MOLECULAR AND FUNCTIONAL CHARACTERIZATION OF TWO
DISTINCT IGF BINDING PROTEIN-6 GENES IN ZEBRAFISH

Xianlei Wang¹, Ling Lu¹, Yun Li¹, Mingyu Li¹, Chen Chen¹, Qiang Feng¹,
Chunyang Zhang¹, and Cunming Duan¹,²*

1 Laboratory of Molecular Medicine, School of Medicine and Pharmacy, Ocean
University of China, 5 Yushan Road, Qingdao, Shandong 266003, China;
and
2 Department of Molecular, Cellular, and Developmental Biology, University of
Michigan, Ann Arbor, MI 48109, USA

* Corresponding author to: Cunming Duan, PhD

Fax: 734-647-0884
E-mail: cduan@umich.edu

Running head: Duplication and divergence of the IGFBP-6 gene in zebrafish
Abstract

Insulin-like growth factor binding proteins (IGFBPs) are high affinity binding partners for IGFs and play important roles in growth and development by binding to and modulating IGF activities. In this study, we have identified and characterized two functional IGFBP-6 genes in zebrafish. Structural, phylogenetic, and comparative genomic analyses indicate that they are co-orthologs of the human IGFBP-6 gene. To gain insight into how the duplicated genes may have evolved through partitioning of ancestral functions, gene expression and functional studies were carried out. In adult fish, IGFBP-6a mRNA was most abundantly expressed in the muscle. The levels of IGFBP-6a mRNA in non-muscle tissues were very low or barely detectable. In comparison, the levels of IGFBP-6b mRNA were high in the brain, heart, and muscle, but very low or undetectable in other adult tissues. During embryogenesis, the IGFBP-6a mRNA levels were relatively low. The IGFBP-6b mRNA levels were low during the initial 48 h. They became significantly higher at 72 and 96 hpf. Overexpression of zebrafish IGFBP-6a and -6b caused a similar degree of reduction in body size and developmental rates. No notable effects were observed on cell fate or patterning in these transgenic fish. These data suggest that the duplicated igfbp-6 encodes two functionally equivalent proteins, but they have evolved distinct spatial and temporal expression patterns. These findings are consistent with the notion of an additional gene duplication event in teleost fish and have provided novel insight into the structural and functional evolution of the IGFBP gene family.

Keywords: Gene duplication, insulin-like growth factor, binding protein, embryo, growth
Introduction

Insulin-like growth factors (IGFs) are evolutionarily conserved peptides that play fundamental roles in regulating development, growth, reproduction, and longevity. The biological actions of IGFs are mediated by the IGF-1 receptor, which is a member of the receptor tyrosine kinase family (20). In addition to the IGF ligands and receptors, there exists another important component of the IGF system, the IGF binding proteins (IGFBPs). Six distinct IGFBPs, designated as IGFBP-1 to -6, have been isolated and cloned from human and other species and each represents an individual gene product (4, 9). These proteins share relatively high amino acid sequence similarity, but each has distinct structural and biochemical properties and each is subject to differential developmental and hormonal regulation. These IGFBPs can act as carrier proteins in the bloodstream and can control the efflux of IGFs from the vascular space. The circulating IGF/IGFBP complexes prolong the half-life of IGFs and buffer their acute hypoglycemic effects. Since IGFBPs bind to IGFs with either equal or higher affinities than to the IGF receptors, locally expressed IGFBPs provide a means of localizing IGFs in specific cells and alter their biological activity by modulating their interaction with the IGF receptors. IGFBPs have been shown to both enhance and inhibit IGF-mediated cellular proliferation, differentiation, and apoptosis in a variety of mammalian cell cultures. IGF-independent actions have also been demonstrated for some IGFBPs (4, 9).

IGFBP-6 is unique among the IGFBPs in that it has a 50-fold higher affinity for IGF-2 than IGF-1 (1). Because of this binding preference, IGFBP-6 has been shown to inhibit IGF-2-induced cell proliferation, differentiation, and survival, but it has little effect on IGF-1 actions (1). Ligand-independent actions of IGFBP-6 on cell
migration have been documented, although the underlying molecular mechanism(s) remains incompletely understood (10). A recent study has reported that IGFBP-6 has a functional nuclear localization sequence that mediates its nuclear import in cultured cells (15). Overexpression of IGFBP-6 inhibits rhabdomyosarcoma xenograft and neuroblastoma xenograft growth in vivo (11, 12). Transgenic mice overexpressing IGFBP-6 in the central nervous system resulted in postnatal growth, brain development, and reproduction abnormalities (2). Overexpression of IGFBP-6 in the brain resulted in dysregulation of energy homeostasis in mice (3).

Recent studies using the zebrafish model have shed valuable insight into the in vivo functions of several IGFBPs, including IGFBP-1, 2, and -3 (16, 18, 21, 33). Zebrafish embryos develop in water, thus eliminating the complication of maternal compensation. Fast developing and transparent zebrafish embryos make it easy to perform genetic manipulations and observe the phenotypic changes in organ formation in real time. Importantly, major components of the zebrafish IGF signaling pathway, including IGFs, IGF receptors, IGFBPs, and intracellular signal transduction network components, have been characterized and are highly conserved (7, 32). Furthermore, many teleost fish, including zebrafish, may have experienced an additional genome wide duplication event (25, 29). As a result, fish often have two co-orthologs in contrast to a single copy gene in humans and mammals. The retention of a particular pair of genes provided a unique opportunity to gain insights into the impact of evolution on gene structure and function.

The goal of this study was to identify and characterize the zebrafish IGFBP-6 gene(s), determine their spatial and temporal expression patterns, and elucidate their biological actions. A comparison of the structure of the IGFBP-6 gene from different vertebrate species will provide novel insight into our understanding of the evolution
of the IGFBP gene family. An understanding of the zebrafish IGFBP-6 gene expression and function in early development will not only shed new insights into the developmental roles of this unique IGFBP, but also contribute to our knowledge of IGFBP-6 physiology in vertebrates in general.

**Materials and Methods**

**Animals**

Zebrafish (*Danio rerio*) were kept at 28°C under a 14 h/10 h light/dark photoperiod and were fed twice daily. Zebrafish embryos were obtained from natural breeding, and were kept at 28.5°C in embryo rearing solution (26, 31) supplemented with 0.003% (weight/volume) 2-phenylthiourea (Sigma, St. Louis, MO, USA). Embryos were staged following standard criteria (19). At each stage they were fixed in 4% (w/v) paraformaldehyde (PFA, Sigma, St. Louis, MO, USA) in 1×PBS and stored at -20°C in 100% methanol or frozen in liquid nitrogen and stored at -80°C for further analysis. All experiments were conducted at Ocean University of China with approved protocols. These protocols also met the guidelines established by the University of Michigan Committee on the Use and Care of Animals.

**Molecular cloning and sequence analysis**

Using human IGFBP-6 amino acid sequence as a query, we searched the zebrafish genome database (http://www.ensembl.org/Danio_rerio/blastview) by BLASTX and
BLASTN for candidate sequences. Two zebrafish IGFBP-6 like sequences were found. Their full-length cDNAs were determined by 5’- and 3’- rapid amplification of cDNA ends (RACE) as described below. Amino acid sequences were aligned by CLUSTALW using the GCG program FASTA. Best-fit model of protein evolution was obtained by Prottest (http://darwin.uvigo.es/software/prottest.html). Gene structure was analyzed using the Genscan program (http://genes.mit.edu/GENSCAN.html).

Phylogenetic analysis was generated by the Neighbor-Joining method using the MEGA 3.1 software (Biodesign Ins., Tempe, USA). The reliability of the estimated tree was evaluated by the bootstrap method with 5000 replications. We also constructed a phylogenetic tree using the Maximum Likelihood (ML) method with the PhyML V2.4.4 program (13). Synteny analysis was carried out based on Danio rerio Zv6 (http://www.ensembl.org/Danio_rerio/index.html) and Homo sapiens Build 36.3 (http://www.ensembl.org/Homo_sapiens/index.html) and from zebrafish and human synteny map. Genomic structure was determined by comparing the full-length cDNA sequence and the zebrafish genome sequence (http://www.ensembl.org/Danio_rerio/index.html).

For RACE and RT-PCR, total RNA was isolated from adult zebrafish and embryos using TRIzol Reagent (Invitrogen, Carlsbad, CA, USA). One μg RNA was converted into cDNA using M-MLV (Promega, Madison, WI, USA) with oligo(dT)18 primer (Sangon, Shanghai, China). The 5’ and 3’ RACE were performed using the SMART RACE cDNA Amplification kit (Clontech, Mountain View, CA, USA).
following the manufacturer’s instructions. Nested primers used in the IGFBP-6a 5’ RACE were 5’-TGAAGGACTGTAAGAGCTTGTCG-3’ and 5’-CGCTGCTACATGAGATTGCATC-3’; Nested primers used in the IGFBP-6a 3’RACE were 5’-AGCACCAACTACCTCCACTGAACAT-3’ and 5’-TGTCTCTGCCTATTGCTAGCTC-3’; Nested primers used in the IGFBP-6b 5’RACE were 5’-GTCTTGACCAAGCTGAAATCATCC-3’ and 5’-CAAGACACTGTTGACGCTACG-3’; and Nested primers used in the IGFBP-6b 3’RACE were 5’-CCAGAAATAGCAGCCAGCACGTAC-3’ and 5’-AGAACCTGAGACTGGCAAGAT-3’.

Primers used in confirming the extra exon, Extron 5, of the igfbp-6b were 5’-GGATGATTTCAGCTTGGTGCAAGAC-3’ and 5’-AGCGGCATAAAATACATCCAGTTGG-3’.

Quantitative real-time RT-PCR (qRT-PCR) and whole mount in situ hybridization (WISH)

One μg RNA was converted into cDNA using M-MLV and oligo(dT)18 primer. qRT-PCR was carried out on iCycler iQ Multicolor real-time PCR detection system (Bio-Rad, Hercules, CA) according to the manufacturer’s instructions. Primer sequences for qRT-PCR were β-actin: 5’-ACAGGGAAAAGATGACACAG-3’ and 5’-AGAGTCCATCACGATACCAG-3’; IGFBP-6a: 5’-CTCCGCTACAGACTATGA-3’ and 5’-GTCCTGAACCGAGTGAAAT-3’; and
IGFBP-6b: 5’-CCACATCTTCACACAGTGG-3’ and 5’-ACCGACACTGCTTTTTCCTG-3’. qPCR was performed using SYBR Green PCR Master Mix (Applied Biosystems, Foster City, CA). β-actin amplification efficiency was similar to that of IGFBP-6a and IGFBP-6b. Results were analyzed using the $2^{-\Delta\Delta T}$ method (22). Each experiment was repeated in triplicate.

For WISH, embryos were fixed in 4% PFA at 24 hours post fertilization (hpf). Digoxigenin-labeled RNA probes were synthesized in vitro using T7 or T3 RNA polymerase (Promega, Madison, WI, USA). The following probes were used: Egr2b, Emx1, MyoD and RxI. WISH protocols were essentially the same as described in Li et al. (21). Images of zebrafish embryos were recorded with a Nikon DS-U1 digital camera mounted to a Nikon SMZ1500 Fluorescence stereomicroscope (Melville, NY, USA).

Construction of expression plasmids

For in vivo functional studies, several IGFBP-6a and -6b expression constructs were engineered. First, DNA fragments containing the entire ORF of zebrafish IGFBP-6a or -6b, in which the stop codon was mutated, was generated by PCR (IGFBP-6a: forward primer 5’-CCATCGATGCCACCATGTCTCTGCCTCATTTGCT-3’ and reverse primer 5’-CGGAATTCCAGCGCTGCTACATGAGATTGAT-3’; IGFBP-6b: forward primer 5’-CCATCGATGCCACCATGTCTTTCCTTTCCAACTT-3’ and reverse
primer 5’-CGGAATTCCCTCCGGTCACATGGTAAAAT-3’). The amplified products were digested with EcoR I and Cla I and subcloned into the pCS2-eGFP vector (21), resulting in pCS2+IGFBP-6a:eGFP and pCS2+IGFBP-6b:eGFP, respectively. Both constructs were verified by DNA sequencing.

Microinjection experiment

For mRNA synthesis, the plasmids were linearized by digestion with the appropriate restriction enzymes. Capped RNAs were synthesized using a SP6 Message Machine kit (Ambion, Austin, TX, USA) using the linearized plasmid DNA as a template. The mRNA was diluted to 0.4 μg/μl with Phenol red (Sigma, St. Louis, MO, USA) and injected with a volume of 1.5 nl into 1-2 cell stage embryos as previously reported (33). Each experiment carried out with two controls, wild type (no injection) and GFP mRNA injected embryos. After injection, embryos were kept in embryo rearing solution and kept at 28.5°C. Body length was quantified with MB-Ruler 5.1 software (Markus Bader, Germany), measuring the curvilinear distance from the forebrain boundary to the tail. Somite numbers were counted manually.

Western immunoblot

Thirty embryos from each treatment group were dechorionated, deyolked, and homogenized in 20 μl of RIPA buffer (50 mM Tris-HCl, 150 mM NaCl, 2 mM
EGTA, 0.1% Triton X-100, pH 7.5) containing 10 μg/ml aprotinin, 10 μg/ml leupeptin, 10 μg/ml pepstatin, and 100 mM PMSF. The homogenates were centrifuged to pellet cellular debris and the supernatant was retained. Each sample was subjected to SDS-PAGE (12.5%) and transferred to Immobilon-P membrane (Millipore Corp., Billerica, MA, USA). Western immunoblot analysis was performed as described previously (6) using a GFP antibody at a 1:5000 dilution (Torrey Pines Biolabs, Inc., Houston, TX, USA).

Statistics

All the data were expressed as the mean ± standard error (S.E.M.). Differences among groups were analyzed by one-way analysis of variance (ANOVA) followed by Sidak's post-hoc test for multiple comparisons (SPSS, SPSS Inc., Chicago, USA). The criterion for statistical significance was $p < 0.05$.

Results

Identification and cloning of two igfbp-6 genes in zebrafish

We searched the zebrafish GenBank database using the human IGFBP-6 sequence (Genbank accession number: CAA07346.1) as a query and found two distinct IGFBP-6 like proteins (Genbank accession number: XM_679868 and XM_688133). These two predicted proteins both displayed significant sequence identities to human IGFBP-6, but they are clearly distinct from each other. For reasons described
hereafter, we named these two genes as *igfbp-6a* and *igfbp-6b*.

Using 5’- and 3’-RACE, their full-length cDNAs were determined. The full-length IGFBP-6a cDNA is 1249 bp (GenBank accession no. EU005245). It contains an ORF of 585 bp encoding a protein of 194 a.a., a 108 bp 5′-UTR, and a 556 bp 3′-UTR. There is a typical polyadenylation (poly(A)) signal sequence AATAAA and 30 bp of poly(A) tail at the end of the 3’ UTR. The encoded protein has a putative signal peptide of 24 a.a. and a mature protein of 170 a.a. The full-length IGFBP-6b cDNA is 1266 bp and it contains an ORF of 612 bp, which encodes a protein of 203 a.a. (GenBank accession no. EU005246). The encoded protein has a 26 a.a. signal peptide and a mature protein of 177 a.a. The 5’-UTR is 159 bp and 3′-UTR is 495 bp, including a typical poly(A) signal sequence AATAAA and 26 bp of poly(A) tail.

Comparison of the primary sequence of zebrafish IGFBP-6a and -6b with those of other known vertebrate IGFBPs revealed that they share the highest sequence identity to that of IGFBP-6. These two zebrafish IGFBPs are clearly distinct to the known zebrafish IGFBPs. Zebrafish IGFBP-6a is only 26%, 27%, 34%, and 34% identical to zebrafish IGFBP-1s, IGFBP-2, IGFBP-3 and IGFBP-5 at the amino acid level (Table 1). Zebrafish IGFBP-6b shares 24-28%, 25%, 35%, and 32% sequence identity to that of zebrafish IGFBP-1s, IGFBP-2, IGFBP-3, and IGFBP-5, respectively. Among the known teleost IGFBPs, the IGFBP-6a shares 49% identity to that of rainbow trout and Atlantic salmon IGFBP-6. The IGFBP-6b has 76% and 75% identity to that of trout and salmon IGFBP-6. The two zebrafish IGFBP-6s show comparable sequence identities to human IGFBP-5 and -6 (37~32%). Their sequence identities to human IGFBP-1 and -2 are much lower (Table 1).

While all mammalian IGFBP-6 have 10 cysteine residues in the N-domain, both zebrafish IGFBP-6s only have 6 cysteine residues in this region. Zebrafish IGFBP-6a
and IGFBP-6b also have 6 cysteine residues in the conserved C-terminal domain, and 1 in the signal peptide (Fig. 1A). Like mammalian IGFBP-6, both zebrafish IGFBP-6s have a thyroglobulin type 1 motif and a CWCV motif in the C-terminal domain, but they lack the GCGCCXXC motif present in IGFBP-1～5. Zebrafish IGFBP-6a has an estimated molecular mass of 18.7 kDa. IGFBP-6b has an estimated molecular mass of 19.5 kDa. There are 5 O-glycosylation sites in the L-domain region of human IGFBP-6 (24), Analysis using an online program (http://ogpet.utep.edu/OGPET/), which can predict all five O-glycosylation sites in the L-domain region of human IGFBP-6, failed to find any consensus O-glycosylation sites in zebrafish IGFBP-6a and -6b. Analysis using the NetOGlyc 3.1 Server (http://www.cbs.dtu.dk/services/NetOGlyc/), however, suggests that there are several potential O-glycosylation sites: Thr97, Thr105, Thr111, and Thr112 in zebrafish IGFBP-6a and Thr57, Thr105, Ser106, Thr108, Thr114, Ser119, and Ser121 in IGFBP-6b. There is no consensus N-glycosylation site in mature IGFBP-6a and -6b.

We next performed phylogenetic analysis using the full-length amino acid sequences of known vertebrate IGFBPs. As shown in Fig. 1B, a phylogenetic tree generated by the Neighbor-Joining method grouped the two newly identified proteins into the IGFBP-6 clade with high bootstrap value, indicating that they are indeed co-orthologs of human IGFBP-6 (Fig. 1B). Similar results were obtained using the Maximum Likelihood method (Supplemental Fig. 1). In agreement with a recent report (26), our phylogenetic analysis indicated that the previously reported rainbow trout and salmon IGFBP-3 are grouped into the IGFBP-2 clade rather than in the IGFBP-3 clade.
The two zebrafish IGFBP-6 genes have similar genomic structure and show high synteny to the human IGFBP-6 gene

The genomic structure of zebrafish \textit{igfbp-6a} and \textit{igfbp-6b} are shown in Fig. 2A. Like their mammalian counterparts (8), zebrafish \textit{igfbp-6a} contains 4 exons and 3 introns and spans 6.1 kb (Fig. 2A). Exon 1 contains the 5’ UTR, signal peptide, all of the N-domain, and part of L-domain. Exon 2 is very small and it encodes part of the L-domain. Exon 3 and 4 contain the C domain and the 3’ UTR. The zebrafish \textit{igfbp-6b} is approximately 3.9 kb. Comparison of the exon lengths of the two zebrafish \textit{igfbp-6} genes indicates a high degree of conservation. The zebrafish \textit{igfbp-6b}, however, has an extra exon in its 3’ end, which encodes a portion of the 3’ UTR (Fig. 2A).

To explore the possible synteny relationship among the zebrafish \textit{igfbp-6}s and the human \textit{IGFBP-6}, we analyzed genes surrounding the \textit{igfbp-6} loci in the zebrafish and human genomes. While \textit{igfbp-6a} is located to linkage group (LG) 11, \textit{igfbp-6b} is mapped to LG 23 (Fig. 2B). The human IGFBP-6 gene resides on chromosome 12 (8). Further analysis of the surrounding regions revealed that \textit{igfbp-6a} and 12 other zebrafish genes (CYP27B1, NPFF1, HNRNPA1, PIP5K2C, DGKA, TMEM 4, SMUG 1, HOXC-6b, HOXC-11b, HOXC-12b, HOXC-13b and RARG) located on LG11 have orthologs on human chromosome 12 (8). Similarly, in the proximity of the \textit{igfbp-6b} locus, 14 other zebrafish genes (HOXC-4a, HOXC-5a, HOXC-6a, HOXC-8a, HOXC-9a, HOXC-10a, HOXC-11a, HOXC-13a, PIP5K2C, ADMR, ZBTB 39, CSAD, RARG, SILV b) have their orthologs located near the human \textit{IGFBP-6} locus
on chromosome 12. This conserved synteny strongly suggests that *igfbp-6a* and *igfbp-6b* are co-orthologs of the human IGFBP-6 gene.

*The two igfbp-6 genes show distinct temporal and spatial expression patterns*

The temporal and spatial expression patterns of *igfbp-6a* and *-6b* were determined by qRT-PCR and are shown in Fig. 3. Both IGFBP-6a and IGFBP-6b transcripts were detectable by qRT-PCR throughout embryogenesis and the larval stages, albeit at low levels (Fig. 3A). The IGFBP-6a mRNA levels were relatively constant between 24 to 96 hpf. The IGFBP-6b mRNA levels were low during the initial 48 h. They became significantly higher at 72 and 96 hpf. Whole mounted *in situ* hybridization analysis failed to show any specific hybridization signal (data not shown) despite considerable efforts, probably due to the relatively low levels of expression of these genes during early development. In the adult stage, IGFBP-6a mRNA was most abundantly expressed in the muscle (Fig. 3B). The levels of IGFBP-6a mRNA were very low in the gill and testis, approximately 10 fold lower than those of muscle. The IGFBP-6a mRNA levels were barely detectable in brain, eye, heart, stomach + intestine, kidney, liver, ovary, and kidney (Fig. 3B). In comparison, the levels of IGFBP-6b mRNA were very high in the brain, heart, and muscle. IGFBP-6b mRNA was very low or barely detectable in other tissues examined (Fig. 3B).

*Overexpression of IGFBP-6a and IGFBP-6b inhibits growth and development*
To determine and compare the biological activities of zebrafish IGFBP-6a and -6b, capped mRNAs encoding zebrafish IGFBP-6a and -6b were generated and introduced into zebrafish embryos by microinjection. The successful expression of IGFBP-6a and -6b were verified by examining the GFP signal under fluorescent microscopy (Fig. 4A) and by Western immunoblot (Fig. 4B). As shown in Fig. 4C, embryos injected with either zebrafish IGFBP-6a or -6b mRNA were morphologically normal, but they were smaller in size and developmentally delayed compared to GFP mRNA injected or wild type control embryos. The body lengths of IGFBP-6a and IGFBP-6b mRNA injected embryos were 2.163 ±0.015 mm (n = 48) and 2.121 ±0.009 mm (n = 47) (Fig. 4D), both were significantly smaller than the wild type control group (2.249±0.013 mm, n = 46) and the GFP mRNA injected group (2.277±0.012 mm, n = 52) (p<0.001). As shown in Fig. 4E, injection of IGFBP-6a or IGFBP-6b mRNA also significantly (p < 0.001) reduced the somite numbers from 22.67±0.36 somites of the wild type control group (n = 46) and 22.95±0.38 of the GFP mRNA injected group (n = 52) to 21.31±0.34 (n = 48) somites in the IGFBP-6a mRNA and 20.26 ±0.39 (n = 47) in the IGFBP-6b mRNA group.

We next performed in situ hybridization to determine any potential defects in patterning. Overexpression of either IGFBP-6a or IGFBP-6b did not alter the brain patterning, as indicated by mRNA expression patterns of egr2b (labeling third and fifth rhombomeres), emx1 (forebrain), rx1 (retina), and myoD (somite) (Fig. 5). These data suggest that the two IGFBP-6 isoforms have comparable biological activities. When expressed, they both inhibited embryonic growth and developmental rates without notable effects on cell fate determination or patterning.
Discussion

In this study, we have identified and characterized two structurally distinct IGFBP-6 genes from zebrafish. Structural, phylogenetic, and comparative genomic analyses indicate that they are co-orthologs of the human IGFBP-6 gene. Our molecular and functional analysis suggests that the duplicated zebrafish IGFBP-6 genes have undergone subfunctionalization partitioning by evolving distinct gene expression patterns.

The conclusion that zebrafish *igfbp-6a* and *-6b* are two co-orthologs of human IGFBP-6 is supported by several lines of evidence. Comparison of their primary sequence with those of other known fish IGFBPs revealed that they share the highest sequence identity to that of IGFBP-6. Phylogenetic analysis using two independent methods have firmly placed the zebrafish *igfbp-6a* and *-6b* gene clusters with other vertebrate IGFBP-6s. In humans, the *IGFBP-6* is located on Chromosome 12, next to the RARG gene. In its proximity are the HOXC gene cluster, PIP5K2C, CYP27B1, among others. In the zebrafish genome, zebrafish *igfbp-6a* and *-6b* are located on LG 11 and 23 respectively. In their neighborhoods, there are RARG, PIP5K2C, CYP27B1, and a cluster of HOX-C genes. In addition to these genes, we were able to identify other orthologous genes between zebrafish LG9 and human Chr. 2. This conserved synteny has provided very strong evidence that zebrafish *igfbp-6a* and *igfbp-6b* are co-orthologs of the human IGFBP-6 gene. The presence of two duplicated IGFBP-6 genes may not be unique to zebrafish. While zebrafish IGFBP-6b has 76% and 75% identity to that in trout and salmon IGFBP-6, zebrafish IGFBP-6a only shares 49% identity to that of rainbow trout and Atlantic salmon IGFBP-6. Judging from these divergences, the published rainbow trout IGFBP-6 and salmon IGFBP-6 are more
closely related to zebrafish IGFBP-6b. It is possible that there may be more IGFBP-6 in salmonid fish.

The finding of two functional zebrafish IGFBP-6 genes has provided us with the opportunity to gain insight into how the IGFBP-6 genes may have evolved through partitioning of ancestral functions. After duplication, each gene copy can follow a separate evolutionary fate. It has been postulated that non-functionalization is the fate for the majority of the duplicate genes with only 20% or so being retained (25, 29). Subfunction partitioning is common among the retained duplicated genes (25, 29).

Diversified expression profiles have been documented in duplicated zebrafish igfbp-1 genes (18). In those cases, it has been suggested that the changes in the regulatory modules in each of the two duplicates may have been sufficient to allow their distinct expression patterns. The duplicated zebrafish igfbp-6 genes are yet another example of temporal and spatial subfunctioning partitioning. In mammals and other teleost fish, IGFBP-6 mRNA is found in a wide variety of tissues. For instance, rainbow trout IGFBP-6 mRNA was expressed in almost all tissues analyzed with the strongest expression in the spleen, gills, brain, and stomach (17). IGFBP-6 mRNA also has a ubiquitous presence in rat and human (5, 28). In the adult zebrafish, IGFBP-6a mRNA is highly expressed in the muscle. Only very low levels of IGFBP-6a mRNA are found non-muscle tissues. In comparison, IGFBP-6b mRNA is abundantly expressed in brain, heart, and muscle. During embryogenesis, the two genes exhibit distinct temporal patterns. The IGFBP-6a mRNA levels are low during embryogenesis. The IGFBP-6b mRNA levels, on the other hand, are low during the first 48 h. They
become higher at 48 hpf and were maintained at high levels thereafter. These data suggest that the duplicated *igfbp-6* genes have evolved distinct spatial and temporal expression patterns. Compared with other IGFBPs (21, 33, 34), however, the levels of IGFBP-6a and -6b mRNA appeared to be considerably lower. In fact, they were under the detection limit of in situ hybridization and were detected by RT-PCR with very high cycles (data not shown). Maures et al. (23) reported both IGF-1 and IGF-2 mRNAs are expressed in most, if not all cells, throughout the embryogenesis. In adult zebrafish, IGF-2 mRNA is highly expressed in the skeletal muscle, brain, and heart. It is possible that the high levels of IGFBP-6a and -6b in these tissues may act as local regulators of IGF availability and biological activity in these adult zebrafish tissues.

Subfunction partitioning can involve protein structural changes (25). The duplicated zebrafish IGFBP-6s have clearly diverged as they are only 50% identical to each other in primary sequence. Despite divergences in their primary sequences, important motifs and domains are clearly conserved in the two zebrafish IGFBP-6s. Like all known IGFBPs, they possess a CWCV motif and 6 cysteine residues in their C domains. Mutational studies with human IGFBP-6 have indicated that Pro93/Leu94/Leu97/Leu98 are critical for IGF binding because changing them into Ala resulted in a more than 10,000-fold reduction in binding affinity for IGFs (10). These residues are perfectly conserved in the zebrafish IGFBP-6a (Pro80/Leu81/Leu84/Leu85) and IGFBP-6b (Pro88/Leu89/Leu92/Leu93). IGFBP-6 is unique among the IGFBP family in that it preferentially binds to IGF-2 (1). Headey et al. (14) have mapped the IGF-2 binding surface in human IGFBP-6. It includes
Leu174, Val178, Gly181, Val187, Asn189, Cys190, Gly194, Arg199, Ser203, Ser204, Gln205, and Gly206. In addition, Arg164, Leu167, Leu171, Ala182, Pro188, and Cys201 are also believed to be part of the binding site but this could not be verified experimentally. Most of the above residues are conserved in zebrafish IGFBP-6a and -6b. The perfectly conserved residues are: Arg121 (corresponding to Arg164 in human), Leu124, Leu128, Val135, Pro145, Asn146, Cys147, Gly151, Cys158, Ser160, Ser161, and Gly163 in zebrafish IGFBP-6a; Arg130, Leu133, Leu137, Leu140, Pro154, Asn155, Cys156, Gly160, Cys167, Ser169, Ser170, and Gly172 in zebrafish IGFBP-6b. There are also several conserved substitutions. Val187 in human IGFBP-6 is changed into Ile in IGFBP-6a (residue 144) and -6b (residue 153) and Arg199 is changed into Lys in zebrafish IGFBP-6a (Lys156) and IGFBP-6b (Lys165). Likewise, Val135 in human is changed into Ile144 in zebrafish IGFBP-6b and Leu174 in human is changed into Ile131 in zebrafish IGFBP-6a. These results suggest that the IGF-2 specific binding site is highly conserved in zebrafish IGFBP-6a and -6b. Human IGFBP-6 has been reported to possess a functional nuclear localization signal (HRGFYRKQQCRSSSQGQRRG-the bold letters indicate charged residues) in the N domain, which mediates its nuclear import (15). Similar sequences are found in zebrafish IGFBP-6a (148DKQGSRKQCRSSRGMQRGHC169) and IGFBP-6b (157DTRGFYRKQCRSSKGMQRGHC178). Whether the zebrafish IGFBP-6 genes can be found in the nucleus also awaits further functional tests.

*In vitro* studies using a variety of mammalian cell culture systems have indicated that mammalian IGFBP-6 has a predominantly inhibitory effect on IGF-2 actions (1),
although IGF-independent action in promoting cancer cell migration has also been reported (10). As mentioned in the Introduction, published literature suggests that mammalian IGFBP-6 is primarily an inhibitory IGFBP in vivo. In good agreement with these findings, injection of IGFBP-6a and -6b mRNA into zebrafish embryos caused significant decreases in body size and somite number. These overexpression experiments suggest that IGFBP-6a and -6b are inhibitory proteins and that they probably act by binding to and inhibiting IGF actions in vivo. Overexpression of either zebrafish IGFBP-6a or IGFBP-6b had no notable effects on cell fate determination or patterning, suggesting that they are primarily involved in growth and temporal development. Our functional analysis using transgenic zebrafish suggests that the inhibitory action of IGFBP-6 is conserved across species. More importantly, overexpression of zebrafish IGFBP-6a and -6b in zebrafish embryos had nearly identical phenotypes, i.e., they caused a similar decrease in the growth and developmental rates. Our results suggest that IGFBP-6a and 6b are not major players in the IGF signaling pathway during early embryogenesis because of their low expression levels during early development. In addition, overexpression of IGFBP-6a and -6b only caused modest decreases in growth and developmental rates compared to zebrafish IGFBP-1 and IGFBP-2 (18, 34).

Genomic studies have suggested that a genome duplication event occurred in the teleost lineage ~350 million years ago, prior to the beginning of the teleost radiation (25, 29). Although most of the duplicated genes have been lost, a substantial percentage of the duplicates have been retained. Recent studies have shown that most, if not all the major components in the fish IGF signaling pathway, are retained as duplicates in the
zebrafish genome. Zebrafish contain two \textit{igf-1r} genes and two \textit{insulin receptor} genes (27, 30). Recently we have found that there are two \textit{igfbp-1} genes (18), two \textit{igfbp-2} genes, and two \textit{igfbp-5} genes in zebrafish (34). In this study, we have demonstrated the presence of two functional IGFBP-6 genes in zebrafish. This high rate of gene retention in genes belonging to the IGF signaling pathway in the ray-fin fish is intriguing. Most teleost fishes do not reach a static adult size but continue to grow indeterminately well beyond puberty. Given the known importance of the IGF signaling pathway in growth regulation, the retention of the duplicated IGF ligands, receptors, and IGFBPs may play important roles in the continuous growth pattern in teleosts.

\textbf{Perspective and significance}

IGFBP-6 is unique among IGFBPs because it preferentially binds to IGF-2 over IGF-1. In this study, we identified and characterized two functional IGFBP-6 genes in zebrafish. Structural, phylogenetic, and comparative genomics analyses indicate that they are co-orthologs of the human IGFBP-6 gene, likely resulting from a genome wide duplication event before the beginning of the teleost radiation. Gene expression and functional analysis suggest that the two IGFBP-6 genes have undergone subfunctionalization partitioning by evolving distinct gene expression patterns, while their protein functionalities remain similar. Zebrafish IGFBP-6a and -6b are conserved regulatory protein and they inhibit embryo growth and development when overexpressed. This study has not only provided novel insights into the functional evolution of IGFBP-6, but also demonstrated the utility of the zebrafish model system
in studying IGFBP physiology. Further studies in zebrafish, focusing on determining whether the two zebrafish IGFBP-6 genes possess different ligand-binding, cellular localization, and IGF-independent properties, will be needed in order to elucidate the molecular basis of the duplicated IGFBP-6 genes.

Acknowledgements

We thank Mr. Justin Lenhart, University of Michigan, for proof reading this manuscript. This study was supported by NSF Research Grant IOB 0110864 and the Changjiang Scholars Program to C.D., the Natural Scientific Foundation of China (#30800838) and a fellowship from the Shandong Province Postdoctoral Fund (#200603073) to X.W.

Reference

1. **Bach LA.** IGFBP-6 five years on; not so ‘forgotten’? *Growth Horm IGF Res* 15: 185-192, 2005.


**Figure legends**
Fig. 1. A) Amino acid sequence alignment of zebrafish (zf) IGFBP-6a and -6b, rainbow trout (rt) IGFBP-6, Atlantic salmon (as) IGFBP-6, murine (mu) IGFBP-6, and human (hu) IGFBP-6. The residues conserved across species are shaded. B) A phylogenetic tree of the IGFBP gene family. Amino acid sequences of full-length IGFBPs were analyzed using the Neighbor-Joining method. Values on branches are percentages of times that the two clades branched as sisters (5000 runs). The results indicate that the two newly identified IGFBPs are most closely related to those in the IGFBP-6 subgroup.

Fig. 2. Genomic structure and chromosomal location of zebrafish igfbp-6a and igfbp-6b. A) A schematic diagram showing the structure of zebrafish igfbp-6a and igfbp-6b. Exons are shown as boxes and introns are shown as lines. B) Zebrafish igfbp-6a and igfbp-6b are syntenic to the human IGFBP-6 genes. Zebrafish igfbp-6a is located on linkage group (LG) 11; igfbp-6b is located on LG23. The human IGFBP-6 is located on chromosome 12. Note the highly conserved synteny in the surrounding genes.

Fig. 3. Distinct temporal and spatial expression patterns of zebrafish igfbp-6a and -6b. A) Temporal expression patterns of IGFBP-6a (left panel) and IGFBP-6b mRNA (right panel) during early development. Total RNA was isolated from 3 independent groups of embryos, each contain 20-30 embryos of the indicated stages. The developing stages are shown as hours post fertilization (hpf). Values are means ± S.E.M., n = 3, p < 0.05. B) Tissue distribution patterns of IGFBP-6a (left panel) and
IGFBP-6b mRNA (right panel). Total RNA was isolated from brain (Br), eye (Ey), gill (Gi), heart (He), intestine and stomach (I+S), kidney (Ki), liver (Li), muscle (Mu), ovary (Ov), spleen (Sp), and testis (Te) from adult zebrafish. The IGFBP-6a and -6b mRNA levels measured by qRT-PCR. The values are expressed as relative value to the β-actin mRNA levels. Values shown are means ± S.E.M., n = 4, * p < 0.05.

Fig. 4. IGFBP-6a and IGFBP-6b both inhibit embryo growth and development. A) Representative images of wild-type (WT), GFP mRNA (600 pg/embryo), IGFBP-6a mRNA (600 pg/embryo), or IGFBP-6b mRNA (600 pg/embryo) injected embryos. Upper panels are bright field views and lower panels show fluorescence views. Scale bar = 500 m. B) Western immunoblot analysis result. Lysates prepared from zebrafish embryos injected with IGFBP-6a-GFP (BP6a), IGFBP-6b-GFP (BP6b), GFP mRNA (GFP) or wild type (WT) embryos were subjected to Immunoblot analysis using a polyclonal anti-GFP antibody. C) Representative images of wild-type, GFP mRNA, IGFBP-6a mRNA, or IGFBP-6b mRNA injected embryos at 24 hpf. D) Body length of the groups indicated. Numbers of embryos are shown in the column. Values are represented as means ± S.E.M. *, p < 0.001 compared to the control GFP mRNA group. Similar results were obtained from three independent microinjection experiments. E) Somite numbers of the groups indicated. Numbers of embryos are shown in the column. Values are represented as means ± S.E.M. *, p < 0.001 compared to the control GFP mRNA group. Similar results were obtained from three independent microinjection experiments.
Fig. 5. Ectopic expression of IGFBP-6a and IGFBP-6b has no effect on cell fate and patterning. Whole mount *in situ* hybridization images of wide type, GFP mRNA, IGFBP-6a, and IGFBP-6b mRNA-injected embryos at 24 hpf; *egr2b* expression in the third and fifth rhombomeres of the hindbrain; *emx1* expression in the forebrain; *rx1* expression in the retina; *myoD* expression in the somatic myotome. Scale bar = 200 µm. Similar patterns were observed in all embryos examined in each group (n = 20-25).

Supplemental Fig. 1. Phylogenetic tree of the IGFBP family constructed by the Maximum Likelihood (ML) method with the PhyML V2.4.4 program. Values on branches are percentages of times that the two clades branched as sisters. The results indicate that the two newly identified IGFBPs are most closely related to those in the IGFBP-6 subgroup.
Table 1. Amino acid sequence identities between zebrafish IGFBP-6a/IGFBP-6b and other IGFBPs.

<table>
<thead>
<tr>
<th>Protein Species#</th>
<th>IGFBP-1</th>
<th>IGFBP-2</th>
<th>IGFBP-3</th>
<th>IGFBP-4</th>
<th>IGFBP-5</th>
<th>IGFBP-6</th>
</tr>
</thead>
<tbody>
<tr>
<td>human</td>
<td>26/25</td>
<td>26/29</td>
<td>33/33</td>
<td>34/28</td>
<td>37/37</td>
<td>32/34</td>
</tr>
<tr>
<td>mouse</td>
<td>28/27</td>
<td>26/30</td>
<td>28/35</td>
<td>29/29</td>
<td>37/33</td>
<td>30/36</td>
</tr>
<tr>
<td>rat</td>
<td>28/27</td>
<td>26/30</td>
<td>28/35</td>
<td>29/27</td>
<td>36/35</td>
<td>29/35</td>
</tr>
<tr>
<td>bovine</td>
<td>25/24</td>
<td>26/30</td>
<td>29/34</td>
<td>34/27</td>
<td>37/37</td>
<td>31/33</td>
</tr>
<tr>
<td>chicken</td>
<td>31/31</td>
<td>28/29</td>
<td>-</td>
<td>28/26</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>xenopus</td>
<td>29/27</td>
<td>24/31</td>
<td>-</td>
<td>-</td>
<td>35/34</td>
<td>-</td>
</tr>
<tr>
<td>trout</td>
<td>28/22</td>
<td>27/32</td>
<td>25/24</td>
<td>26/25*</td>
<td>38/33</td>
<td>49/76</td>
</tr>
<tr>
<td>salmon</td>
<td>29/22</td>
<td>27/29</td>
<td>25/24</td>
<td>-</td>
<td>35/33</td>
<td>49/75</td>
</tr>
<tr>
<td>zebrafish</td>
<td>26/24**</td>
<td>26/28***</td>
<td>27/25</td>
<td>34/35</td>
<td>-</td>
<td>34/32</td>
</tr>
</tbody>
</table>

The left numbers show identities between zebrafish IGFBP-6a and other IGFBPs, and the right number show identities between zebrafish IGFBP-6b and other IGFBPs;
* Compared with corresponding partial sequences;
** Identities between zebrafish IGFBP-1a and zebrafish IGFBP-6s;
*** Identities between zebrafish IGFBP-1b and zebrafish IGFBP-6s;
**** Identity between zebrafish IGFBP-6a and IGFBP-6b.
A

IGFBP-6a

5' 108 316 2K 15 128 2K TAA 3' 556

IGFBP-6b

5' 154 143 15 981 128 114 TGA 3' 334 161

400bp

200bp

B

Zebrasfish LG 11

200K

Human Chr. 12

500K

Zebrasfish LG 23

200K