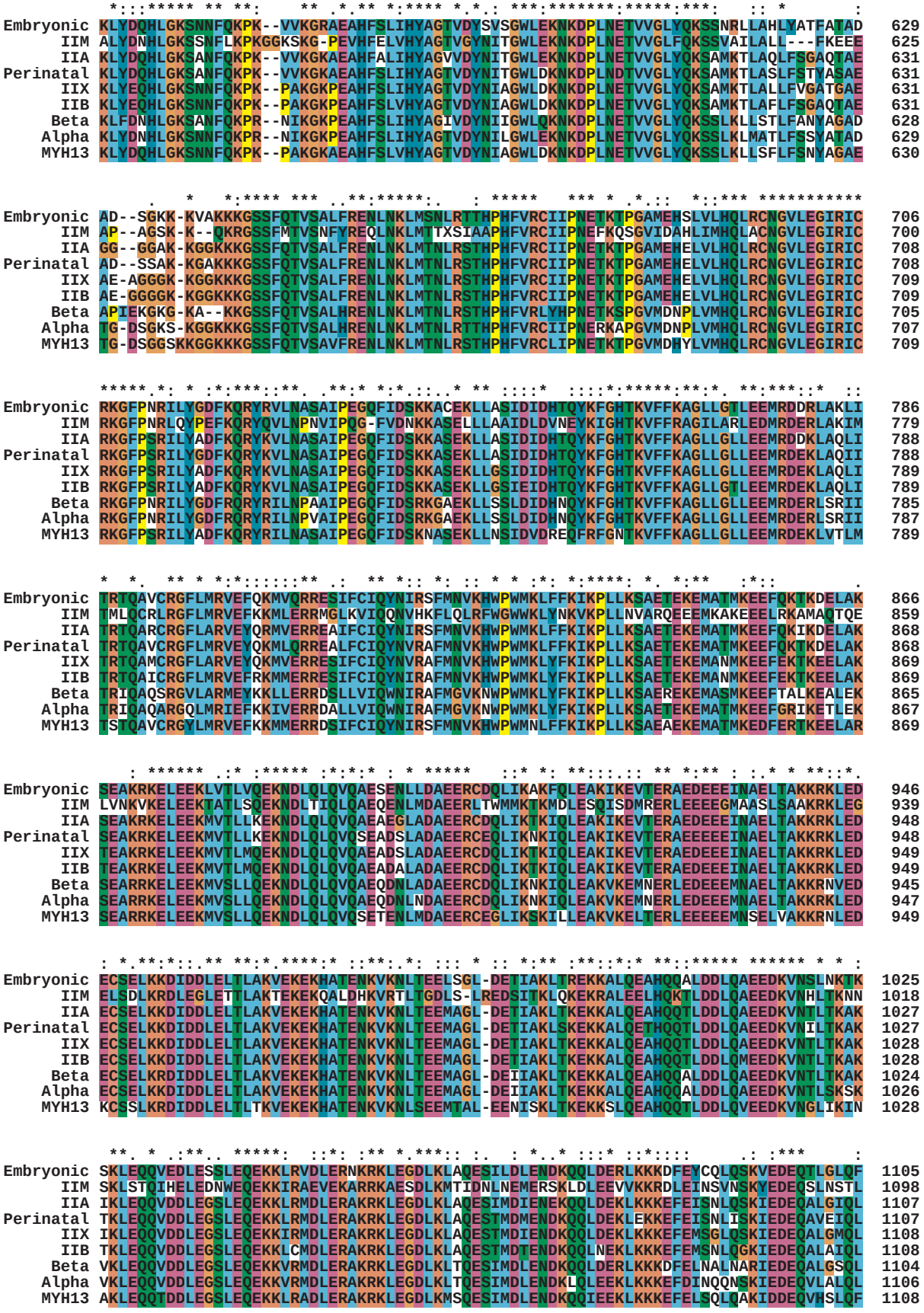
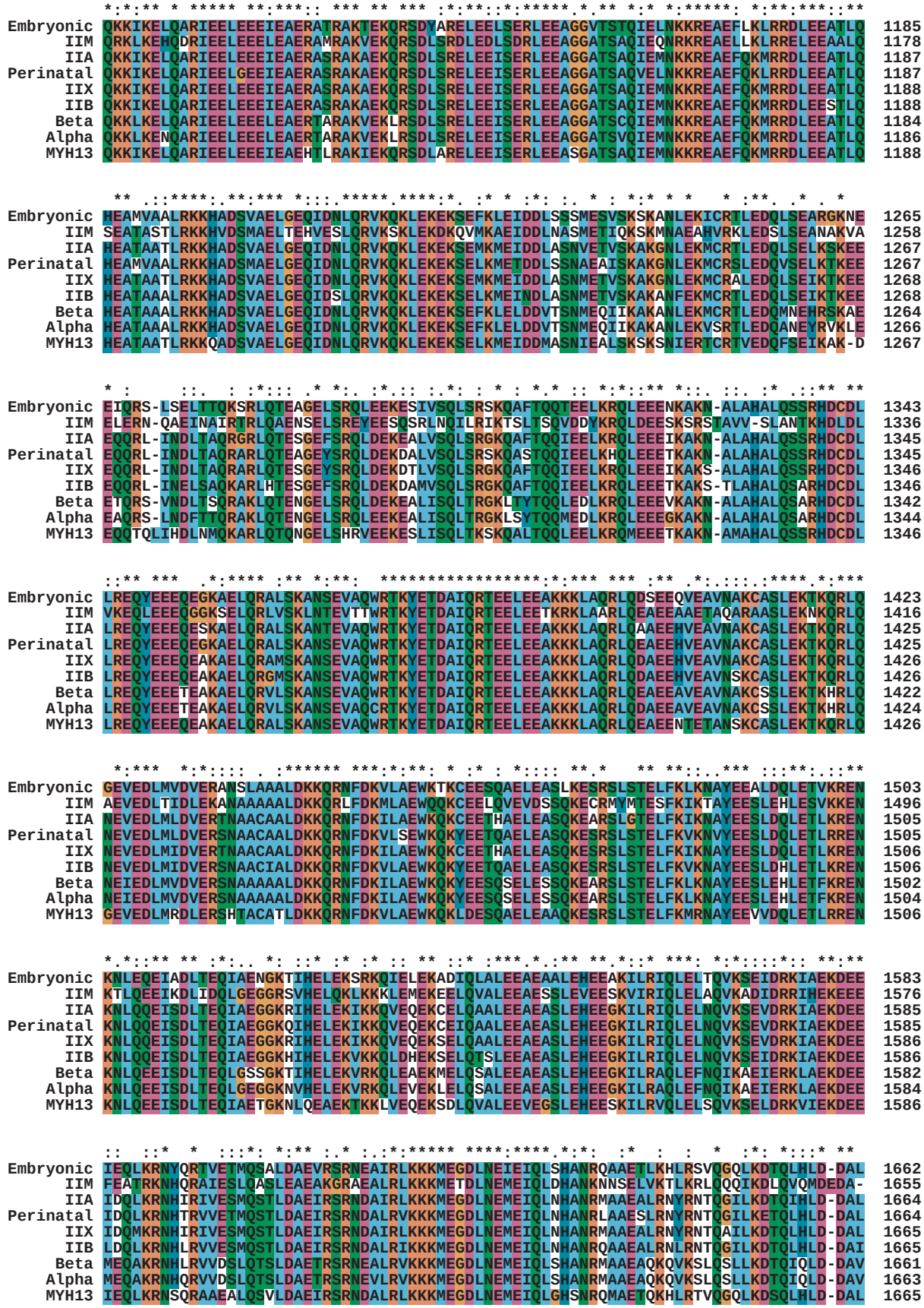


Fig. 2. For legend see p. 2195


Embryonic 1185
IIM 1178
IIA 1187
Perinatal 1187
IIX 1188
IIB 1188
Beta 1184
Alpha 1186
MYH13 1188
Embryonic 1265
IIM 1258
IIA 1267
Perinatal 1267
IIX 1268
IIB 1268
Beta 1264
Alpha 1266
MYH13 1267
Embryonic 1343
IIM 1336
IIA 1345
Perinatal 1345
IIX 1346
IIB 1346
Beta 1342
Alpha 1344
MYH13 1346
Embryonic 1423
IIM 1416
IIA 1425
Perinatal 1425
IIX 1426
IIB 1426
Beta 1422
Alpha 1424
MYH13 1426
Embryonic 1503
IIM 1496
IIA 1505
Perinatal 1505
IIX 1506
IIB 1506
Beta 1502
Alpha 1504
MYH13 1506
Embryonic 1583
IIM 1576
IIA 1585
Perinatal 1585
IIX 1586
IIB 1586
Beta 1582
Alpha 1584
MYH13 1586
Embryonic 1662
IIM 1655
IIA 1664
Perinatal 1664
IIX 1665
IIB 1665
Beta 1661
Alpha 1663
MYH13 1665

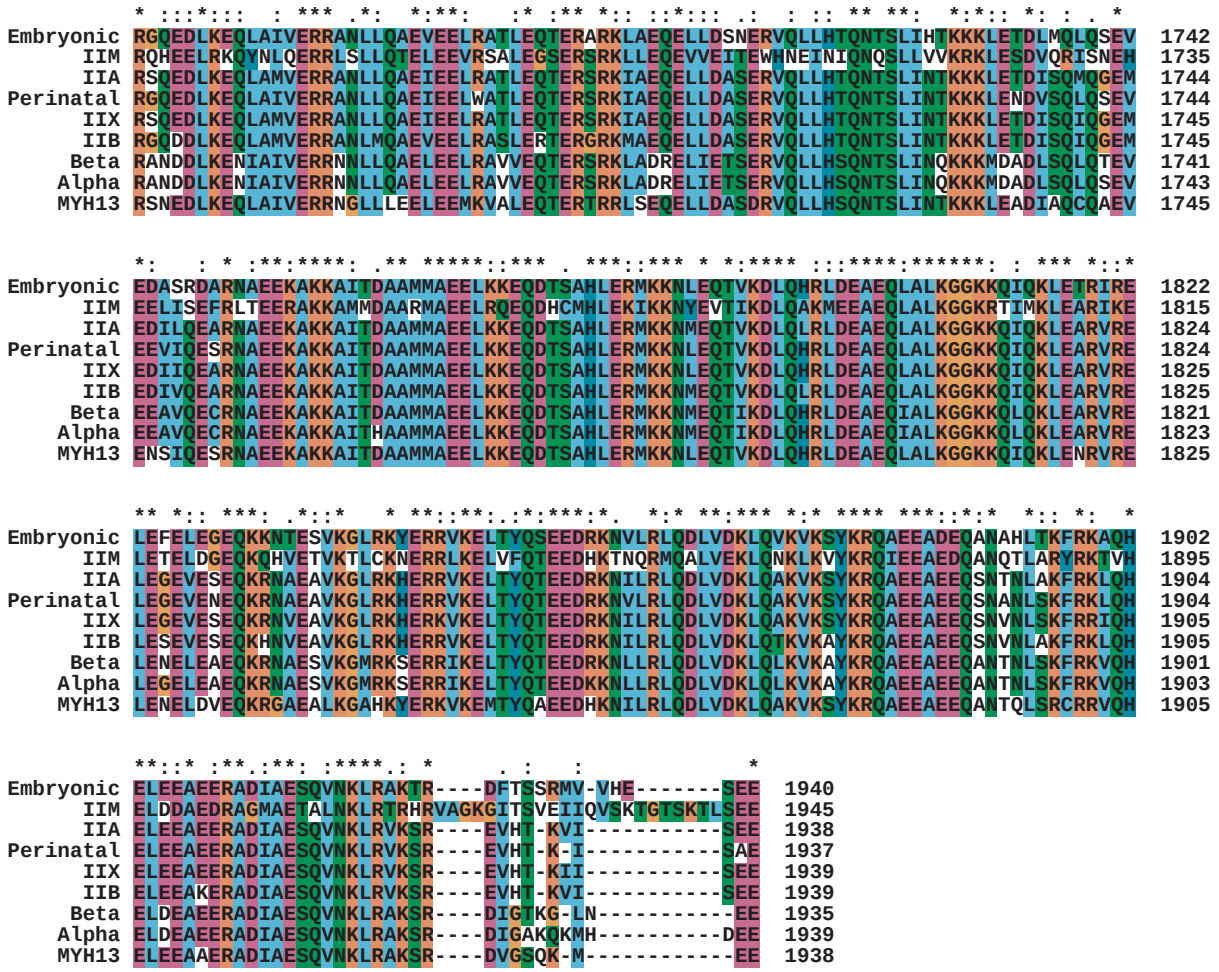


Fig. 2. Alignment of the fast/developmental and IIM MYH genes. To construct a sequence phylogeny, the amino acid sequences of the human embryonic, IIA, IIX, IIB, IIM, perinatal, cardiac α and β and MYH13 MYH gene sequence were aligned using Clustal X (Thompson et al., 1997). The high level and continuity of sequence conservation among striated muscle myosins made it possible to use low penalties for gap introduction and extension without introducing spurious gaps. Low values were necessary to properly align both the N- and C-terminal sequences. *, colon and stop mark positions with identical, strongly or weakly conserved amino acids, respectively. Color indicates the chemical characteristics of amino acids: light blue, hydrophobic; purple, acidic; dark blue, H and Y; green, N, Q, S and T; dark brown, R, K and C; light brown, G; yellow, P. A white background marks non-conserved residues.

The EO MYH occupies a pivotal position in the phylogeny of mammalian striated myosin heavy chains

Ten striated myosin heavy chain genes have been localized in the human genome: the six members of the fast/developmental MYH gene cluster (the embryonic, IIA, IIX, IIB, perinatal and extraocular genes) (Weiss et al., 1999a); two slow/cardiac MYH genes α and β on human chromosome 14 (Mahdavi et al., 1984; Saez et al., 1987); the masticatory myosin (IIM), an orphan striated muscle MYH gene on chromosome 7 (Hoh et al., 1999), which appears to be a pseudogene encoding the IIM myosin in humans (Schachat and Briggs, 1999), and a gene that probably encodes the slow-tonic MYH on chromosome 20 (Berg et al., 2001). Fig. 2 presents an amino acid sequence alignment of the striated MYHs using Clustal X (Thompson et al., 1997), and Fig. 3 shows the phylogeny inferred by neighbor-joining analysis. MYH IIM, the most divergent MYH gene sequence, was used as an outlier and, as depicted in the figure, bootstrapping analysis indicates that the phylogenetic tree is well-supported. In this phylogeny, MYH13 occupies a critical position. It arose from the precursor of the fast/developmental MYH genes shortly after the fast/developmental and slow cardiac precursors diverged. Thus, MYH13 represents an ancient MYH lineage. It appears to be the first member of the fast/developmental cluster to diverge and specialize, having arisen even before the developmental MYH genes.

Support for that ancient lineage is suggested by phylogenetic analysis of the two functionally distinct domains of striated muscle myosin. Previous studies have found that the sequences of other striated muscle myosin head or motor domains and their coiled-coil rod or tail domains generate similar phylogenetic relationships – implying that the these functionally distinct domains have evolved at similar rates (Korn, 2000). However, as shown in Fig. 4, the motor and rod

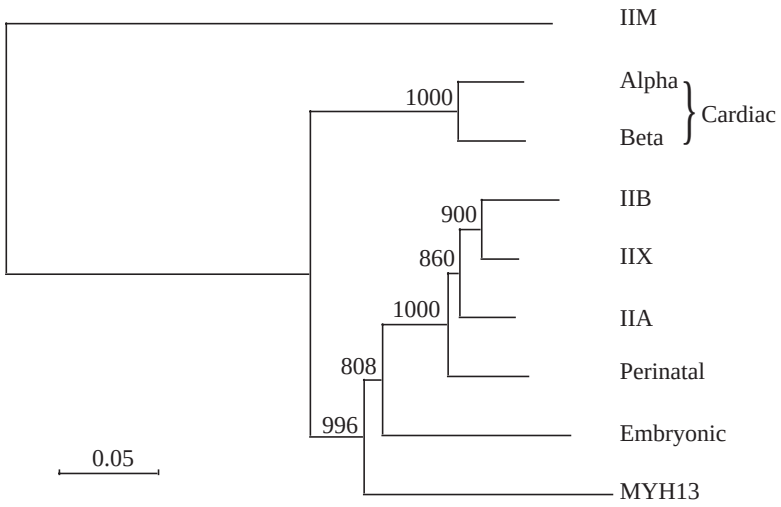


Fig. 3. Phylogenetic relationships among the striated MYH genes. The phylogenetic tree is based on neighbor-joining (N-J) analysis of the MYH amino acid sequences aligned in Fig. 2. All gapped positions were omitted for the N-J analysis. Because it was the most divergent of the striated MYHs, the IIM gene was designated as the outlier. The number of times 1000 independently chosen subsequences yielded the same tree are indicated at the node points. This bootstrap analysis indicates that the tree is well supported. A partial human IIM sequence was assembled by analysis of the chromosome 7 DNA sequence using GeneQuest. Because the human IIM genomic sequence is currently incomplete, a full-length IIM sequence was generated by fusing the human sequence to amino acid residues 1–525 of the cat IIM sequence (Hoh et al., 1999). A frameshift we identified in the coding region of the human IIM gene probably explains the absence of IIM expression in human (Rowlerson et al., 1983).

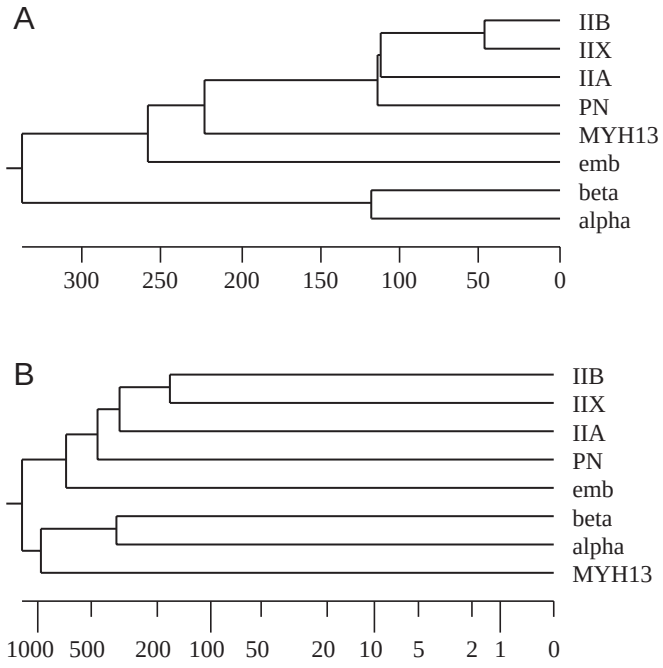


Fig. 4. Differences in selective pressures on the two functional domains of myosin. Sequence relatedness analysis using the entire amino acid sequence for the fast/developmental MYH genes generated different phylogenetic relationships for (A) the myosin head or motor domain, amino acid residues 1–841, and (B) the α-helical coiled-coil rod or tail domain, residues 842 to approximately 1938. The motor sequences were more similar to the adult fast MYH genes, indicative of both the role of the MYH13 gene in the superfast contractions of extraocular muscle and the two regions of exon homology evident in Fig. 1. In contrast, the rod domain exhibited a closer relationship to the α and β cardiac sequences, reaffirming the inference that MYH13 diverged and specialized before the precursors of the slow/cardiac and fast/developmental clusters had significantly diverged. This is a modified version of a figure from Briggs and Schachat (2000). emb, embryonic; PN, perinatal.

domains of MYH13 do not follow this relationship: the rod is more closely related to the slow/cardiac MYHs than to the fast/developmental genes, while the motor domain is more similar to the fast/developmental MYHs. Coupled with the sequence phylogeny, this finding implies that the MYH13 gene arose before the fast/developmental and cardiac MYH genes had significantly diverged, possibly at a time when the precursors of the slow/cardiac and fast/developmental gene clusters were linked.

Independent support for the early divergence of the MYH13: exon organization

The exon organization of the MYH13 gene provides independent support for its strategic position in the radiation of mammalian striated muscle MYH genes. Its exon boundaries are identical in position and phase (reading frame) with those previously described for the rat and human embryonic MYHs (Stedman et al., 1990; Strehler et al., 1986). However, MYH13 differs from the adult fast and perinatal MYH genes in its 3′ coding exons (Fig. 1B; Table 1). The 49 C-terminal amino acid residues of MYH13 are encoded by two exons, exons 40 and 41; but the homologous sequences of the adult fast and perinatal MYHs are encoded by a single exon, exon 40. This single-exon pattern is also shared with the slow/cardiac genes (Epp et al., 1993; Liew et al., 1990), providing independent support for the inference from sequence phylogeny that MYH13 and the slow/cardiac genes arose from duplications of a common ancestor – one that possessed a ‘primitive’ 3′-exon organization rather than the ‘derived’ single-exon organization of the perinatal and adult fast MYHs (Schachat and Briggs, 1999).

Distinctive features of full-length MYH13 transcripts suggest an early divergence and specialization

In addition to its 3′-exon organization and the ‘insulation’ suggested by its exon spacing, MYH13 exhibits distinctive

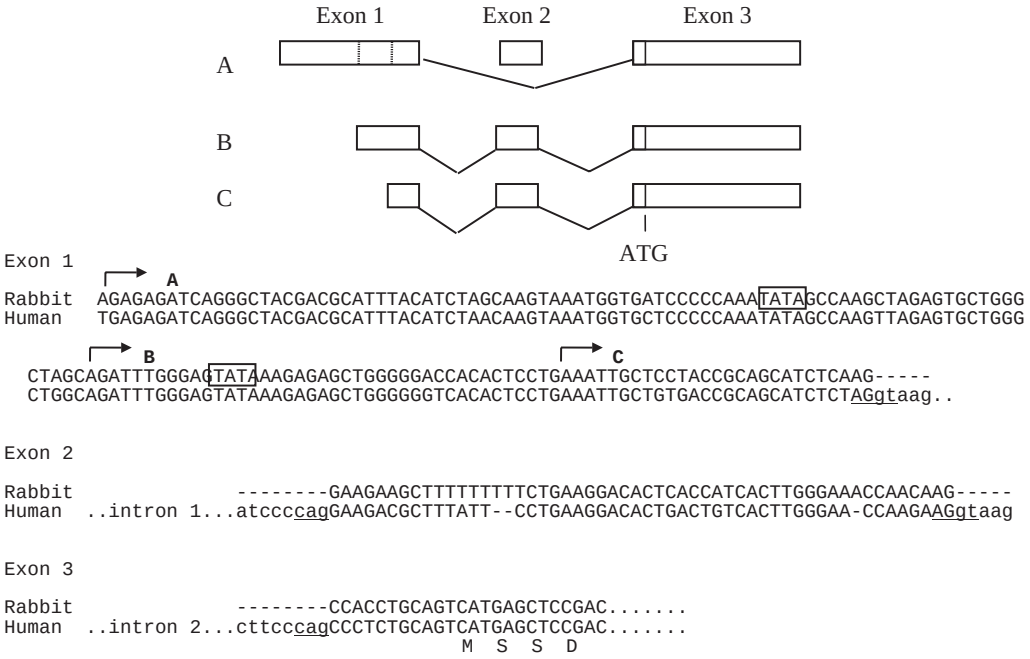


Fig. 5. MYH13 uses multiple transcription initiation sites and generates alternatively spliced mRNAs. The alternative splicing patterns of three full-length cDNAs of extraocular myosin are shown. Transcript A skips exon 2. Transcription of B and C begin downstream of A in exon 1, and they both include exon 2. The nucleotide sequence of the longest rabbit cDNA is aligned with the human gene sequence to show the predicted exon boundaries. Short stretches of the introns are shown in lower case, and the consensus splice sites are marked by underlining. The locations of the transcription start sites A, B and C are indicated by horizontal arrows, and the TATA boxes upstream of B and C are boxed. Only part of exon 3 is shown. From Briggs and Schachat (2000).

transcriptional features that distinguish it from the other members of the fast/developmental gene cluster and from the slow/cardiac MYH genes. Analysis of rabbit full-length MYH13 cDNAs (Briggs and Schachat, 2000) indicates that transcription can initiate at three sites (Fig. 5). Transcripts A and B generate 5′-UTRs that are significantly longer than the single 5′-UTRs reported for other MYH genes. Comparison of transcripts B and C with A shows they are shorter and each incorporates an additional 50-nucleotide insertion that corresponds to a 5′-non-coding exon present in the human EO myosin cDNA (Weiss et al., 1999b). Although alternative splicing of the 5′-UTR sequences and multiple transcription initiation sites have been described for several muscle genes including enolase, cardiac troponin T and acetylcholinesterase, where they are important for the developmental or regional regulation of gene expression (Giallongo et al., 1993; Jin et al., 1995; Li et al., 1993; Takechi et al., 1994), this is the only striated muscle MYH gene for which it has been described.

Phylogenetic footprinting suggests that transcription of MYH13 is regulated by a specialized network of transcriptional activators

Genomic sequence analysis also provides insights into the mechanisms that may account for the tissue-restricted patterns of MYH13 expression. Promoter and enhancer analysis of the embryonic, cardiac and IIB MYH genes has identified transcription factors that regulate striated-muscle-specific expression. These include the myogenic factors MyoD, myogenin and myf-5, which bind E-box sequences, and binding sites for other key transcription factors such as MEF2, TEF, SP1 and Oct1 (Bouvagnet et al., 1987; Firulli and Olson, 1997; Gossett et al., 1989; Gupta et al., 1998; Shimizu et al., 1992; Swoap, 1998; Takeda et al., 1992). Generalizing from the organization of regulatory elements described for the IIB and cardiac MYH genes, as well as many other genes (Blackwood and Kadonaga, 1998), the binding sites for these factors generally cluster in a proximal promoter within several hundred nucleotides upstream of the transcription initiation site and distal regions that function as transcriptional enhancers or suppressors. These regions typically lie 1–2 kb upstream of the promoter and in the first or second intron, although they may reside beyond the 3′ end of the gene, as is the case for myosin light chain 1f/3f (Donoghue et al., 1988). The initial inspection of these regions of the MYH13 gene in the human genome suggested that it was regulated differently (Briggs and Schachat, 2000). Strikingly, only a single binding site for these factors, an E-box, was present in the proximal promoter region (Fig. 6). This led us to propose that MYH13 expression might be directed by a novel set of transcription factors.

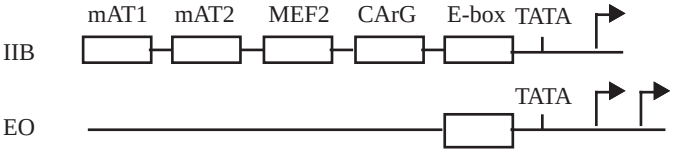


Fig. 6. Comparison of the density of myogenic factor binding sites in the proximal promoter regions of the IIB and MYH13 genes. Phylogenetic footprinting shows that the extraocular (EO) MYH proximal promoter lacks most of the functionally important binding sites for myogenic factors identified in the IIB gene.

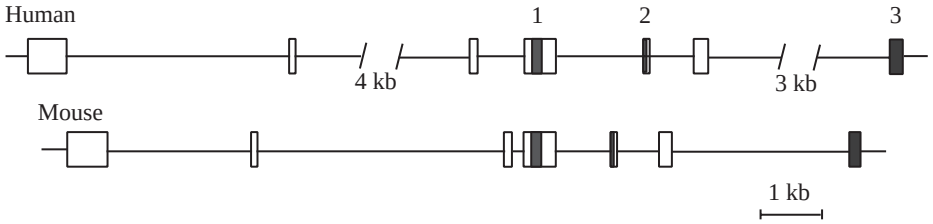


Fig. 7. The position of potential MYH13 regulatory regions based on phylogenetic footprinting. Phylogenetic footprinting identifies several highly conserved upstream and intronic regions between the human and mouse extraocular (EO) MYH genes. Conserved regions are depicted by open boxes. The positions of exons 1, 2 and 3 are indicated as filled boxes. The mouse genomic sequence spanning the myh13 gene (AC019008) was identified using Blast, and the exon boundaries were determined using GeneQuest.

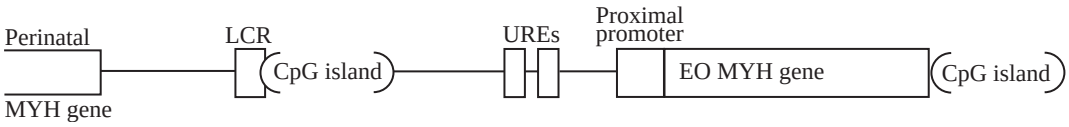


Fig. 8. A model for the MYH13 gene chromosomal domain. Conserved sites identified by phylogenetic footprinting (see Fig. 7) and two bounding CpG island domains revealed by GeneQuest are shown on a map of the human chromosomal region from the end of the upstream perinatal gene to a region downstream of the MYH13 gene. Here, the sites are tentatively assigned functions as the proximal promoter, upstream regulatory elements (UREs) and a locus control region (LCR) by analogy with the model for gene organization presented by Blackwood and Kadonaga (1998). This region is drawn to scale.

Because its expression is restricted to extraocular and laryngeal muscle, regulation of the EO MYH must differ from that of the other fast/developmental and slow/cardiac MYHs. Minimally, this means that factors other than the myogenic transcription factors must also be involved in determining when and where MYH13 is transcribed. The co-expression of MYH13 with adult fast myosins in extraocular muscle and laryngeal muscle fibers, coupled with its absence from limb fibers expressing the adult fast myosins, serves to emphasize its requirement for regulatory factors other than those that drive the trunk and limb muscle MYH genes.

Insights into the nature of that regulation come from the identification of conserved sequences in the 5′ flanking and intronic regions of the human MYH13 and mouse myh13 genes. This comparative genomic technique, known as phylogenetic footprinting, is based on the observation that important regulatory sequences are conserved across species and has been used to detect muscle-specific regulatory sequences (Wasserman and Fickett, 1998). Using BLAST 2 (Tatusova and Madden, 1999), several highly conserved regions were identified, including a likely proximal promoter, which includes 250 bp upstream of the transcription start sites, as well as several possible enhancer or repressor regions upstream and in the first two introns (Fig. 7).

The potential transcription factor binding sites in these regions were characterized with the matInspector 2.2 (Quandt et al., 1995) and GeneQuest (DNAstar, Madison, WI, USA) programs, which compare a test sequence with a large database of consensus binding sites. Many potential sites are clustered in the conserved regions. Among the most frequent consensus sequences present in both the mouse and human genomic sequences are those for NF1, which directs aldolase expression in a subset of fast fibers (Salminen et al., 1996), Pitx2, which is necessary for EOM embryonic development (Kitamura et al., 1999), NKX2.5, which is a mammalian homolog of tinman, a regulator of cardiac development (Schwartz and Olson, 1999), TCF11, which binds a DNase-hypersensitive site in the β-globin cluster (Johnsen et al., 1998), and delta EF1, which inhibits myoD-promoted myogenesis (Sekido et al., 1994). Some myogenic factor binding sites are also present, but they do not generally occur in the clusters that allow for myogenic factor dimerization that is critical for activating other MYH genes. The Pitx2, NKX2.5, TCF11 and delta EF1 consensus binding sites are particularly interesting given the early specialization of the EO MYH13 gene, its phylogenetic relationship to the slow/cardiac MYH genes, whose transcription is not regulated by myogenic factors, and the myoD-independent development of EOM revealed in studies on the myoD/myf5 knockout mouse (Tajbakhsh et al., 1997). The presence of these conserved sites suggests that MYH13 expression may be integrated into a broader more primitive developmental program that directs the formation of EOMs or of the eyes themselves. This is consistent with the finding that NKX2.5, which regulates cardiac development, is expressed at high levels in extraocular muscle and the fact that mammalian homologs of factors that control eye development in *Drosophila*, including Dach2, Six1, Eya2 and Pax3, have been found to participate in a regulatory network that directs myogenic factor expression in vertebrate somites (Heanue et al., 1999).

Indications of hierarchical regulation of MYH13 by chromatin structure

In addition to promoter and enhancer elements, some genes are regulated by additional elements that control their chromatin conformation. An open chromatin conformation is a necessary prerequisite for transcription. Open conformations may be generated locally in the proximal promoter by the

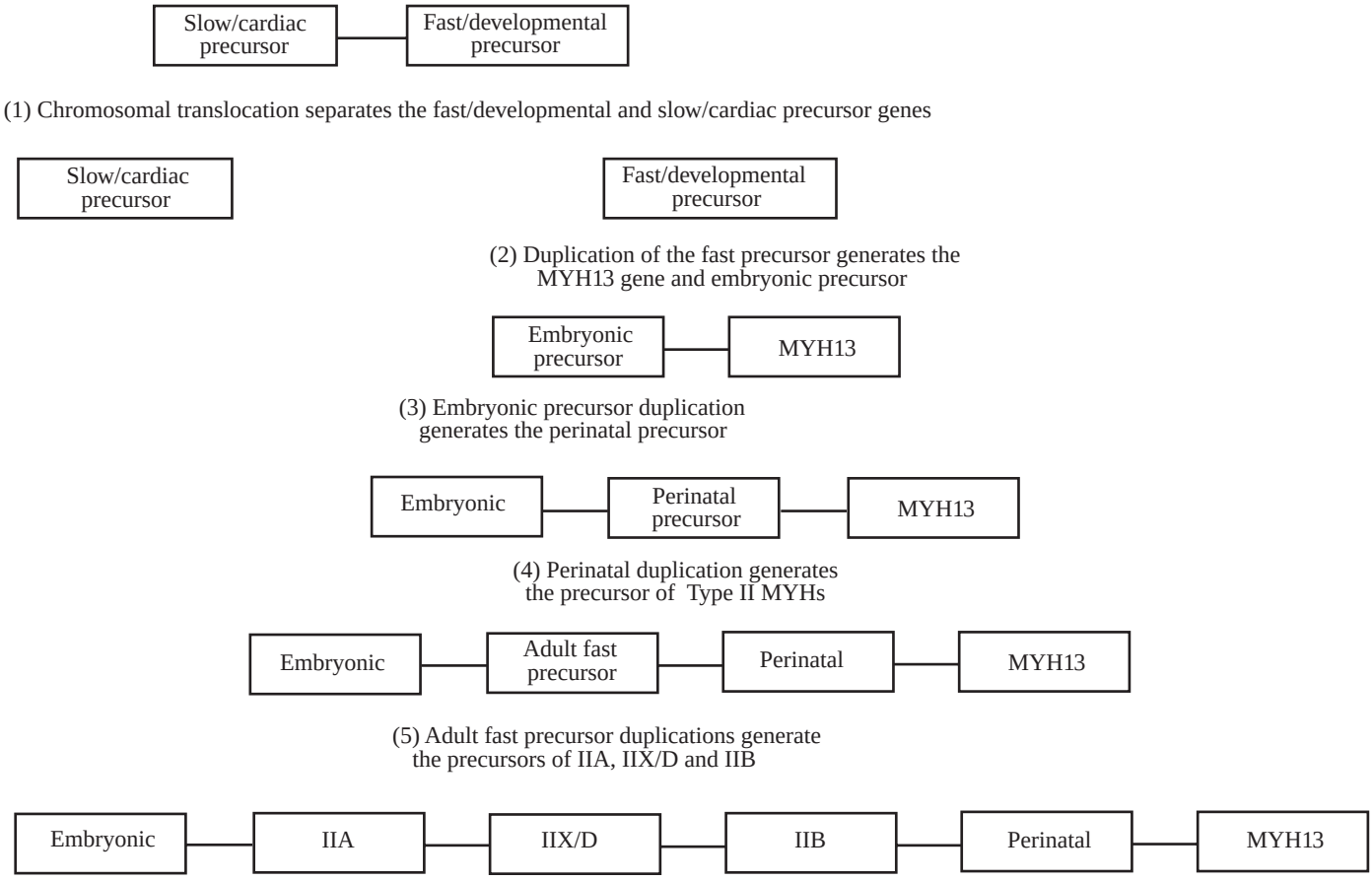


Fig. 9. Modeling the natural history of the muscle fast/developmental gene cluster. This model of gene duplication integrates the genomic organization (Fig. 1) and phylogeny of the fast/developmental MYH cluster.

action of histone deacetylases associated with the transcription complex, but the chromatin conformation of some tissue-specific genes is controlled by distinct upstream and downstream elements that specify the conformation of entire chromosomal domains of 30–100 kb or more (Blackwood and Kadonaga, 1998). The boundaries of these chromodomains (as well as active proximal promoters) are characterized by DNase hypersensitivity and may also be marked by CpG islands (Blackwood and Kadonaga, 1998; Lewin, 1994). Phylogenetic footprinting of MYH13 gene structure reveals the presence of such chromodomain-delimiting CpG islands (Fig. 8). Interestingly, the upstream CpG island (11 kb upstream from the human proximal promoter) overlaps with the 3′ region of an 800-nucleotide non-coding sequence that exhibits greater than 75 % sequence identity between human and mouse. This surprising homology, coupled with the apparent 'insulation' of the extraocular myosin heavy chain gene from the other members of the fast/developmental MYH cluster, suggests that MYH13 may reside in a chromodomain that occurs in a constitutively closed conformation that represses its transcription in trunk and limb muscle. The analysis of the chicken β-globin genes provides precedence for such chromatin-dependent regulation of a single member of a multigene cluster (Felsenfeld, 1993; Hebbes et al., 1994).

A model for the natural history of the skeletal muscle fast/developmental gene cluster

In Fig. 9, the phylogenetic relationships and gene order of the fast/developmental cluster are modeled as a series of spatially ordered gene duplications that can generate the fast/developmental MYH cluster (Fig. 1A). Beginning with the duplication of the precursor of the slow/cardiac and fast/developmental clusters, it proceeds to duplications that generate the extraocular MYH gene, then the developmental MYH genes and finally the adult fast MYH genes. This model for the origin of the fast/developmental MYHs differs from previous models for the origin of the fast/developmental cluster because it takes into account the critical position of MYH13 near the divergence of the slow/cardiac and fast/developmental MYH genes.

We wish to thank Mr Amar Nijagal for his assistance and the National Eye Institute (USA) for award EY11377 which funded this research.

References

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

Blackwood, E. M. and Kadonaga, J. T. (1998). Going the distance: a current view of enhancer action. *Science* **281**, 61–63.

Bouvagnet, P. F., Strehler, E. E., White, G. E., Strehler-Page, M. A., Nadal-Ginard, B. and Mahdavi, V. (1987). Multiple positive and negative 5′ regulatory elements control the cell-type-specific expression of the embryonic skeletal myosin heavy-chain gene. *Mol. Cell. Biol.* **7**, 4377–4389.

Briggs, M. and Schachat, F. (1999). Full-length cDNA sequence, predicted exon structure and phylogenetic analysis of the novel myosin in extraocular and laryngeal muscle. *Mol. Biol. Cell* **10**, 34a.

Briggs, M. and Schachat, F. (2000). Early specialization of the superfast myosin in extraocular and laryngeal muscles. *J. Exp. Biol.* **203**, 2485–2494.

Briggs, M. and Schachat, F. (2001). Extraocular myosin is localized near the innervation zone of rabbit extraocular muscle. *Biophys. J.* **(Annual Meeting Abstracts)**, 575a.

Close, R. L. and Luff, A. R. (1974). Dynamic properties of inferior rectus muscle of the rat. *J. Physiol., Lond.* **236**, 259–270.

Donoghue, M., Ernst, H., Wentworth, B., Nadal-Ginard, B. and Rosenthal, N. (1988). A muscle-specific enhancer is located at the 3′ end of the myosin light-chain 1/3 gene locus. *Genes Dev.* **2**, 1779–1790.

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

Felsenfeld, G. (1993). Chromatin structure and the expression of globin-encoding genes. *Gene* **135**, 119–124.

Firulli, A. B. and Olson, E. N. (1997). Modular regulation of muscle gene transcription: a mechanism for muscle cell diversity. *Trends Genet.* **13**, 364–369.

Giallongo, A., Venturella, S., Oliva, D., Barbieri, G., Rubino, P. and Feo, S. (1993). Structural features of the human gene for muscle-specific enolase. Differential splicing in the 5′-untranslated sequence generates two forms of mRNA. *Eur. J. Biochem.* **214**, 367–374.

Goldberg, S. J. and Shall, M. S. (1999). Motor units of extraocular muscles: recent findings. *Prog. Brain Res.* **123**, 221–232.

Gossett, L. A., Kelvin, D. J., Sternberg, E. A. and Olson, E. N. (1989). A new myocyte-specific enhancer-binding factor that recognizes a conserved element associated with multiple muscle-specific genes. *Mol. Cell. Biol.* **9**, 5022–5033.

Gupta, M., Zak, R., Libermann, T. A. and Gupta, M. P. (1998). Tissue-restricted expression of the cardiac alpha-myosin heavy chain gene is controlled by a downstream repressor element containing a palindrome of two ets-binding sites. *Mol. Cell. Biol.* **18**, 7243–7258.

Heanue, T. A., Reshef, R., Davis, R. J., Mardon, G., Oliver, G., Tomarev, S., Lassar, A. B. and Tabin, C. J. (1999). Synergistic regulation of vertebrate muscle development by Dach2, Eya2 and Six1, homologs of genes required for *Drosophila* eye formation. *Genes Dev.* **13**, 3231–3243.

Hebbes, T. R., Clayton, A. L., Thorne, A. W. and Crane-Robinson, C. (1994). Core histone hyperacetylation co-maps with generalized DNase I sensitivity in the chicken beta-globin chromosomal domain. *EMBO J.* **13**, 1823–1830.

Hoh, J. F. Y., Qin, H., Hsu, M., Yu, J., Eyre, H. and Sutherland, G. (1999). Cloning of the full-length cat superfast myosin heavy chain cDNA and mapping of its human homologue to chromosome 7q22. *J. Muscle Res. Cell Motil.* **20**, 80.

Jacoby, J., Ko, K., Weiss, C. and Rushbrook, J. I. (1990). Systematic variation in myosin expression along extraocular muscle fibers of the adult rat. *J. Muscle Res. Cell Motil.* **11**, 25–40.

Jin, J. P., Zhang, J. and Wang, J. (1995). An embryonic alternative transcription initiation site and the 5′-upstream structure of mouse cardiac troponin T gene. *Biochem. Biophys. Res. Commun.* **214**, 1168–1174.

Johnsen, O., Murphy, P., Prydz, H. and Kolsto, A. B. (1998). Interaction of the CNC-bZIP factor TCF11/LCR-F1/Nrf1 with MafG: binding-site selection and regulation of transcription. *Nucleic Acids Res.* **26**, 512–520.

Kitamura, K., Miura, H., Miyagawa-Tomita, S., Yanazawa, M., Katoh-Fukui, Y., Suzuki, R., Ohuchi, H., Suehiro, A., Motegi, Y., Nakahara, Y., Kondo, S. and Yokoyama, M. (1999). Mouse Pitx2 deficiency leads to anomalies of the ventral body wall, heart, extra- and periocular mesoderm and right pulmonary isomerism. *Development* **126**, 5749–5758.

Korn, E. D. (2000). Coevolution of head, neck and tail domains of myosin heavy chains. *Proc. Natl. Acad. Sci. USA* **97**, 12559–12564.

Lewin, B. (1994). Chromatin and gene expression: constant questions, but changing answers. *Cell* **79**, 397–406.

Li, Y., Camp, S., Rachinsky, T. L., Bongiorno, C. and Taylor, P. (1993). Promoter elements and transcriptional control of the mouse acetylcholinesterase gene. *J. Biol. Chem.* **268**, 3563–3572.

Li, Z. B., Rossmanith, G. H. and Hoh, J. F. (2000). Cross-bridge kinetics of rabbit single extraocular and limb muscle fibers. *Invest. Ophthalmol. Vis. Sci.* **41**, 3770–3774.

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

Lucas, C. A., Rughani, A. and Hoh, J. F. (1995). Expression of extraocular myosin heavy chain in rabbit laryngeal muscle. *J. Muscle Res. Cell Motil.* **16**, 368–378.

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

McLoon, L. K., Rios, L. and Wirtschafter, J. D. (1999). Complex three-dimensional patterns of myosin isoform expression: differences between and within specific extraocular muscles. *J. Muscle Res. Cell Motil.* **20**, 771–783.

Noden, D. M., Marcucio, R., Borycki, A. G. and Emerson, C. P., Jr (1999). Differentiation of avian craniofacial muscles. I. Patterns of early regulatory gene expression and myosin heavy chain synthesis. *Dev. Dyn.* **216**, 96–112.

Peachey, L. (1971). The structure of the extraocular muscle fibers of mammals. In *The Control of Eye Movements* (ed. P. Bach-y-rita, C. C. Collins and J. E. Hyde), pp. 47–66. New York: Academic Press.

Quandt, K., Frech, K., Karas, H., Wingender, E. and Werner, T. (1995). MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucleic Acids Res.* **23**, 4878–4884.

Rowlerson, A., Mascarello, F., Veggetti, A. and Carpene, E. (1983). The fibre-type composition of the first branchial arch muscles in Carnivora and Primates. *J. Muscle Res. Cell Motil.* **4**, 443–472.

Rubinstein, N. A. and Hoh, J. F. (2000). The distribution of myosin heavy chain isoforms among rat extraocular muscle fiber types. *Invest. Ophthalmol. Vis. Sci.* **41**, 3391–3398.

Saez, L. J., Gianola, K. M., McNally, E. M., Feghali, R., Eddy, R., Shows, T. B. and Leinwand, L. A. (1987). Human cardiac myosin heavy chain genes and their linkage in the genome. *Nucleic Acids Res.* **15**, 5443–5459.

Salminen, M., Lopez, S., Maire, P., Kahn, A. and Daegelen, D. (1996). Fast-muscle-specific DNA–protein interactions occurring *in vivo* at the human aldolase A M promoter are necessary for correct promoter activity in transgenic mice. *Mol. Cell. Biol.* **16**, 76–85.

Schachat, F. and Briggs, M. (1999). Identification of two patterns of exon organization in the human striated muscle myosin heavy chain genes. *Mol. Biol. Cell* **10**, 34a.

Schwartz, R. J. and Olson, E. N. (1999). Building the heart piece by piece: modularity of cis-elements regulating Nkx2-5 transcription. *Development* **126**, 4187–4192.

Sekido, R., Murai, K., Funahashi, J., Kamachi, Y., Fujisawa-Sehara, A., Nabeshima, Y. and Kondoh, H. (1994). The delta-crystallin enhancer-binding protein delta EF1 is a repressor of E2-box-mediated gene activation. *Mol. Cell. Biol.* **14**, 5692–5700.

Shimizu, N., Prior, G., Umeda, P. K. and Zak, R. (1992). cis-acting elements responsible for muscle-specific expression of the myosin heavy chain beta gene. *Nucleic Acids Res.* **20**, 1793–1799.

Spencer, R. F. and Porter, J. D. (1988). Structural organization of the extraocular muscles. In *Neuroanatomy of the Oculomotor System* (ed. J. A. Buttner-Ennever), pp. 33–79. Amsterdam: Elsevier Science Pub.

Stedman, H. H., Eller, M., Jullian, E. H., Fertels, S. H., Sarkar, S., Sylvester, J. E., Kelly, A. M. and Rubinstein, N. A. (1990). The human embryonic myosin heavy chain. Complete primary structure reveals evolutionary relationships with other developmental isoforms. *J. Biol. Chem.* **265**, 3568–3576.

Strehler, E. E., Strehler-Page, M. A., Perriard, J. C., Periasamy, M. and Nadal-Ginard, B. (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.

Swoap, S. J. (1998). *In vivo* analysis of the myosin heavy chain IIB promoter region. *Am. J. Physiol.* **274**, C681–C687.

Tajbakhsh, S., Rocancourt, D., Cossu, G. and Buckingham, M. (1997). Redefining the genetic hierarchies controlling skeletal myogenesis: Pax-3 and Myf-5 act upstream of MyoD. *Cell* **89**, 127–138.

Takechi, H., Hosokawa, N., Hirayoshi, K. and Nagata, K. (1994). Alternative 5′ splice site selection induced by heat shock. *Mol. Cell. Biol.* **14**, 567–575.

Takeda, S., North, D. L., Lakich, M. M., Russell, S. D. and Whalen, R. G. (1992). A possible regulatory role for conserved promoter motifs in an adult-specific muscle myosin gene from mouse. *J. Biol. Chem.* **267**, 16957–16967.

Tatusova, T. A. and Madden, T. L. (1999). BLAST 2 Sequences, a new tool for comparing protein and nucleotide sequences. *FEMS Microbiol. Lett.* **174**, 247–250.

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. and Higgins, D. G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **25**, 4876–4882.

Wasserman, W. W. and Fickett, J. W. (1998). Identification of regulatory regions which confer muscle-specific gene expression. *J. Mol. Biol.* **278**, 167–181.

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

Weiss, A., Schiaffino, S. and Leinwand, L. A. (1999b). Comparative sequence analysis of the complete human sarcomeric myosin heavy chain family: implications for functional diversity. *J. Mol. Biol.* **290**, 61–75.

Winters, L. M., Briggs, M. M. and Schachat, F. (1998). The human extraocular muscle myosin heavy chain gene (MYH13) maps to the cluster of fast and developmental myosin genes on chromosome 17. *Genomics* **54**, 188–189.

Yoon, S.-J., Seiler, S. H., Kucherlapati, R. and Leinwand, L. (1992). Organization of the human skeletal myosin heavy chain gene cluster. *Proc. Natl. Acad. Sci. USA* **89**, 12078–12082.