J. Fish. Sci. Technol. 11(3), 177-181, 2008

JOURNAL OF
FISHERIES SCIENCE
AND TECHNOLOGY

Notes

Evolutionary History of Two Paralogous Glyceraldehyde 3-Phosphate Dehydrogenase Genes in Teleosts

Keun-Yong Kim and Yoon Kwon Nam*

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

Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) is a key enzyme for carbohydrate metabolism in most living organisms. Recent reports and our own searches of teleost species in publicly available genomic databases have identified at least two distinct *GAPDH* genes in a given species. The two *GAPDH* genes are located on the same chromosome in teleosts, whereas they are located on the different chromosomes in mammals. Thus, we reconstructed a phylogenetic tree to better understand the evolutionary history of the *GAPDH* genes in the vertebrate lineage. Our phylogenetic analysis revealed unambiguously that the two *GAPDH* genes of teleosts are phylogenetically closely affiliated to one of the cytosolic *GAPDH* and spermatogenic *GAPDH-S* of mammals. This indicates that the two paralogous *GAPDH* genes shared a common ancestor and subsequently underwent a gene duplication event during early vertebrate evolution. However, *GAPDH-S* of teleosts showed significant differences in the polypeptide residues and tissue distribution of its mRNA transcripts from that of mammals, implying they have undergone a different history of functionalization.

Key words: Evolutionary history, *GAPDH*, *GAPDH-S*, Gene duplication, Teleost

Introduction

Glyceraldehyde 3-phosphate dehydrogenase (GAPDH; EC 1.2.1.12) is a key enzyme for carbohydrate metabolism in most living organisms. Due to their housekeeping role, GAPDHs of diverse species belonging to a wide variety of taxa possess highly conserved features in their structures and glycolytic functions. Based on this concept, the *GAPDH* gene has been used extensively as an invariant internal control for various mRNA expression assays. However, recent reports have indicated that this classic protein should no longer be considered a unifunctional glycolytic protein, but rather a multiplayer that modulates diverse cellular functions, especially those involved in the induction of apoptosis (Chuang et al., 2005; Sirover, 2005). In addition, the second type of *GAPDH* discovered in mammals (i.e., *GAPDH-S*) is sperm-specific and plays a crucial role in the normal development and fertility of male reproduction (Miki et al., 2004; Welch et al., 2006). Genetic information on *GAPDHs* is available from a number of teleost species, and bioinformatic searches of public databases indicate that at least two distinct *GAPDH* genes exist in the teleost lineage (e.g., Table 1). However, barring one recent study, the isolation and characterization of two paralogous *GAPDH* transcripts from a given fish species have not yet been obtained in detail. The only study was on a flatfish, the Senegalese sole (*Solea senegalensis*), in which two *GAPDH* mRNA species are differentially expressed in adult tissues and developing larvae (Manchado et al., 2007). Unlike in mammals, *GAPDH-S* in the Senegalese sole was not sperm-specific, although it was predicted to be orthologous to *GAPDH-S* in mammals. This suggests that functional differentiation between mammalian and teleost *GAPDH-Ss* may have occurred in their evolutionary path, although their physiological or cellular roles remain to be further explored.

Recently, we identified two distinct *GAPDH* transcripts in the barred knifejaw *Oplegnathus fasciatus*, and found that they are differentially expressed in response to various stimuli in a tissue-specific manner, implying their different and multifunctional roles (unpublished data). Here we deciphered a complete set of teleost *GAPDH* genes in public databases including GenBank (http://www.ncbi.nlm.nih.gov) and Ensembl (http://www.ensembl.org), and reconstructed a molecular phylogenetic tree to gain insights into

*Corresponding author: yoonknam@pknu.ac.kr

Table 1. Information on *GAPDH* genes for five teleosts deciphered from the genomic database Ensembl by ortholog and paralog predictions

Species	Common name	Gene	Location	Ensembl gene ID
Danio rerio	Zebrafish	*gapdh*	chromosome 16	ENSDARG00000043457
		gapdhs	chromosome 16	ENSDARG00000039914
Gasterosteus aculeatus	Stickleback	*GAPDH* (1 of 2)	chromosome groupXX	ENSGACG00000010219
		GAPDH (2 of 2)	scaffold_208	ENSGACG00000018725
		GAPDHS	chromosome groupXX	ENSGACG00000005864
Oryzias latipes	Medaka	*GAPDH*	chromosome 16	ENSORLG00000012224
		GAPDHS	chromosome 16	ENSORLG00000006033
Takifugu rubripes	Japanese pufferfish	*GAPDH*	scaffold_205	ENSTRUG00000003708
		GAPDHS	scaffold_473	ENSTRUG00000017358
Tetraodon nigroviridis	Spotted green pufferfish	no description	chromosome 8	GSTENG00015338001
		no description	chromosome 8	GSTENG00035345001

their comprehensive evolutionary history in the vertebrate lineage.

Materials and Methods

Complete nucleotide sequences encoding *GAPDH* genes for all fish species, and representative sequences of terrestrial vertebrates, arthropods, and nematodes were compiled from GenBank by a BLAST search, and from the genomic database Ensembl by ortholog and paralog predictions. Information on the *GAPDH* genes of five teleosts deciphered from Ensembl is summarized in Table 1. Nucleotide sequence alignment of the *GAPDH* genes was created based on alignments of the corresponding proteins in DAMBE (Xia and Xie, 2001).

Nucleotide sequences of the *GAPDH* genes from the same species that were clustered into the same branches were excluded in the final phylogenetic analysis to reduce the time of computation. The ascidian *Ciona intestinalis* was removed in the final phylogenetic analysis due to its high sequence divergence, resulting in the long-branch-attraction effect and collapsing the overall tree topology. A cytosolic *GapC* gene of the land plant *Zea mays* and a *GAPDH* gene of the baker's yeast *Saccharomyces cerevisiae* were used as outgroups.

The final alignment of the coding nucleotides contained a total of 1,017 positions including gaps, after excluding the proline-rich lead sequences of mammalian *GAPDH-S*s and the ambiguously aligned positions at the 5' and 3' termini. The aligned sequence matrix was subjected to neighbor-joining (NJ) and maximum parsimony (MP) analyses in PAUP* 4.0b10 (Swofford, 2002). Three nucleotide codon positions in the reading frame were differentially weighted in a ratio of 2:3:1, because of saturation of synonymous substitutions in the third codon positions (Martin et al., 1993). The NJ tree was reconstructed with the Kimura 2-parameter model. The MP tree was reconstructed using the heuristic search option with random addition of sequences (10 replicates) and tree-bisection-reconnection (TBR) branch swapping. Robustness of tree topologies was evaluated by bootstrap analyses with 1,000 pseudoreplicates (Felsenstein, 1985).

Results and Discussion

Searches in publicly available genomic databases for five teleosts (the zebrafish *Danio rerio*, stickleback *Gasterosteus aculeatus*, medaka *Oryzias latipes*, Japanese pufferfish *Takifugu rubripes*, and spotted green pufferfish *Tetraodon nigroviridis*) identified at least two distinct *GAPDH* genes (Table 1), as in the Senegalese sole (Manchado et al., 2007) and barred knifejaw (unpublished data). Exceptionally, two isoforms of *GAPDH* (1 of 2 and 2 of 2), as well as one isoform of *GAPDH-S* were predicted in the genomic database of the stickleback. *GAPDH* and *GAPDH-S* were located on the same chromosome in zebrafish (chromosome 16), medaka (chromosome 16), and spotted green pufferfish (chromosome 8). One of the *GAPDH* isoforms (*GAPDH* 1 of 2) and *GAPDH-S* were also located on the same chromosome in the stickleback (chromosome groupXX). However, the two genes in each fish species were not close neighbors. On the other hand, the chromosomal locations of *GAPDH* and *GAPDH-S* in the Japanese pufferfish are not yet available; they were located at relatively small-sized scaffolds or contigs. Although the genomic sequence data were incomplete, the chromosomal distributions of the two *GAPDH* genes in teleosts were different from those in mammals; *GAPDH* and *GAPDH-S* were located on different chromosomes in human (chromosomes 12 and 19, respectively) and

mouse (chromosomes 6 and 7, respectively). Furthermore, primate and murine species possess multiple copies of processed *GAPDH* pseudogenes, and their chromosomal distributions and clustering are quite species-specific.

To trace the evolutionary history of teleost *GAPDH* genes, we performed the phylogenetic analysis together with corresponding homologs from metazoans represented by nematodes, arthropods (crustaceans and insects), and terrestrial vertebrates (birds and mammals). The two phylogenetic methods of NJ and MP used for tree reconstruction produced similar tree topologies, and only the NJ tree is shown in Fig. 1. Our phylogenetic tree revealed that metazoans were recovered as the monophyletic group with 85 and 92% bootstrap values with respect to outgroups (i.e., the land plant *Z. mays* and baker's yeast *S. cerevisiae*). Within this metazoan clade, three nematodes and arthropods composed of crustaceans and insects formed monophyletic groups, giving rise to the monophyletic vertebrate lineage with 72 and 62% bootstrap values. Thereafter, all vertebrates were phylogenetically subdivided into two distinct monophyletic groups according to types of *GAHDH* genes with strong bootstrap supports.

In the *GAPDH* clade, terrestrial vertebrates and teleosts further ramified into two distinct monophyletic branches. Within the terrestrial vertebrate branch, two *Xenopus* species (Amphibia) occupied the basal position, giving rise to groups of Amniota composed of birds and mammals. Meanwhile, in the teleost branch, the stickleback (*GAPDH* 2 of 2) occupied the most basal positions followed by the monotypic Japanese eel *Anguilla japonica* and the monophyletic group of the other teleosts. Tree topologies within this clade were not well resolved except the branches of the orders Cypriniformes, Salmoniformes, Pleuronectiformes, and Tetraodontiformes.

The *GAPDH-S* clade also bifurcated into monophyletic branches of terrestrial vertebrates and teleosts. This phylogenetic relationship strongly suggested that the second type of *GAPDH* of teleosts is orthologous to the spermatogenic *GAPDH-S* of mammals, despite the absence of the characteristic proline-rich lead sequence at the N-terminus. The monophyletic *GAPDH-S* clade of teleosts was supported by 100% bootstrap values. Within this teleost clade, the cyprinid *D. rerio* and two salmonoid *Oncorhynchus* species were recovered as a monophyly with high bootstrap supports, and they were phylogenetically separated from nine fish species of the Percomorpha, forming a monophyletic group.

Overall, our phylogenetic tree revealed the monophyletic nature of two paralogous *GAPDH* genes of all vertebrates. In addition, *GAPDH* of teleosts and terrestrial vertebrates, and *GAPDH-S* of teleosts and mammals each formed distinct monophyletic groups. This result strongly suggests that the *GAPDH* and *GAPDH-S* genes originated from an early common vertebrate ancestor, probably by a whole-genome duplication event. Gene duplication, as well as endosymbiotic gene transfer of *GAPDH*s, have frequently occurred across diverse eubacterial and eukaryotic lineages (Martin et al., 1993; Figge et al., 1999; Liaud et al., 2000; Petersen et al., 2003), of which gene diversity might donate multifunctional activities to *GAPDH* genes barring the classic glycolytic-glycongeogenetic function (Sirover, 1999). Gene duplication events of *GAPDH*s within each phylum or division are also quite prominent in fruit fly (Tso et al., 1985), nematode (Huang et al., 1989), yeast (Holland et al., 1983), and land plant (Petersen et al., 2003).

While GAPDH is located in the cytosol, GAPDH-S specifically localizes and functions at the principal piece of a sperm flagellum with the help of the proline-rich lead sequence at the N-terminus in mammals (Miki et al., 2004; Welch et al., 2006). However, no GAPDH-S proteins of teleosts are known to contain this transit peptide, despite their close phylogenetic affiliation to their mammalian homologs. This implies that the transit peptide of *GAPDH-S* was acquired specifically during mammalian evolution. Similarly, eukaryotic organelle-targeting *GAPDH* genes [e.g., *GapA/B*, *GapC-I*, *GapC-III*; see Liaud et al. (2000) for nomenclature] have independently acquired transit peptides in the N-terminus, which are necessary for reimportation into the organelles of their origins (i.e., mitochondria and plastids; Liaud et al., 2000; Petersen et al., 2003). Given that the whole-genome duplication additionally occurred in the Actinopterygii (ray-finned fishes; Blomme et al., 2006; Steinke et al., 2006), there is a strong possibility for the presence of more divergent *GAPDH* or *GAPDH-S* isoforms, including one specifically targeting to fish testis with the transit peptide. For example, there are two divergent isoforms of *GAPDH* in the stickleback (Table 1 and Fig. 1).

In summary, bioinformatic searches in publicly available genomic databases revealed that *GAPDH* and *GAPDH-S* are located on the same chromosome in teleosts, whereas they are located on the different chromosomes in mammals. The phylogenetic analysis revealed unambiguously that the two distinct *GAPDH* genes of teleosts are highly homologous to the cytosolic *GAPDH* and spermatogenic *GAPDH-S* of mammals. The two paralogous *GAPDH* genes shared

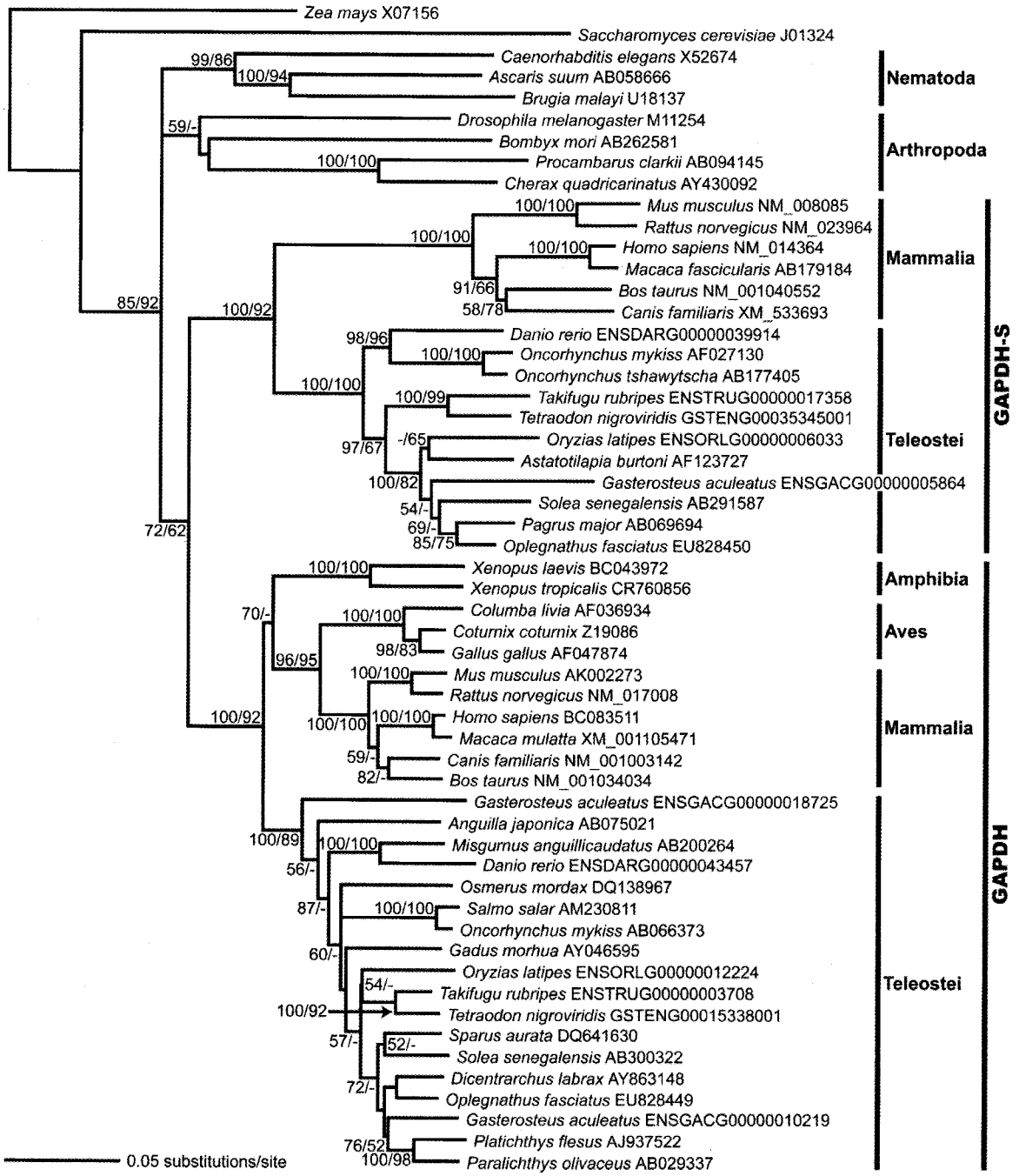

Fig. 1. The neighbor-joining tree inferred from coding nucleotide sequences of *GAPDH* genes. Representative metazoan taxa are included for the comparative phylogenetic analysis. The land plant *Zea mays* and baker's yeast *Saccharomyces cerevisiae* are used as outgroups. Numbers at each branch node indicate bootstrap values above 50% in neighbor-joining and maximum parsimony methods, respectively. Taxonomic placements and types of *GAPDH* genes are indicated in each clade.

a common ancestor during early vertebrate evolution and thereafter underwent a gene duplication event with functionalization. However, considering the previous report on the wide tissue distribution of *GAPDH-S* mRNA transcripts in the stickleback, rather than being sperm-specific as in mammals (Manchado et al., 2007), extensive expression analysis should be challenged in the future study in order to better understand the physiological roles of *GAPDH-S* in teleosts.

Acknowledgements

This study was supported by the funds from the Korea Institute of Marine Science & Technology Promotion.

References

Blomme, T., K. Vandepoele, S. De Bodt, C. Simillion, S. Maere and Y. Van de Peer. 2006. The gain and loss of genes during 600 million years of vertebrate evolution. Genome Biol., 7, R43.

Chuang, D.M., C. Hough and V.V. Senatorov. 2005. Glyceraldehyde 3-phosphate dehydrogenase, apoptosis, and neurodegenerative diseases. Annu. Rev. Pharmacol. Toxicol., 45, 269-290.

Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution, 39, 783-791.

Figge, R.M., M. Schubert, H. Brinkmann and R. Cerff. 1999. Glyceraldehyde 3-phosphate dehydrogenase gene diversity in eubacteria and eukaryotes: evidence for intra- and inter-kingdom gene transfer. Mol. Biol. Evol., 16, 429-440.

Holland, J.P., L. Labieniec, C. Swimmer and M.J. Holland. 1983. Homologous nucleotide sequences at the 5' termini of messenger RNAs synthesized from the yeast enolase and glyceraldehyde-3-phosphate dehydrogenase gene families: the primary structure of a third yeast glyceraldehydes 3-phosphate dehydrogenase gene. J. Biol. Chem., 258, 5291-5299.

Huang, X.-Y., L.A.M. Barrios, P. Vonkhorporn, S. Honda, D.G. Albertson and R.M. Hecht. 1989. Genomic organization of the glyceraldehydes 3-phosphate dehydrogenase gene family of *Caenorhabditis elegans*. J. Mol. Biol., 206, 411-424.

Liaud, M.-F., C. Lichtlé, K. Apt, W. Martin and R. Cerff. 2000. Compartment-specific isoforms of TPI and GAPDH are imported into diatom mitochondria as a fusion protein: evidence in favor of a mitochondrial origin of the eukaryotic glycolytic pathway. Mol. Biol. Evol., 17, 213-223.

Manchado, M., C. Infante, E. Asensio and P. Canavate. 2007. Differential gene expression and dependence on thyroid hormones of two glyceraldehydes 3-phosphate dehydrogenases in the flatfish Senegalese sole (*Solea senegalensis* Kaup). Gene, 400, 1-8.

Martin, W., H. Brinkmann, C. Savonna and R. Cerff. 1993. Evidence for a chimeric nature of nuclear genomes: cubacterial origin of eukaryotic glyceraldehydes 3-phosphate dehydrogenase genes. Proc. Natl. Acad. Sci. USA, 90, 8692-8696.

Miki, K., W. Qu, E.H. Goulding, W.D. Willis, D.O. Bunch, L.F. Strader, S.D. Perreault, E.M. Eddy and D.A. O'Brien. 2004. Glyceraldehyde 3-phosphate dehydrogenase-S, a sperm-specific glycolytic enzyme, is required for sperm motility and male fertility. Proc. Natl. Acad. Sci. USA, 101, 16501-16506.

Petersen, J., H. Brinkmann and R. Cerff. 2003. Origin, evolution, and metabolic role of a novel glycolytic GAPDH enzyme recruited by land plant plastids. J. Mol. Evol., 57, 16-26.

Sirover, M.A. 1999. New insights into an old protein: the functional diversity of mammalian glyceraldehydes 3-phosphate dehydrogenase. Biochim. Biophys. Acta, 1432, 159-184.

Sirover, M.A. 2005. New nuclear functions of the glycolytic protein, glyceraldehydes 3-phosphate dehydrogenase, in mammalian cells. J. Cell. Biochem., 95, 45-52.

Steinke, D., S. Hoegg, H. Brinkmann and A. Meyer. 2006. Three rounds (1R/2R/3R) of genome duplications and the evolution of the glycolytic pathway in vertebrates. BMC Biol., 4, 16.

Swofford, D.L. 2002. PAUP: Phylogenetic Analysis Using Parsimony (and Other Methods). Ver 4.0b10, Sinauer Associates Inc., Sunderland, MA, USA.

Tso, J.Y., X.H. Sun and R. Ray. 1985. Structure of two unlinked *Drosophila melanogaster* glyceraldehydes 3-phosphate dehydrogenase genes. J. Biol. Chem., 260, 8220-8228.

Welch, J.E., R.R. Barbee, P.L. Magyar, D.O. Bunch and D.A. O'Brien. 2006. Expression of the spermatogenic cell-specific glyceraldehyde 3-phosphate dehydrogenase (GAPDS) in rat testis. Mol. Reprod. Dev., 73, 1052-1060.

Xia, X. and Z. Xie. 2001. DAMBE: Data analysis in molecular biology and evolution. J. Hered., 92, 371-373.

(Received June 2008, Accepted September 2008)