ISSN 1062-3590, Biology Bulletin, 2006, Vol. 33, No. 1, pp. 30–34. © Pleiades Publishing, Inc., 2006.
Original Russian Text © N.S. Mugue, N.D. Ozernyuk, 2006, published in Izvestiya Akademii Nauk, Seriya Biologicheskaya, 2006, No. 1, pp. 38–43.

GENETICS

Comparative Structural Analysis of Myosin Light Chains and Gene Duplication in Fish

N. S. Mugue and N. D. Ozernyuk

Kol'tsov Institute of Developmental Biology, Russian Academy of Sciences, ul. Vavilova 26, Moscow, 119991 Russia
e-mail: ozernyuk@mail.ru

Received July 8, 2005

Abstract—Origin and structure of myosin light chain (MLC) proteins have been studied by comparative analysis of fish *mlc1, mlc2*, and *mlc3* genes encoding MLC1, MLC2, and MLC3, respectively. The exon–intron structure of these genes has been analyzed in zebrafish *Danio rerio*, loach *Misgurnus fossilis*, fugu *Takifugu rubripes*, and Nile puffer *Tetraodon fahaka.* We propose that *mlc1* and *mlc3* are homologues genes originated by fish-specific whole genome duplication (paralogs). This is supported by high sequence similarity between mlc1 and mlc3 as well as by the exon–intron structure of these genes and their localization on different chromosomes. Exons 2 to 5 of *mlc1* and *mlc3* are highly conserved and have similar splicing sites. A paralog gene of *mlc2* resulting from a similar duplication event has been identified in zebrafish genome. Expression of *mlc1* paralog is limited to the larval stages of *Danio rerio* and to regenerating tissues of the adult fish. There is a possibility that the paralog of *mlc1* encodes larval myosin light chain protein (larval MLC) previously reported in a number of fish species.

DOI: 10.1134/S1062359006010043

The expression control of contractile protein, particularly myosins, in animals involves different mechanisms. These differences are observed in the case of both heavy (MHC) and light chains (MLC) of myosin. MHC isoforms are encoded by different genes in most animals, while in *Drosophila* they are generated by alternative splicing (Rozek and Davidson, 1986; George *et al*., 1989; Collier *et al*., 1990). Such pattern is also typical for MLC isoforms. In nematode *Caenorhabditis elegans, Drosophila*, and lower vertebrates, MLCs are encoded by different genes; while in higher vertebrates, they are both synthesized from separate genes (*mlc1*) and are generated by alternative splicing (*mlc1* and *mlc3*) (Hirayama *et al*., 1997; Xu *et al*., 1999, 2000; Thiebaud *et al*., 2001; Moutou *et al*., 2001).

Understanding the structural and functional organization of myosins and their origin can be approached by analysis of the role of gene duplication in different organisms. Some researchers believe that the emergence of multicellular animals and later vertebrates would be impossible without the duplication and further functional diversification of genes (Ohno, 1970; Spring, 1997; Hoegg *et al*., 2004). Complete or partial genome duplications clearly generate a great number of genes capable to provide for new functions in evolution. For instance, over 28 and 33% of the genome resulted from gene duplications in *Mycoplasma pneumoniae* and *Mycobacterium tuberculosis*, respectively (Jorden *et al*., 2001). Over 40% of human genes are represented by two or more copies (Spring, 1997; Lynch and Conery, 2000; Wang and Gu, 2000; Friedman and Hughes, 2001; Lynch, 2001; Wolfe, 2001; Gu and Huang, 2002; Samonte and Eichler, 2002).

Fishes play a particular role in the analysis of the genome duplication problem. For instance, a hypothesis of fish-specific whole-genome duplication has been proposed (Amores *et al*., 1998; Wittbrodt *et al*., 1998; Ohno, 1999; Taylor *et al*., 2001, 2003; *Genome Evolution…*, 2003). The time of gene duplication correlates with the time of teleost fish diversification (Hoegg *et al*., 2004). Such correlation was demonstrated for the *Hox* gene clusters. In fish, the duplication of several tens of genes has been demonstrated largely in zebrafish, fugu, and puffer as well as in medaka, killifish, and swordtail (Taylor *et al*., 2001). At the same time, the duplication of contractile proteins has not been analyzed in fish. In this context, different mechanisms of origin of myosins in different animals and the presence of several myosin isoforms substantiate the analysis of possible duplications of myosin genes in fish skeletal muscle and possible consequences of such duplications.

The goal of this work was to evaluate sequence similarity between *mlc1, mlc2*, and *mlc1* genes from skeletal muscle of loach, zebrafish, fugu, and Nile puffer and to comparatively analyze the exon–intron structure of individual MLC types in these fish species in order to clarify the fate of the duplicated genes in the genome.

MATERIALS AND METHODS

White (fast) skeletal muscles of adult loach *M. fossilis* were used. The isolation of total RNA from the

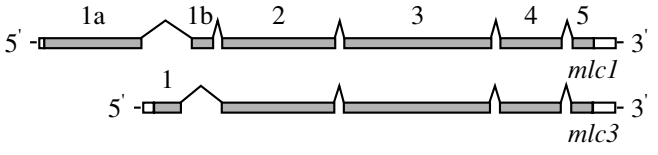

Fig. 1. Exon–intron organization of *mlc1* and *mlc3* genes from zebrafish skeletal muscle. Numbers above indicate exon numbers.

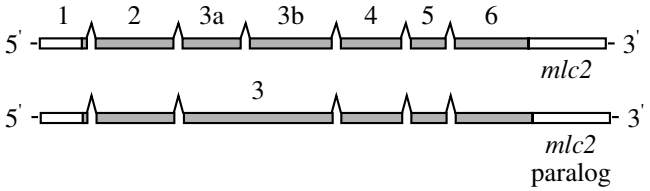

Fig. 2. Exon–intron organization of *mlc1* gene and its paralog from zebrafish skeletal muscle.

muscle tissue and of the mRNA fraction as well as cDNA synthesis, polymerase chain reaction (PCR), and analysis of PCR products were carried out as described elsewhere (Mugue *et al*., 2005).

Synthetic oligonucleotides used as PCR primers were designed from the conserved DNA sequences of *mlc1, mlc2*, and *mlc1* available for carp and zebrafish (Hirayama *et al*., 1997); their nucleotide sequences were reported elsewhere (Mugue *et al*., 2005). PCR fragments were extracted from agarose gels using a magnetic particle kit (Sileks M, Russia) according to the manufacturer's instructions. Sequencing was performed on an ABI 3100 sequencer using the same primers and a BigDye Kit v.1.1. Sequence fragments were merged and multiple alignments were generated using the DNAStar package (Lasergene, United States). The sequences have been deposited in GenBank (acc. nos. DQ068406–DQ068408).

The structure of *mlc1, mlc2*, and *mlc1* genes in zebrafish *Danio rerio*, fugu *Takifugu rubripes*, and Nile puffer *Tetraodon fahaka* was studied by the analysis of mRNA and genomic DNA sequences deposited in GenBank (www.ncbi.nlm.nih.gov).

RESULTS

Since different evolutionary mechanisms were realized to form MLCs in different animals, we compared the degree of sequence similarity between *mlc1, mlc2*, and *mlc1* genes from the skeletal muscle of loach as well as the exon–intron structure of all MLC types in different fish species to understand the origin of isoforms of these proteins.

Comparison of the primary structure of MLC genes from the skeletal muscle of loach demonstrated high similarity (77.3%) between *mlc1* and *mlc3*. Note the presence of homologous EF-hand calcium-binding domains in their structure. In contrast, sequence similarity was much lower between *mlc2* and *mlc1* or *mlc3*: 39.9 and 41.9%, respectively. High sequence similarity between *mlc1* and *mlc3* can point to their close phylogenetic relationship.

In order to analyze the degree of similarity between MLCs of other fish species (zebrafish, fugu, and puffer), we also studied the exon–intron structure of their genes. The sequences of both mRNA and genomic DNA from these species available in GenBank were analyzed. Figure 1 shows the exon–intron structure of *mlc1* and *mlc3* from zebrafish. Most of their exons have a similar structure, which agrees with a high degree of similarity between their sequences. Four out of five exons have identical structure in these genes and only the 5'-region has two exons (1a and 1b) in *mlc1* but a single exon (1) with a different structure in *mlc3*. These genes are localized on different chromosomes. A similar exon–intron structure was demonstrated for *mlc1* and *mlc3* genes of puffer. Only *mlc1* with an exon–intron structure identical to those in zebrafish and puffer has been revealed in fugu.

The exon structure of zebrafish *mlc1* significantly differed from that of *mlc1* and *mlc3* (Figs. 1, 2). Hence, a high degree of similarity between *mlc1* and *mlc3* and the similarity of their exon–intron structure confirms the phylogenetic relations between these two genes in fish species and supports their origin by gene duplication.

Analysis of *mlc1* gene structure allowed us to reveal its duplication in the genome of *D. rerio*. A paralog of this gene was localized to another chromosome (paralogs are homologous genes formed by duplication; Koonin *et al*., 2004). Four out of six exons have identical structure in these genes, while two exons (3a and 3b) in *mlc1* and one exon (3) in the *mlc1* duplicate are different (Fig. 2). Analysis of the expression of the paralog of *mlc1* in the Unigene bank (cluster Dr. 36460) demonstrated its expression limited to the embryo, embryonic heart, and regenerating fin tissue samples. Hence, this gene can be a template for the synthesis of an embryonic MLC.

In order to estimate the time of *mlc1/mlc3* duplication, we comparatively analyzed sequences of these genes in fish species representing Salmoniformes, Clupeiformes, Cypriniformes, Beloniformes, Mugiliformes, Gadiformes, and Perciformes as well as in amphibians, birds, and mammals available in GenBank (Fig. 3). The presence of two copies (*mlc1* and *mlc3*) was observed in all studied fish species suggesting that the duplication preceded the formation of the major teleost orders (Hoegg *et al*., 2004).

DISCUSSION

Presently, genome duplication is considered as one of the most important large-scale evolutionary processes (Ohno, 1970; Spring, 1997; *Genome Evolution…*, 2003; Hegg *et al*., 2004). Gene duplications

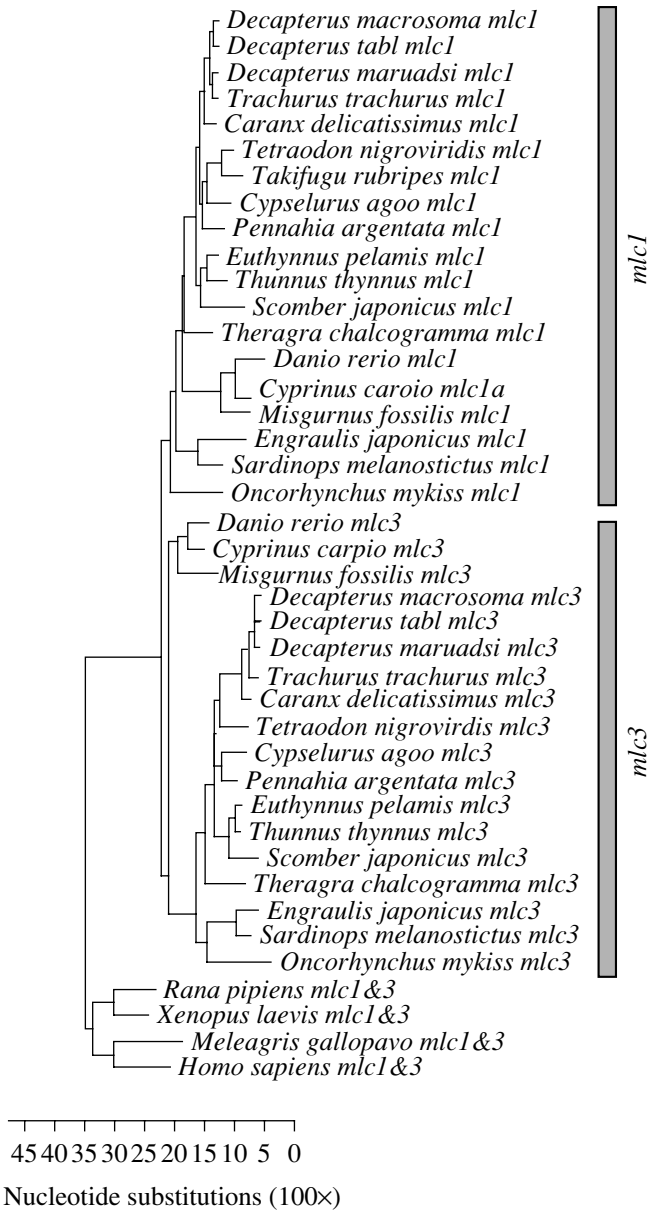

Fig. 3. Phylogenetic reconstruction based on nucleotide sequences of *mlc1* and *mlc3* genes from fish, amphibians, birds, and human. Only conserved homologous exons 2–5 were considered.

clearly introduce new adaptive functions to the organisms. This can be exemplified by antifreeze proteins in Antarctic fishes (Cheng and Chen, 1999), thermal adaptations in *Escherichia coli* (Riehle *et al*., 2001), color vision in New World primates (Dulai *et al*., 1999), etc.

Analysis of duplications of myosin genes and, particularly, MLCs is of considerable interest in terms of the actively discussed fish-specific genome duplication (Wittbrodt *et al*., 1998; Ohno, 1999; Taylor *et al*., 2001, 2003; *Genome Evolution…*, 2003), which could play a crucial role in fish diversification (Amores *et al*., 1998; Taylor *et al*., 2003; Hoegg *et al*., 2004).

The revealed high degree of sequence similarity between loach *mlc1* and *mlc3* as well as the similarity in the exon–intron organization of these MLCs in zebrafish (Fig. 1) and puffer indicate close phylogenetic relationship between these genes, which can be attributed to their duplication. Note that four homologous exons of *mlc1* and *mlc3* form an EF-hand calcium-binding domain and the differences between these genes are localized to the 5'-region. Essentially, the degree of *mlc1* similarity with two other MLC genes is significantly lower than between *mlc1* and *mlc3*; in addition, the exon–intron structure of *mlc2* notably differs from those of *mlc1* and *mlc3* (Figs. 1, 2). Hence, new calcium-containing protein of the fish muscle system MLC3 likely appeared as a result of *mlc1* gene duplication. This conclusion agrees with the views on gene duplications as a mechanism to increase the number of adaptive functions in evolution through the emergence of new proteins (Hoegg *et al*., 2004). Note in this context that fish MLC3 is a result of *mlc1* gene duplication, while MLC3 of vertebrates resulted from alternative splicing of *mlc1* (Weeds and Loweley, 1971; Frank and Weeds, 1974). Similar mechanisms of new protein generation in evolution have been demonstrated for the genes encoding synapsin (*syn*) and tissue inhibitor of metalloproteinases (*timp*) (Yu *et al*., 2003). In fugu, isoforms *syn2A* and *syn2B* were formed by the duplication of ancestor gene *syn*; while humans utilized alternative splicing of *syn2* gene.

Clearly, both gene duplication and alternative splicing represent mechanisms to increase the number of proteins, which extends the functions in evolution. This is commonly applied to closely related genes with similar functions. Kopelman *et al.* (2005) proposed that gene duplication and alternative splicing, being the main mechanisms providing for gene redundancy required for evolutionary processes, inversely correlate. This idea is further confirmed by the obtained and published data on MLC formation in fish and mammals (Weeds and Loweley, 1971; Frank and Weeds, 1974).

Another duplication was observed for zebrafish *mlc2* gene (Fig. 2). Note that mammalian MLC2, in contrast to MLC1 and MLC3, is a product of a separate gene too (Weeds and Loweley, 1971; Frank and Weeds, 1974). Amazing data were obtained concerning the pattern of tissues where the paralog of *mlc2* was expressed. Its expression in the embryo and embryonic heart suggests this gene as a candidate template for the synthesis of the embryonic MLC in fish (Mugue *et al*., 2005). The embryonic and larval forms of MLC have been previously revealed in fish (Nareiko, 1988; Focant *et al*., 1992; Ozernyuk, 2004; Ozernyuk *et al*., 2004) and higher vertebrates (Whalen *et al*., 1978, 1981; Hoh and Yeoh, 1979; Gauthier *et al*., 1982; Lowey *et al*., 1982; Takano-Ohmuro *et al*., 1982; Crow *et al*., 1983; Butler-Browne and Whalen, 1984). However, the genes encod-

ing the embryonic and larval MLC have not been identified in fish. The revealed homolog of *mlc2* gene and its expression pattern suggests it as a candidate template for the synthesis of embryonic MLC forms in fish. Similar to *mlc3*, a paralog of *mlc2* increases the number of protein products, which commonly have similar functions.

Comparative analysis of the structure of *mlc1* and *mlc3* in representatives of seven teleost orders suggests that the duplication of these genes preceded the radiation of these orders (Fig. 3). Detailed analysis of the duplication problem using the expression pattern of *fzd, sox11*, and tyrosinase gene demonstrated that the genome duplication in fish took place after the separation of Acipenseriformes and Semionotiformes from the teleost lineage, on the one hand, but before the divergence of Osteoglossiformes, on the other hand (Hoegg *et al*., 2004).

Note in conclusion that the analysis of specific consequences of gene duplication in fish suggest that, in the case of MLC1, it gave rise to a new gene and protein MLC3; while in the case of MLC2, it gave rise to a paralog of *mlc2* gene which can code for embryonic MLC.

REFERENCES

Amores, A., Force, A., Yan, Y.L., *et al.*, Zebrafish *hox* Clusters and Vertebrate Genome Evolution, *Science*, 1998, vol. 282, pp. 1711–1714.

Butler-Browne, G.S. and Whalen, R.C., Myosin Isozyme Transitions Occurring during the Postnatal Development of the Soleus Muscle, *Devel. Biol.*, 1984, vol. 102, pp. 324–332.

Cheng, C.-H. and Chen, L., Evolution of an Antifreeze Glycoprotein, *Nature*, 1999, vol. 401, pp. 443–444.

Collier, V.I., Kronert, W.A., O'Donnell, P.T., *et al.*, Alternative Myosin Hinge Regions Are Utilized in a Tissue-Specific Fashion That Correlates with Muscle Contraction Speed, *Genes Dev.*, 1990, vol. 4, pp. 885–895.

Crow, M.T., Olson, P.S., and Stockdale, F.E., Myosin Light Chain Expression during Avian Muscle Development, *J. Cell Biol.*, 1983, vol. 96, pp. 736–744.

Dulai, K.S., von Dornum, M., Mollon, J.D., *et al.*, The Evolution of Trichromatic Colour Vision by Opsin Gene Duplication in New World and Old World Primates, *Genome Res*, 1999, vol. 9, pp. 629–638.

Focant, B., Huriaux, F., Vandewalle, P., *et al.*, Myosin, Parvalbumin and Myofibril Expression in Barbel (*Barbus barbus*, L.) Lateral White Muscle during Development, *Fish Physiol. Biochem.*, 1992, vol. 10, pp. 133–143.

Frank, D. and Weeds, A.G., The Amino Acid Sequences of the Alkali Light Chains of Rabbit Skeletal-Muscle Myosin, *Eur. J. Biochem.*, 1974, vol. 44, pp. 317–334.

Friedman, R. and Hughes, A.L., Pattern and Timing of Gene Duplication in Animal Genomes, *Genome Res.*, 2001, vol. 11, pp. 1842–1847.

Gauthier, C.F., Lowey, S., Bormioli, P.A., *et al.*, Distribution and Properties of Myosin Isozymes in Developing Avian and Mammalian Skeletal Muscle Fibres, *J. Cell Biol.*, 1982, vol. 92, pp. 471–484.

Genome Evolution; Gene and Genome Duplications and the Origin of Novel Gene Functions, Meyer, A. and van de Peer, Y., Eds., Dordrecht: Kluwer Academic, 2003.

George, E.L., Ober, M.B., and Emerson, C.P., Functional Domains of the *Drosophila melanogaster* Muscle Myosin Heavy-Chain Gene Are Encoded by Alternatively Spliced Exons, *Mol. Cell. Biol.*, 1989, vol. 9, pp. 2957–2974.

Gu, X. and Huang, W., Testing the Parsimony Test of Genome Duplications: A Counter Example, *Genome Res*, 2002, vol. 12, pp. 1–2.

Hirayama, Y., Kantoh, S., Nakaya, M., *et al.*, The Two Essential Light Chains of Carp Fast Skeletal Myosin, LC1 and LC3, Are Encoded by Distinct Genes and Change Their Molar Ratio Following Temperature Acclimation, *J. Exp. Biol.*, 1997, vol. 200, pp. 693–701.

Hoegg, S., Brinkmann, H., Taylor, J.S., *et al.*, Phylogenetic Timing of the Fish-Specific Genome Duplication Correlates with the Diversification of Teleost Fish, *J. Mol. Evol.*, 2004, vol. 59, pp. 190–203.

Hoh, J.F.Y. and Yeoh, G.P.S., Rabbit Skeletal Myosin Isoenzymes from Foetal, Fast-Twitch and Slow-Twitch Muscles, *Nature*, 1979, vol. 280, pp. 321–323.

Jorden, I.K., Makarova, K.S., Spouge, J.L., *et al.*, Lineage-Specific Gene Expansions in Bacterial and Archaeal Genomes, *Genome Res.*, 2001, vol. 11, pp. 555–565.

Koonin, E., Fedorovo, N., Jackson, J., *et al.*, A Comprehensive Evolutionary Classification of Proteins Encoded in Complete Eukaryotic Genomes, *Genome Biol.*, 2004, vol. 5, no. 2.

Kopelman, N.M., Lancet, D., and Yanai, I., Alternative Splicing and Gene Duplication Are Inversely Correlated Evolutionary Mechanisms, *Nat. Gen.*, 2005, vol. 37, no. 6, pp. 588–589.

Lowey, S. Benfiels, P.A., le Blanc, D.D., *et al.*, Characterization of Myosin from Embryonic and Development: Molecular and Cellular Control, New York: Cold Spring Harbor Lab, 1982, pp. 15–24.

Lynch, M., The Molecular Natural History of the Human Genome, *Trends Ecol. Evol.*, 2001, vol. 16, pp. 420–422.

Lynch, M. and Conery, J.S., The Evolutionary Fate and Consequences of Duplicate Genes, *Science*, 2000, vol. 290, pp. 1151–1155.

Moutou, K.A., Canario, A.V.M., Mamuris, Z., *et al.*, Molecular Cloning and Sequence of *Sparus aurata* Skeletal Myosin Light Chains Expressed in White Muscle: Developmental Expression and Thyroid Regulation, *J. Exp. Biol.*, 2001, vol. 204, pp. 3009–3018.

Mugue, N.S., Tikhonov, A.V., and Ozernyuk, N.D., Ontogenetic and Phylogenetic Analysis of Myosin Light Chain Proteins from Skeletal Muscles of Loach *Misgurnus fossilis*, *Izv. Akad. Nauk. Ser. Biol.*, 2005, no. 5, pp. 573–577.

Nareiko, V.G. and Ozernyuk, N.D., Differences in the Composition of Contractile Proteins from Red and White Skeletal Muscles of Teleost Fishes, *Dokl. Akad. Nauk SSSR*, 1988, vol. 298, no. 4, pp. 1022–1024.

Ohno, S., *Evolution by Gene Duplication*, New York: Springer-Verlag, 1970.

Ohno, S., Gene Duplication and the Uniqueness of Vertebrate Genomes circa 1970–1999, *Semin. Cell Dev. Biol.,* 1999, vol. 10, pp. 517–522.

Ozernyuk, N.D., Comparative Properties of Myogenesis in Invertebrates and in Lower and Higher Vertebrates, *Ontogenez*, 2004, no. 6, pp. 441–450.

Ozernyuk, N.D., Smirnova, Yu.A., Nareiko, V.G., and Zinov'eva, R.D., Pattern of Skeletal Muscle Differentiation in Fish: Molecular and Biological Approaches, *Izv. Akad. Nauk, Ser. Biol.*, 2004, no. 3, pp. 261–268.

Riehle, M.M., Bennette, A.F., and Long, A.D., Genetic Architecture of Thermal Adaptation in *Escherichia coli, Proc. Natl. Acad. Sci. USA*, 2001, vol. 98, pp. 525–530.

Rozek, C.E. and Davidson, N., Differential Processing of RNA Transcribed from Single-Copy *Drosophila* Myosin Heavy Chain Gene Produces Four mRNA That Encode Two Polypeptides, *Proc. Natl. Acad. Sci. USA,* 1986, vol. 83, pp. 2128–2132.

Spring, J., Vertebrate Evolution by Interspecific Hybridization—Are We Polyploid?, *FEBS Lett.*, 1997, vol. 400, pp. 2–8.

Samonte, R.V. and Eichler, E., Segmental Duplications and the Evolution of the Primate Genome, *Nat. Genet. Rev.*, 2002, vol. 3, pp. 65–72.

Takano-Ohmuro, H., Obinata, T., Masaki, T., *et al.*, Changes in Myosin Isozymes during Development of Chicken Breast Muscle, *J. Biochem.*, 1982, vol. 91, pp. 11305–11309.

Taylor, J.S., van de Peer, Y., Braasch, I., *et al.*, Comparative Genomics Provides Evidence for an Ancient Genome Duplication Event in Fish, *Philos. Trans. R. Soc*, 2001, vol. 356, pp. 1661–1679.

Taylor, J.S., Braasch, I., Frickey, T., *et al.*, Genome Duplication, A Trait Shared by 22000 Species of Ray-Finned Fish, *Genome Res.*, 2003, vol. 13, pp. 382–390.

Thiebaud, P., Rescan, P.-Y., Barillot, W., *et al.*, Developmental Program Expression of Myosin Alkali Light Chain and Skeletal Actin Genes in Rainbow Trout *Oncorhynchus mykiss, Biochim. Biophys. Acta*, 2001, vol. 1519, pp. 139–142.

Wang, Y. and Gu, X., Evolution Patterns of Gene Families Generated in the Early Stages of Vertebrates, *J. Mol. Evol.*, 2000, vol. 51, pp. 88–96.

Weeds, A.G. and Lowey, S., Substructure of the Myosin Molecule. II. The Light Chain of Myosin, *J. Mol. Biol.*, 1971, vol. 61, pp. 701–725.

Whalen, R.G., Butler-Browne, G.S., and Gros, F., Identification of a Novel Form of Myosin Light Chain Present in Embryonic Muscle Tissue and Cultured Muscle Cells, *J. Mol. Biol.*, 1978, vol. 126, pp. 415–431.

Whalen, R.G., Sell, S.M., Butler-Browne, G.S., *et al.*, Three Myosin Heavy Chain Isoenzymes Appear Sequentially in Rat Muscle Development, *Nature*, 1981, vol. 292, pp. 805–809.

Wittbrodt, J., Meyer, A., and Schartl, M., More Genes in Fish?, *Bio-Essays*, 1998, vol. 20, pp. 511–515.

Wolfe, K.H., Yesterday's Polyploids and the Mystery of Diploidization, *Nat. Genet. Rev.*, 2001, vol. 2, pp. 333–341.

Xu, Y., He, J., Tian, H.L., *et al.*, Fast Skeletal Muscle-Specific Expression of a Zebrafish Myosin Light Chain 2 Gene and Characterization of Its Promoter by Direct Injection into Skeletal Muscle, *DNA Cell Biol.*, 1999, vol. 18, pp. 85–95.

Xu, Y., He, J., Lim, T.M., *et al.*, Asynchronous Activation of 10 Muscle-Specific Protein (MSP) Genes during Zebrafish Somitogenesis, *Devel. Dyn.*, 2000, vol. 219, pp. 201–215.

Yu, W.-P., Brenner, S., and Venkatesh, B., Duplication, Degeneration and Subfunctionalization of the Nested Synapsin-Timp Genes in Fugu, *Trends Genet.*, 2003, vol. 19, no. 4, pp. 180–183.