



Article

A Comparative Genomic and Transcriptional Survey Providing Novel Insights into Bone Morphogenetic Protein 2 (*bmp2*) in Fishes

Guang Yang^{1,2}, Zhendong Qin¹, Hongyan Kou¹, Rishen Liang¹, Lijuan Zhao¹, Shoujia Jiang³, Li Lin^{1,*} and Kai Zhang^{1,4,*}

¹ Guangdong Provincial Water Environment and Aquatic Products Security Engineering Technology, Research Center, Guangzhou Key Laboratory of Aquatic Animal Diseases and Waterfowl Breeding, Zhongkai University of Agriculture and Engineering, Guangzhou 510225, China; yangguang@m.scnu.edu.cn (G.Y.); qinzhendongsc@163.com (Z.Q.); langfanghongyan@163.com (H.K.); cheetahliang@126.com (R.L.); zhaolijuan4234@163.com (L.Z.)

² Guangdong Provincial Key Laboratory for Healthy and Safe Aquaculture, College of Life Science, South China Normal University, Guangzhou 510631, China

³ Shenzhen Key Lab of Marine Genomics, Guangdong Provincial Key Lab of Molecular Breeding in Marine Economic Animals, BGI Academy of Marine Sciences, BGI Marine, BGI, Shenzhen 518083, China; jiangshoujia@genomics.cn

⁴ Division of Life Science, Hong Kong University of Science and Technology, Hong Kong 93117, China

* Correspondence: linli@zhku.edu.cn (L.L.); zhangkai@zhku.edu.cn (K.Z.);
Tel.: +86-133-4280-5517 (L.L.); +86-185-9810-9029 (K.Z.)

Received: 12 November 2019; Accepted: 3 December 2019; Published: 5 December 2019



Abstract: Intermuscular bones (IBs) are only found in the muscles of fish. Bone morphogenetic protein 2 (*bmp2*) is considered to be the most active single osteogenesis factor. It promotes cell proliferation and differentiation during bone repair, as well as inducing the formation of bones and cartilages in vivo. However, detailed investigations of this family in fish are incredibly limited. Here, we have used a variety of published and unpublished *bmp2* sequences for teleosts and cartilage fish in order to explore and expand our understanding of *bmp2* genes in fish. Our results confirmed that teleost genomes contain two or more *bmp2* genes, and the diversity of *bmp2* genes in vertebrates appears to be as a result of a combination of whole genome duplication (WGD) and gene loss. Differences were also observed in tissue distribution and relative transcription abundance of the *bmp2s* through a transcriptomic analysis. Our data also indicated that *bmp2b* may play an important role in the formation of IBs in teleosts. In addition, protein sequence alignments and 3D structural predictions of *bmp2a* and *bmp2b* supported their similar roles in fishes. To summarize, our existing work provided novel insights into the *bmp2* family genes in fishes through a mixture of comparative genomic and transcriptomic analysis.

Keywords: IBs; *bmp2*; molecular evolution; comparative genomic; transcriptional level

1. Introduction

Most aquaculture freshwater fish, in particular the cyprinidae species, contain a specific amount of IBs which are hard-boned spicules that can be found in the muscle tissue on both sides of the vertebrae [1]. IBs are only present in the myosepta of lower teleosts, and these are viewed as ossified myoseptal tendons which develop directly from mesenchymal condensation [2]. Previous studies have discovered that the appearance of intermuscular bones (IBs) can enhance the power of herbivorous fish to a certain extent, which means that these fish can better adapt to the environment and evade

natural predators [3–5]. The presence of IBs influences the farming and processing of teleosts, to some degree, as well as decreasing the economic value and edibility of the species, due to the fact that they are difficult to remove [6]. While observing the potential value of the eradication of IBs, researchers initiated the research of IBs in fish in the early 1960s [7]. Until now, almost all current research has only focused on the IBs' morphology in species of fish [8–10]. A very small amount of detailed research has been conducted and little information is known in relation to the molecular mechanism for the development of IBs.

Bone morphogenetic proteins (bmps) are active proteins which were extracted from adult bone tissue by Urist in 1965. Apart from bmp1, other bmps are members of the transforming growth factor-beta (TGF- β) superfamily, which promote the formation of cartilage and bone tissue [11,12]. It was originally believed that bmps only play a role in the formation of bone and cartilage in vertebrates, but now it is believed that bmps play biological roles in various cell types [13]. In addition to inducing the formation of bone and cartilage, bmps are also involved in regulating the development of teeth, kidneys, skin, hair, muscles, hematopoiesis, and neurons, as well as maintaining iron metabolism and vascular homeostasis in the body [14,15]. Bmp2 is considered to be the most active single factor of osteogenesis [16]. Bmp2 has been widely studied in the bmp family since 1988, and it has played an important role in bone induction, and become a therapeutic agent for the restoration of bones and teeth [17,18]. Most data indicate that bmp2 is associated with the initial growth of various organs and tissues of vertebrates, and that it is also involved in the early development of the skeletal system and the formation of various organs [19,20]. Most current research into the *bmp2* gene focuses on its role in bone fragmentation and repair. The role of the *bmp2* gene is still unclear in relation to the formation of IBs in teleosts.

In recent years, the study of bmp2 has been extended to fish models in bone biology, in particular the zebrafish (*Danio rerio*). Mowbray defined the mRNA expression patterns in the ears of developing zebrafish as bmp2b, and provided supporting evidence that bmps play a vital role in a vertebrate's ear development [21]. Spatiotemporal expression patterns of bmp2 in zebrafish and Senegalese sole (*Solea senegalensis*) revealed that all isoforms are stated in calcified tissues, but at a variety of levels [22,23]. Also, bmp2a and bmp2b have been recognized as being part of zebrafish fin regeneration or development [24,25]. Zebrafish swirl mutant which is triggered by alterations in the *bmp2b* gene [26] indicates that bmp2b is vital during the early dorsoventral patterning. The studies of components of *bmp2* genes in other species of fish relating to bone biology are inadequate, and most studies relate to the copy number of *bmp2* genes. The *bmp2* family is present in different members of different species. Most tetrapods harbor a single *bmp2* gene, while two *bmp2* genes are reportedly presented in teleosts [27]. Previous studies have shown that the *bmp2b* gene has the highest expression level in the muscles of common carp (*Cyprinus carpio*) [28]. Furthermore, the *bmp2a* gene is highly expressed in calcified bone tissue of the gilthead seabream (*Sparus aurata*) [29] and in the ovaries of African clawed frogs (*Xenopus laevis*) [30].

Although expression patterns of *bmp2* genes have been reported in a few boned fishes, a comprehensive and comparative genomic and transcriptomic survey of these family genes in various species of fish has not yet been undertaken. In this study, we extracted fish *bmp2* genes (and their encoding sequences) from several published genome sequences, in order to allow us to investigate the presence or absence of species-specific *bmp2* isoforms and their sequence differences throughout various species. Comprehensive analysis was conducted for *bmp2* genes in fish, including the analysis of molecular structures, gene copy numbers, and phylogenetic relationships. Furthermore, the tissue distribution of these genes was determined. Our results lay a solid foundation for further functional studies of *bmp2* genes in fish.

2. Results

2.1. Variation of *bmp2* Copy Number in Different Fish Species for a Phylogenetic Analysis

This study analyzed the *bmp2a* and *bmp2b* sequences (including gene sequence information and protein sequence information) of 37 different vertebrates, including 28 species of fish. Table 1 displays information of the *bmp2* gene copy number which was obtained from the genome of 28 species of fish and one species of amphibian using bioinformatics. Table S1 shows the publicly-reported sequence information of *bmp2a* and *bmp2b* from 16 species of vertebrate (including eight species of fish and eight species of mammal).

Table 1. The copy number of *bmp2* genes from the fishes sequenced by our lab or collaborators.

Class	Common Name	Species Name	<i>bmp2a</i>	<i>bmp2b</i>
Actinopterygii	Snakehead fish	<i>Channa argus</i>	0	1
	Large yellow croaker	<i>Larimichthys crocea</i>	0	1
	Giant-fin mudskipper	<i>Periophthalmus magnuspinnatus</i>	0	1
	Black porgy	<i>Acanthopagrus schlegelii</i>	0	1
	Cave-restricted golden-line fish (Sa)	<i>Sinocyclocheilus anshuiensis</i>	2	2
	Semi-cave-dwelling golden-line fish (Sr)	<i>Sinocyclocheilus rhinocerosus</i>	2	2
	Surface-dwelling golden-line fish (Sg)	<i>Sinocyclocheilus grahami</i>	2	1
	Kanglang fish	<i>Anabarilius sgrahami</i>	1	1
	Catfish	<i>Silurus asotus</i>	1	1
	Channel catfish	<i>ictalurus punctatus</i>	1	1
	Lined_seahorse	<i>Hippocampus erectus</i>	0	1
	Longjaw grenadier anchovy	<i>Coilia macrognathos</i>	1	0
	Spotted gar	<i>Lepisosteus oculatus</i>	0	1
	Golden arowana	<i>Scleropages formosus</i>	0	1
	Bluntnose black bream	<i>Megalobrama amblycephala</i>	1	1
	Atlantic herring	<i>Clupea harengus</i>	1	1
	Hilsa shad	<i>Tenuulosa ilisha</i>	2	1
	Mexican tetra	<i>mexicanus</i>	1	1
	Zebrafish	<i>Danio rerio</i>	1	1
	Japanese medaka	<i>Oryzias latipes</i>	1	1
	Nile tilapia	<i>Oreochromis niloticus</i>	1	1
	Gilthead seabream	<i>Sparus aurata</i>	1	1
	Tongue sole	<i>Cynoglossus semilaevis</i>	1	1
	Japanese flounder	<i>Paralichthys olivaceus</i>	1	1
	Atlantic salmo	<i>Salmo salar</i>	2	2
	Fugu rubripes	<i>Takifugu rubripes</i>	1	1
	Sarcopterygii	Coelacanth	<i>Latimeria chalumnae</i>	0
Holocephali	Australian ghostshark	<i>Callorhynchus milii</i>	1	1
Mammalis	Platypus	<i>Ornithorhynchus anatinus</i>	1	1

We performed gene copy number analysis of the *bmp2* gene in fish genomes. Based on the results of our analysis, the copy number of the *bmp2* genes from different species has a certain variation. There is only one *bmp2* gene in the tetrapod genome (*Homo sapiens*, *Mus musculus*, *Ornithorhynchus anatinus*, and *Pan troglodytes*), whereas there are two or more in the teleost genome. There are two *bmp2* genes (*bmp2a* and *bmp2b*) in the genome of diploid teleosts, such as Nile tilapia (*Oreochromis niloticus*) and bluntnose black bream (*Megalobrama amblycephala*), yet there are exceptions as there are four *bmp2* genes in the Atlantic salmon (*Salmo salar*) genome (two *bmp2a* and two *bmp2b*). We discovered multiple copies in tetraploid teleosts, including surface-dwelling golden-line fish (*Sinocyclocheilus grahami*, Sg), semi-cave-dwelling golden-line fish (*Sinocyclocheilus rhinocerosus*, Sr), and cave-restricted golden-line fish (*Sinocyclocheilus anshuiensis*, Sa). Golden-line fish have three or four *bmp2* copies. In addition, we analyzed the copy number of *bmp2* genes in some ancient fish and discovered that only one copy of the *bmp2b* gene was present in the spotted gar (*Lepisosteus oculatus*) and golden arowana (*Scleropages formosus*).

In addition, we found that both diploid and polyploid fishes contain at least one copy of the *bmp2b* gene. Interestingly, we discovered that some marine fish species lost their *bmp2a* genes, including the

large yellow croaker (*Larimichthys crocea*) and black porgy (*Acanthopagrus schlegelii*). This could have been due to gene loss during evolution or the inability to excavate the *bmp2b* gene from these species due to genome assembly problems.

We constructed a phylogenetic tree, using all the acquired protein sequences of the *bmp2* genes, as displayed in Figure 1. Our results showed that all genes were clustered into three main branches. The *bmp2a* genes of all fish and the *bmp2* genes of tetrapods clustered into a large branch. However, this large branch formed two small branches, namely the *bmp2a* gene branch of the teleosts and the *bmp2* genes branch of the tetrapods. All *bmp2b* genes of the teleosts were clustered into a branch. The phylogenetic tree of the *bmp2* gene family was constructed using phylogenetic analysis in order to determine the specific subtype and evolutionary status of the *bmp2* genes of certain species.

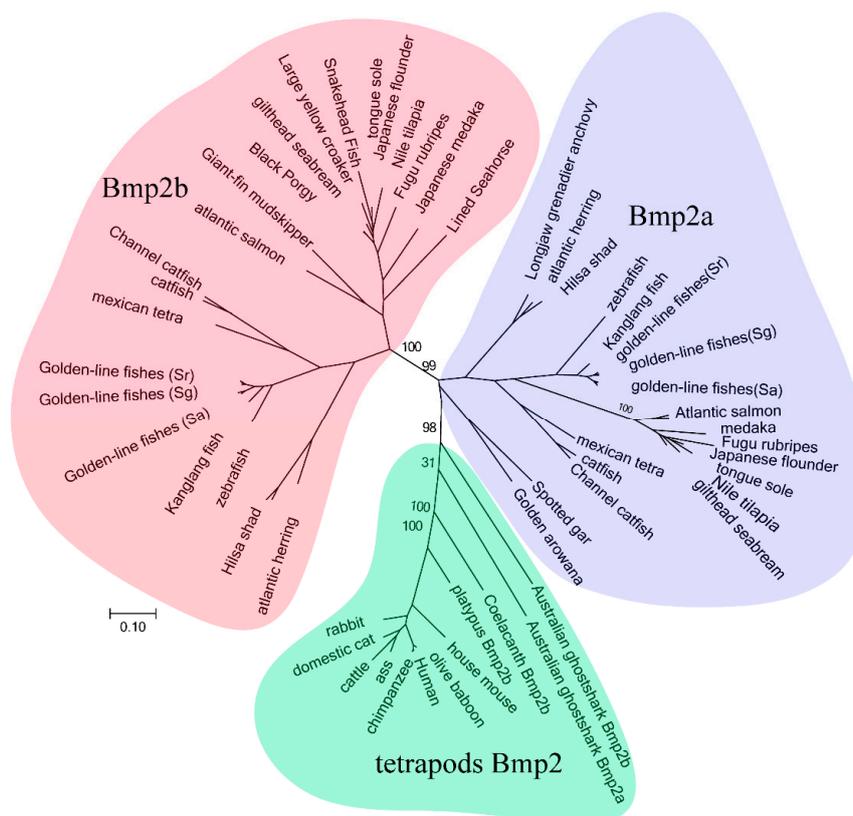


Figure 1. Clustering of *bmp2* genes. The phylogenetic analysis was performed by RAxML8.0.17. Numbers on branches are bootstrap values. Scale bar indicates the rate of amino acid substitution per residue. Although the MEGA tree is not shown here, it possesses the same tree topology. Different *bmp2* subfamilies are displayed with various colors: blue, teleost *bmp2a*; green, tetrapod *bmp2*; and pink, teleost *bmp2b*.

2.2. Synteny Data

We used the zebrafish *bmp2a* and *bmp2b* gene sequences as reference points, and then selected multiple upstream and downstream genes. The protein sequences of these genes were downloaded and compared to the genomes of other species in order to find the two genes on the genomes and the location of the upstream and downstream genes. The specific results are shown in the right-hand side of Figure 2. From the results of the collinear analysis, it can be noted that even though certain species have some gene loss, the distribution of *bmp2a* and *bmp2b* genes in the teleost genome is generally conserved. For the most part, the *hao1* gene and the *tmx4* gene are present upstream of the *bmp2a* gene in the genome of teleosts, and the *drv1* gene, the *napba* gene, the *mcm8* gene, and the *crls1* gene are present downstream of the *bmp2a* gene in the teleost genome. The upstream and downstream

genes of the *bmp2b* gene are highly conserved in the teleost genome. The *Gpcpd1* gene, *trmt6* gene, and *fermt1* gene exist upstream of the *bmp2b* gene in the teleost genome. *TAAR12j* genes, *TAAR12e* genes, *TAAR12g* genes, *TAAR12f* genes, *TAAR12a* genes, *TAAR12b* genes, *TAAR12c* genes, and *TAAR10* gene exist in the downstream of *bmp2b* genes in the teleost genome. We found that the distribution of the *bmp2* genes on the mammalian genome is different to the distribution of the *bmp2a* genes in teleosts. Significant gene loss was discovered in the upstream and downstream of the mammalian *bmp2* gene. A significant amount of gene loss was discovered upstream and downstream of the *bmp2a* and *bmp2b* genes of the Australian ghostshark (*Callorhinchus milii*). The collinear analysis also confirmed that the *bmp2a* and *bmp2b* genes we extracted were reliable.

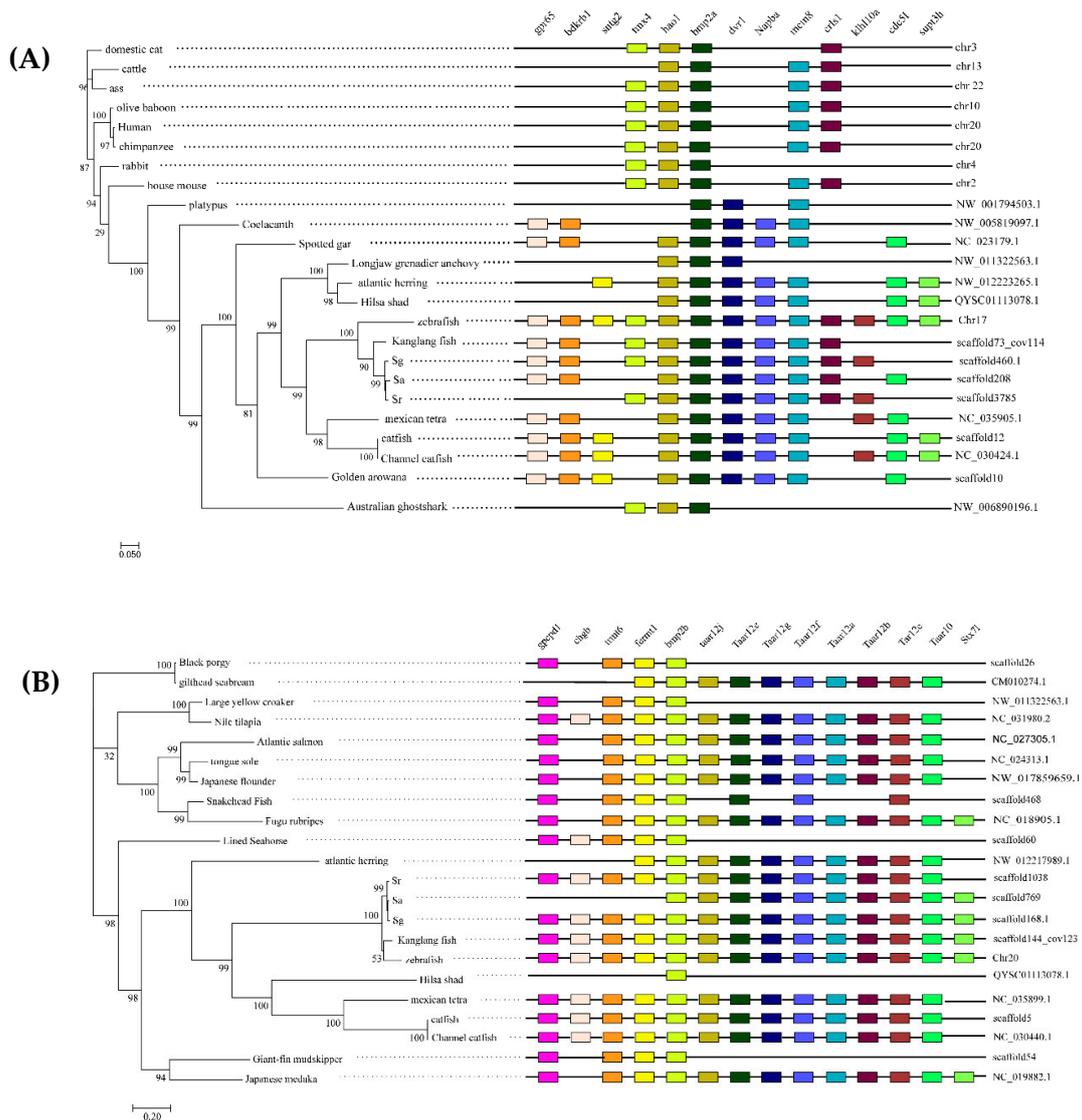


Figure 2. Phylogenetic trees and synteny of *bmp2a* (A) and *bmp2b* (B). These maximum likelihood (ML) trees (on the left-hand side of each figure) were constructed by FastTree v2.1.7. Bootstrap values are shown on branches. The right-hand figures are synteny of the two *bmp2* genes.

2.3. Sequence Alignment and Three-Dimensional (3D) Structure Prediction of Zebrafish *bmp2a* Gene and *bmp2b* Gene

We predicted that the domains of the zebrafish *bmp2a* gene and *bmp2b* gene consisted of three parts, both of which contained the TGFb_propeptide domain and the TGFb domain. Bmp2a also contains a low complexity region, and *bmp2b* contains a signal peptide. TGFb_propeptide represents

the propeptide region of TGF- β forming Latency associated peptide (LAP). TGF- β is concealed as a potential complex, and is comprised of two parts: TGF- β dimer and TGF- β binding protein (LTBP). Following translational processing, TGF- β is concealed as a complex consisting of TGF- β dimer and TGF- β binding protein (LTBP) which are non-covalently bound to LAP. The potential TGF- β can be targeted to the extracellular matrix by LTBP. LAP dimers typically bind to LTBP with disulfide bonds. TGFB belongs to the transmuting growth factor-beta (TGF- β) family. TGFB is a multi-functional peptide which controls the proliferation, differentiation, and other functions of several cell types.

Both the *bmp2a* gene and *bmp2b* gene of the zebrafish contain a common receptor protein with an aserine/threonine kinase activation site in the spatial structure (see Figure 3). The serine/threonine kinase activation site residues on *bmp2a* are located at amino acid positions (286, 288, 291, 293, 310, 313, 315, 317, 351, 353, 356, 378, and 383). The serine/threonine kinase activation site residues on *bmp2b* are located at amino acid positions (311, 313, 316, 318, 335, 338, 340, 342, 376, 378, 381, 384, 403, and 408). Both *bmp2a* and *bmp2b* in zebrafish have a receptor protein with the serine/threonine kinase activation site, which suggests that this site may be significant for maintaining the function of the zebrafish *bmp2* gene family.

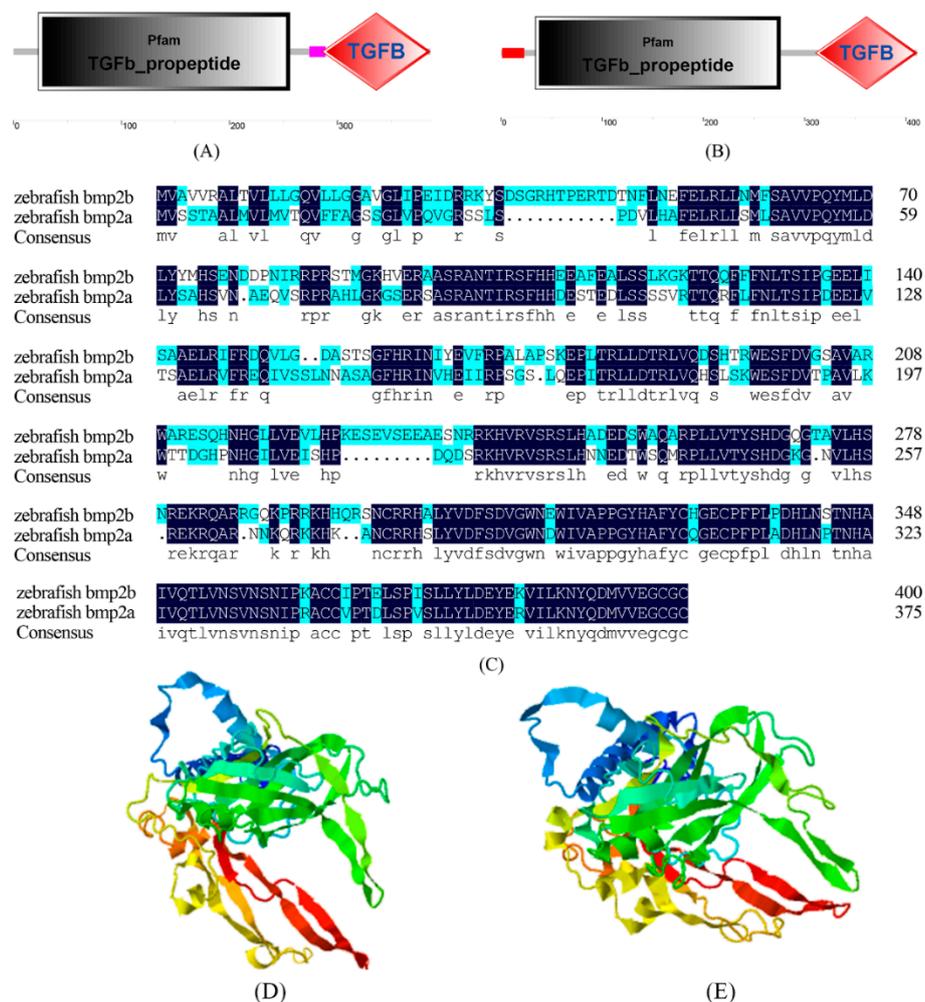


Figure 3. Alignment of *bmp2* protein sequences of zebrafish. (A) The domain of zebrafish *bmp2a*, (B) the domain of zebrafish *bmp2b*, (C) alignment of *bmp2a* and *bmp2b* protein sequences of zebrafish, (D) predicted 3D models of *bmp2a* and *bmp2b* of zebrafish, and (E) predicted 3D models of *bmp2b* of zebrafish.

2.4. Structural Analysis of *bmp2a* and *bmp2b*

Through the structural analysis of the *bmp2a* gene and *bmp2b* gene (see Figure 4), it was discovered that both the *bmp2a* gene and *bmp2b* gene of teleosts contain two exons, and the exons and introns are arranged in a very conserved gene. The model is consistent, and this suggests that *bmp2a* and *bmp2b* genes could have acquainted functions.

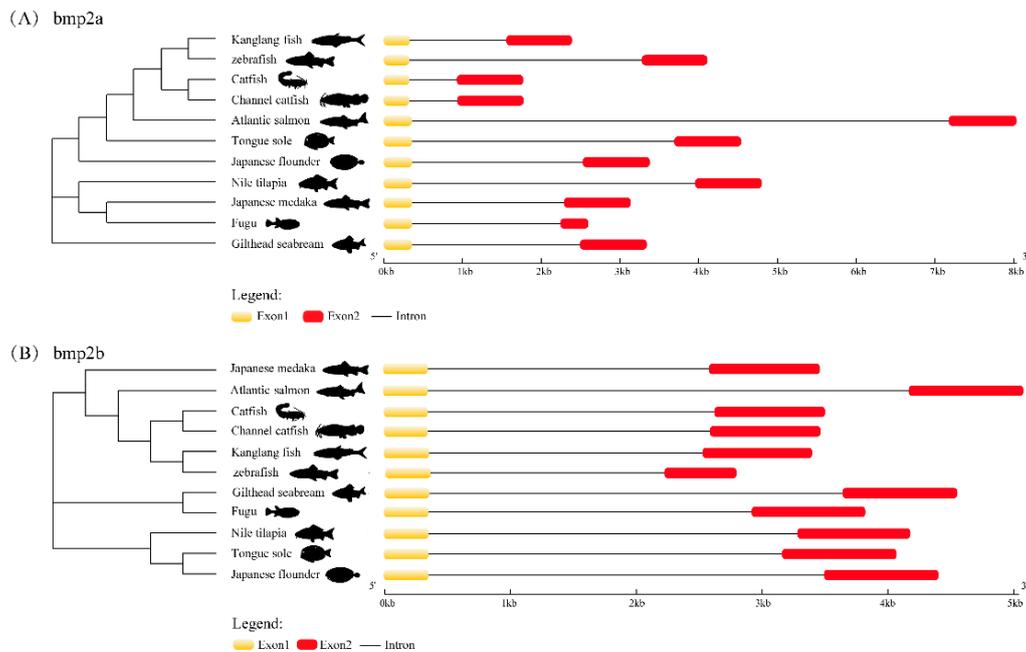


Figure 4. Exon–intron structures of *bmp2a* (A) and *bmp2b* (B). Boxes in red and yellow represent the exons, while gray lines represent the introns.

We downloaded the three-dimensional structure information of human *bmp2* protein from the Protein Data Bank (PDB) database as a template for our analysis. The sequence alignment results and three-dimensional structure analysis results are displayed in Figure 5. We focused on amino acid sites which were thoroughly linked to the function and structure of *bmp2* proteins. It was discovered that certain amino acids around the 2HP ligand binding site were conserved in *bmp2a* and *bmp2b* genes. For example, the two proteins of all species were isoleucine (Ile) residues, proline (Val) residues, alanine (Ala) residues, and proline (Pro) residues at positions 304, 305, 306, and 307 (relative to the *bmp2a* protein of zebrafish, see Figure 5). These amino acid residues can maintain the stability of the 2HP ligand. *Bmp2a* protein and *bmp2b* protein were also conserved at the DIO ligand binding site. For example, two proteins residues are tryptophan (Trp) residues, cysteine (Cys) residues, and tyrosine (Tyr) residues at positions 301, 303, and 363. However, it was very different at positions 37, 60, and 124. The sites are proline (Pro) at position 37 in most mammals, while the *bmp2a* is leucine (Leu) in other teleosts, and the *bmp2b* is phenylalanine (Phe) in most teleosts. The sites are aspartic acid (Asp) at position 60 in most mammals, while the *bmp2a* is serine (Ser) in other teleosts, and the *bmp2b* is glutamine (Glu) in most teleosts. At position 124, mammalian *bmp2* and teleost *bmp2a* are arginine (Arg), while teleost *bmp2b* is glutamine (Glu). These key sites of *bmp2* may be involved in the evolution of *bmp2* in mammals and teleosts.

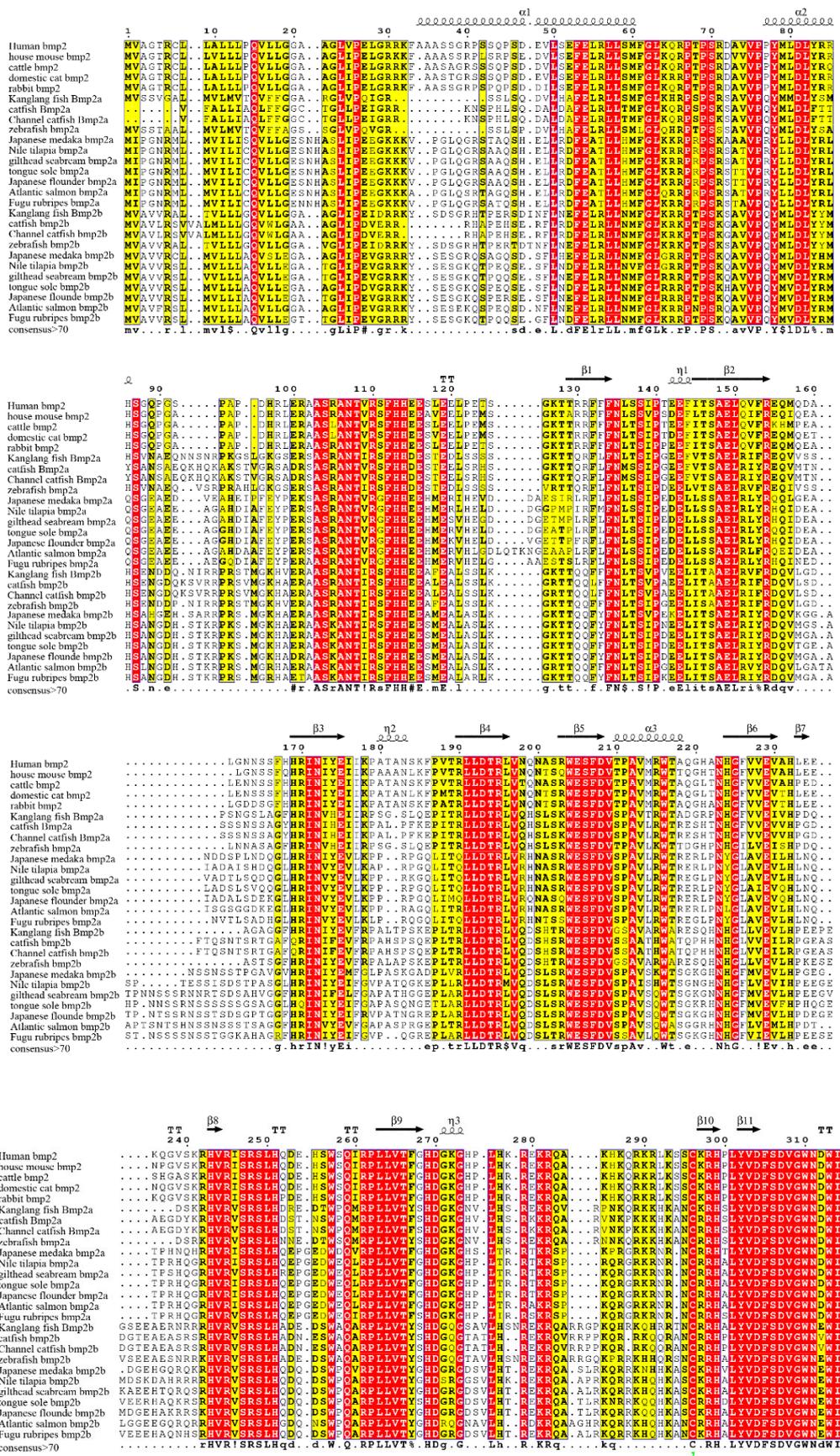


Figure 5. Cont.

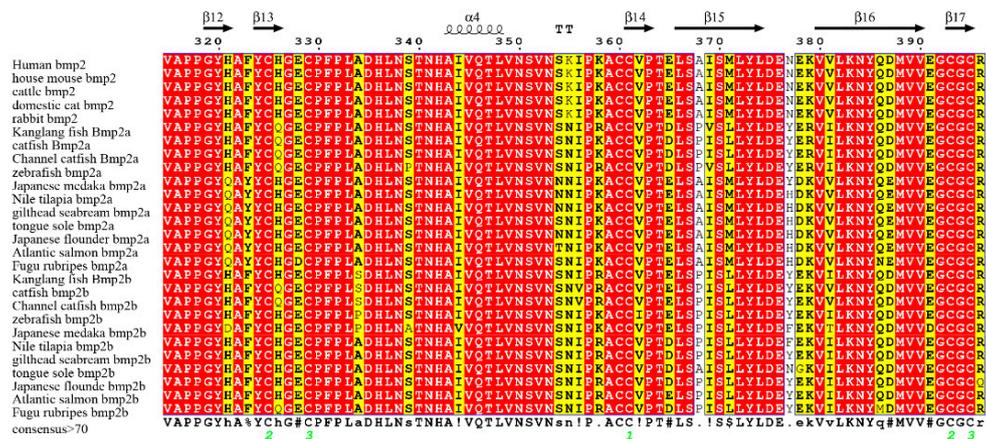


Figure 5. Alignment of bmp2 protein sequences. These sequences were aligned with human bmp2 by multiple sequence alignment based on fast Fourier transform (MAFFT) and colored using TEXshade. The secondary structural elements, alpha helix (α) and beta strand (β), are marked. The color code for the conservation track ranges from red (the most conserved) to yellow (the least conserved) as per TEXshade.

2.5. Transcriptional Level Studies of Various Tissues of Nile Tilapia and Bluntnose Black Bream

We chose two representative teleosts: the Nile tilapia and the bluntnose black bream. Eight tissues, including IBs, spleen, heart, muscle, liver, brain, and fin were selected as research objects for the Nile tilapia and bluntnose black bream. Using β -actin as an internal reference gene, the expression of *bmp2* genes of Nile tilapia and bluntnose black bream in eight tissues which were ribs, intermuscular spines, fins, brain, liver, muscle, heart and spleen was detected by quantitative real-time PCR (qRT-PCR). The results indicated that the *bmp2* genes were expressed in eight tissues of Nile tilapia and bluntnose black bream, but the expression levels were all different (see Figure 6). *Bmp2a* had the highest expression in the spleen, followed by IBs, ribs and fins, and muscle, brain, and liver, with the lowest expression being in the heart of the bluntnose black bream. *Bmp2b* was highly expressed in the IBs, ribs, and fins, and the expression in the liver, brain, and muscle was second. The expression in the heart and spleen was the lowest in the bluntnose black bream. The expression of the *bmp2* genes in various tissues of the Nile tilapia was similar to that in the bluntnose black bream. *Bmp2a* had the highest expression in muscle, IBs, rib, and fin, followed by spleen and brain, with the lowest expression in the liver and heart of the Nile tilapia. *Bmp2b* was highly expressed in the muscle, liver, and the spleen, and second in ribs, IBs, and fins. The expression was lowest in the brain and heart of the bluntnose black bream.

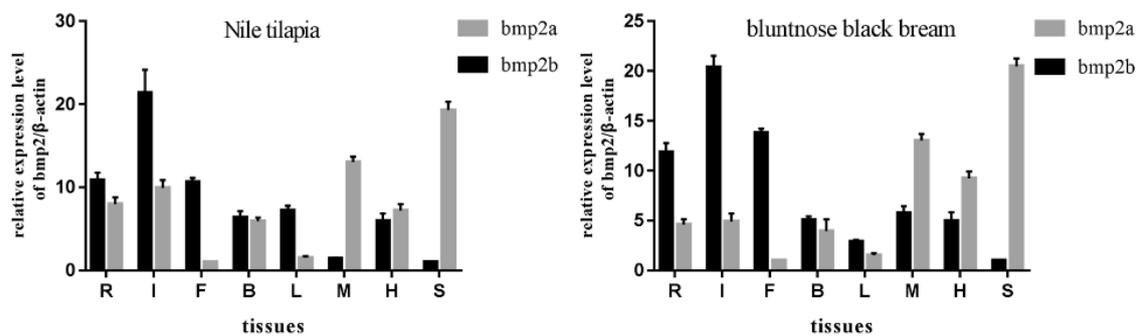


Figure 6. Tissue distribution analysis of *bmp2* of Nile tilapia and Bluntnose black bream were analyzed by quantitative real-time PCR (qRT-PCR). Relative expression levels of *bmp2a* and *bmp2b* in the gill, muscle, stomach, heart, brain, and hemolymph, with β -actin as an internal control. Letters R, I, F, B, L, M, H, S stand for rib, intermuscular bones (IB), fin, brain, liver, muscle, heart, and spleen, respectively.

3. Discussion

In this study, we researched several surfaces of *bmp2* genes (with both *bmp2a* and *bmp2b*), and gained novel insights into the structural variations and diversity of *bmp2* in vertebrates from a genomics point of view.

3.1. Possible Reasons for Copy Number Variations among Vertebrates

We analyzed variations of copy numbers of the *bmp2a* and *bmp2b* genes in different vertebrates and concluded that the main causes of variations were genome-wide replication and gene deletion. As we are aware from the relevant literature and studies, the general forefather of early vertebrates experienced two rounds of whole genome duplication (WGD) [31,32]. All teleosts are generally considered to have experienced at least three whole genome duplication events [33,34], and some fish even experience a fourth genome duplication [35]. Our results suggested that when tetrapods and teleosts are isolated, tetrapods lose the *bmp2b* gene, and there is therefore a *bmp2a* gene in tetrapods. Two *bmp2* genes (*bmp2a* and *bmp2b*) are generally present in the genome of diploid teleosts, such as the bluntnose black bream. Some fish which have undergone a fourth round of genome duplication, such as the tetraploid golden-line fish and Atlantic salmon, contain two *bmp2a* genes and two *bmp2b* genes in their genomes. The copy number of the *bmp2a* gene in tetraploid fish genomes is not 2:1, compared to the copy number of the *bmp2a* genes contained in the diploid fish genome, which indicates that the *bmp2a* genes have been lost in the genome of tetraploid fish. Therefore, we can conclude that the important factors for copy number variation in the *bmp2a* and *bmp2b* genes are genome-wide duplication and gene deletion. Based on our findings, we projected that the important factors for copy number variations of *bmp2* in vertebrates were caused by a combination of gene loss and WGD. The number of copies between diploid and tetraploid does not always correspond to one-to-two, due to the selective loss of genes.

In addition, our synteny analysis results showed that the *bmp2a* and *bmp2b* genes are restricted in the same species with different chromosomes, whereas all *bmp2a* and *bmp2b* genes, across species share a preserved suite of genes binding them on both sides, although some species may display gene loss. Interestingly, we determined that the synteny genes were not conserved between teleosts and tetrapods, indicating that the *bmp2* family genes experienced rearrangement from teleosts to tetrapods.

3.2. Adaptive Evolution of *bmp2s* in Vertebrates

IBs are common and play a role in muscle support and power transmission in the muscles of lower teleosts [36,37]. The reason for the formation of IBs in lower teleosts could be that they are induced by differentiation from muscle fibroblasts, and muscle mesenchymal cells [38,39]. Additionally, the preference for food among lower teleosts is an important biological phenomenon [40–42]. According to our findings, we discovered that herbivorous fish (such as the bluntnose black bream which has more IBs) have more *bmp2* genes than carnivorous fish (such as the snakehead fish which has fewer IBs). Previous studies have reported that the existence of IBs enhances the power of herbivorous fish to a certain extent, which allows these fish to better adapt to their environment and to evade natural predators [3–5]. According to our phylogenetic analysis and protein structure comparisons, we observed that some marine fishes, including the large yellow croaker (*Larimichthys crocea*) and black porgy (*Acanthopagrus schlegelii*), only have one *bmp2* gene (*bmp2a*), and this could result in fewer IBs among marine fish.

Through the analysis of the expression levels of *bmp2a* and *bmp2b* in various teleost tissues, we discovered that *bmp2b* is expressed mostly in the IB, ribs, and fins of the Nile tilapia and *bmp2a* is expressed mostly in the spleen, muscle, and heart of the Nile tilapia, which is consistent with previous studies of zebrafish [22,24,43]. With the bluntnose black bream, *bmp2b* were highly expressed in the IB, ribs and fins. Furthermore, a higher expression level of *bmp2a* was observed in the soft tissue of the spleen, muscle, and heart of the bluntnose black bream. This is also consistent with the previous

findings [27,44]. Our results suggested that *bmp2b* is preferentially expressed in fishbone-related tissue (such as IB, fins, and ribs), while *bmp2a* is mainly expressed in peripheral tissues (such as spleen, muscle, and heart). This could be because *bmp2a* genes and *bmp2b* genes play different roles in teleosts, and *bmp2b* may play an important role in the formation of IBs.

3.3. The Relationship between the Structure and Function of *bmp2*

Generally speaking, the structure of a protein determines its function [45]. In order to explore possible functional differences between the various subtypes of the *bmp2* gene family (*bmp2a* and *bmp2b*), we compared the amino acid sequences in the two proteins of multiple species. Through analysis of the important active sites of the *bmp2a* and *bmp2b* protein sequences of multiple species, we discovered that the *bmp2a* and *bmp2b* proteins in all species are conserved in amino acid around the 2HP ligand binding site, and the three-dimensional structures of *bmp2a* and *bmp2b* are also quite similar. It was also discovered that the spatial structure of zebrafish *bmp2a* and *bmp2b* contains a common receptor protein with a serine/threonine kinase activation site, and this indicates that *bmp2a* and *bmp2b* could have similar functions.

4. Methods and Materials

4.1. Acquisition of *bmp2a* and *bmp2b* for Nucleotide and Protein Sequences

A total of 37 different vertebrate species, containing nine mammals and 28 species of fish were observed for this study. Data were obtained using two different methods. For the first method, unreported *bmp2a* and *bmp2b* sequences from 20 species of fish, which were results from complete genome data produced by us and our collaborators (Table 1) was used. For the second method, published *bmp2a* and *bmp2b* sequences were downloaded from the public databases Ensembl and GenBank (Table S1). Potential homology-based *bmp2a* and *bmp2b* genes were recovered in detail from fish genomes through the use of tBLASTn [46] with an e-value of 10^{-5} . The BLAST results were then treated by Perl script in order to find the best hit for each alignment. Finally, GeneWise v2.2.0 [47] was employed in order to predict the *bmp2a* and *bmp2b* genes from the best hits.

4.2. Phylogenetic Analysis and Sequence Alignment

For further phylogenetic analysis, we utilized protein and nucleotide protein sequences of all collected *bmp2a* and *bmp2b* genes. To summarize, MAFFT software [48] was used in order to align protein sequences of *bmp2* and a maximum likelihood (ML) phylogenetic analysis was conducted, using the RAxML8.0.17 [49,50]. In addition, ML phylogenetic trees of the *bmp2a* and *bmp2b* isotypes were created, using their corresponding coding sequences by FastTree v2.1.7 [51]. We also downloaded a protein model of human *bmp2* from the public Protein Data Bank (PDB) in order to compare structural differences among the fish *bmp2a* and *bmp2b*.

4.3. Analyses of Conserved Synteny and Gene Structures

In order to assess the conservation of *bmp2a* and *bmp2b* genes, we observed different genes, which are present in the downstream and upstream regions of each *bmp2a* and *bmp2b* paralog. Associated genomic data was also received from GenBank and our lab, as mentioned above. The zebrafish genome was utilized as the reference standard for the discovery of any *bmp2a* and *bmp2b* downstream and upstream regions. The genome assemblies of a variety of species of fish were investigated using the BLAST software, and the best hit was chosen, using a Perl script. GeneWise v2.2.0 was applied in order to forecast *bmp2a* and *bmp2b* gene structures.

4.4. qRT-PCR Analysis of *bmp2a* and *bmp2b* for Nile Tilapia and Bluntnose Black Bream

In order to discover the distribution of tissues of *bmp2* genes, we chose two representative fish. The Nile tilapia with fewer IBs and the bluntnose black bream with more IBs were selected. Total RNA

was isolated from rib, IB, heart, spleen, liver, brain, and fin of the adult bluntnose black bream and Nile tilapia. With the use of an innovative method, the total RNA of IBs, ribs, and fins was isolated (Chinese Patent No.201310673534) for bone-related tissues. After using the PrimeScript[®] RT reagent kit, following DNase treatment, 1 µg total RNA was reverse-transcribed to single-strand cDNA (Takara, Dalian). The transcriptional levels of *bmp2a* and *bmp2b* in various adult tissues were measured using a quantitative real-time PCR (qRT-PCR) assay in a Roche LightCycler 480 (Roche, USA). qRT-PCR was conducted using an AceQ[®] qPCR SYBR[®] Green Master Mix (Vazyme, Nanjing, China). The β-actin was used as a control. PCR primers (Table S2) were utilized in order to intensify the blunt snout bream *bmp2a*, *bmp2b*, and the β-actin. PCR reactions were performed for each reaction well containing 20 µL PCR mixture, composed of 1 µL of each specific primer, 1 µL cDNA, 10 µL Green Master Mix (Table S2), and 7 µL ddH₂O. The reaction mixture was initially incubated at 95 °C for 3 min, followed by 40 cycles of 72 °C for 20 s, 60 °C for 30 s, and 95 °C for 15 s, and, finally at 4 °C for 5 min. The comparative expression ratio was normalized for the target genes with an internal reference β-actin gene and expression levels of *bmp2a* and *bmp2b*. Final results were determined using the 2^{-ΔΔCt} method [52].

4.5. Tertiary Structure and Function Prediction of Zebrafish *bmp2* Protein

As mentioned earlier [53], I-TASSER was used to forecast the tertiary functions and structures of *bmp2a* and *bmp2b*. Through C-score, the confidence of models is measured quantitatively, based on the importance of the convergence parameters of the structure assembly simulations and threading template alignments. Ordinarily, a C-score is in a range where a high value supports the corresponding model with high confidence.

5. Conclusions

On the whole, we gained new insights into the fish *bmp2* gene family by comparing genomics and transcriptional studies. We acknowledged the presence of *bmp2a* and *bmp2b* in fish and noted changes in gene copy numbers and the distribution of tissue between the species. We predicted that the spatial structure of *bmp2a* and *bmp2b* in zebrafish would contain a common receptor protein with a serine/threonine kinase activation site, indicating that the *bmp2a* and *bmp2b* genes have high homology, which also implies that the site is likely to be a part of maintaining the function of the *bmp2* gene family in zebrafish. Our results also provide strong evidence to support the findings that *bmp2b* is expressed primarily in fishbone-related tissues (such as IB, fins, and ribs) and *bmp2a* is expressed mostly in peripheral tissues (such as spleen, muscle, and heart), while *bmp2b* may play a significant role in the formation of IBs in teleosts.

Supplementary Materials: The following are available online: <http://www.mdpi.com/1422-0067/20/24/6137/s1>. Table S1, the accession numbers for *bmp2* sequences and Table S2, primers used in the present study.

Author Contributions: K.Z., L.L., and G.Y. conceived and designed the project; G.Y. and K.Z. participated in data analysis and figure preparation; K.Z. and G.Y. prepared the manuscript; K.Z. and L.L. revised the manuscript; and Z.Q., L.Z., R.L., H.K., and S.J. participated in discussions and provided suggestions.

Funding: This work was supported by the Outstanding Doctoral Fund of Zhongkai University of Agriculture and Engineering (No. KA180581215); Guangdong Provincial Bureau of Marine and Fisheries (No. GDME-2018C006), China-ASEAN Maritime Cooperation Fund (No. CAMC-2018F), Foundation of Guangdong Science and Technology Department (No. 2019B020215001).

Acknowledgments: We acknowledge the BGI Marine.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

TGF- β	Transforming growth factor beta
LAP	Latency associated peptide
MAFFT	Multiple sequence alignment based on fast Fourier transform
hao1	Hydroxyacid oxidase 1
tmx4	Thioredoxin related transmembrane protein 4
Napba	N-ethylmaleimide-sensitive factor attachment
crls1	Cardiolipin synthase 1
mcm8	Minichromosome maintenance 8
Gpcpd1	Glycerophosphocholine phosphodiesterase 1
trmt6	tRNA methyltransferase 6
fermt1	Fermitin family member 1
TAAR12j	Trace amine associated receptor 12j
TAAR12e	Trace amine associated receptor 12e
TAAR12g	Trace amine associated receptor 12g
TAAR12f	Trace amine associated receptor 12f
TAAR12a	Trace amine associated receptor 12a
TAAR12b	Trace amine associated receptor 12b
TAAR12c	Trace amine associated receptor 12c
TAAR10	Trace amine-associated receptor 10

References

- Patterson, C.; Johnson, G.D. Intermuscular bones and ligaments of teleostean fishes. *Smithson. Contrib. Zool.* **1995**, *559*, 1–83. [[CrossRef](#)]
- Johnson, G.D.; Patterson, C. The Intermuscular System of Acanthomorph Fishes: A Commentary. *Am. Mus. Novit.* **2001**, *3312*, 1–24. [[CrossRef](#)]
- Danos, B.N.; Staab, K.L. Can mechanical forces be responsible for novel bone development and evolution in fishes? *J. Appl. Ichthyol.* **2010**, *26*, 156–161. [[CrossRef](#)]
- Patricia Hernandez, L.; Bird, N.C.; Staab, K.L. Using zebrafish to investigate cypriniform evolutionary novelties: Functional development and evolutionary diversification of the kinethmoid. *J. Exp. Zool. Part B Mol. Dev. Evol.* **2010**, *308*, 625–641. [[CrossRef](#)] [[PubMed](#)]
- Staab, K.L.; Hernandez, L.P. Development of the cypriniform protrusible jaw complex in *Danio rerio*: Constructional insights for evolution. *J. Morphol.* **2010**, *271*, 814–825.
- Dong, Z.J.; Su, S.Y.; Zhang, J.Q. The research progress on intermuscular bones of teleosts. *Jiangsu Agric. Sci.* **2012**, *40*, 234–235.
- Bing, Z. On the myoseptal spines of the carp (*Cyprinus carpio* L.). *Acta Zool. Sin.* **1962**, *4*, 175–178.
- Ke, Z.H.; Zhang, W.; Jiang, Y.; Bao, B.L. Developmental morphology of the intermuscular bone in *Hypophthalmichthys molitrix*. *Chin. J. Zool.* **2008**, *43*, 88–96.
- Jiang, Y.; Yang, L.L.; Bao, B.L. The epicentrals in several lower teleosts. *J. Shanghai Fish. Univ.* **2008**, *17*, 493–496.
- Fang, L.C.; Li, X.J. A review of research on intermuscular bone formation in lower teleosts. *Fish. Sci.* **2013**, *32*, 749–752.
- Mulsant, P.; Lecerf, F.; Fabre, S.; Schibler, L.; Monget, P.; Lanneluc, I.; Pisselet, C.; Riquet, J.; Monniaux, D.; Callebaut, I.; et al. Mutation in bone morphogenetic protein receptor-IB is associated with increased ovulation rate in Booroola Merino ewes. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 5104–5109. [[CrossRef](#)] [[PubMed](#)]
- Souza, C.; Macdougall, C.C.; Campbell, B.; Mcneilly, A.; Baird, D. The Booroola (FecB) phenotype is associated with a mutation in the bone morphogenetic receptor type 1 B (BMPRI1B) gene. *J. Endocrinol.* **2001**, *169*, 1–6. [[CrossRef](#)] [[PubMed](#)]
- Kanzler, B.; Foreman, R.K.; Labosky, P.A.; Mallo, M. BMP signaling is essential for development of skeletogenic and neurogenic cranial neural crest. *Development* **2000**, *127*, 1095–1104. [[PubMed](#)]
- Miyazono, K.; Kamiya, Y.; Morikawa, M. Bone morphogenetic protein receptors and signal transduction. *J. Biochem.* **2010**, *147*, 35–51. [[CrossRef](#)]

15. Kawabata, M.; Imamura, T.; Miyazono, K. Signal transduction by bone morphogenetic proteins. *Cytokine Growth Factor Rev.* **1998**, *9*, 49–61. [[CrossRef](#)]
16. Riley, E.H.; Lane, J.M.; Urist, M.R.; Lyons, K.M.; Lieberman, J.R. Bone morphogenetic protein-2: Biology and applications. *Clin. Orthop. Relat. Res.* **1996**, *324*, 39–46. [[CrossRef](#)]
17. Wozney, J.M.; Rosen, V.; Celeste, A.J.; Mitscock, L.M.; Whitters, M.J.; Kriz, R.W.; Hewick, R.M.; Wang, E.A. Novel regulators of bone formation: Molecular clones and activities. *Science* **1988**, *242*, 1528–1534. [[CrossRef](#)]
18. Sebald, W.; Nickel, J.; Seher, A.; Mueller, T.D. Bone Morphogenetic Proteins. *Growth Factors* **2003**, *22*, 233–241.
19. Hogan, B.L. Bone morphogenetic proteins: Multifunctional regulators of vertebrate development. *Genes Dev.* **1996**, *10*, 1580–1594. [[CrossRef](#)]
20. Kingsley, D.M. The TGF-beta superfamily: New members, new receptors, and new genetic tests of function in different organisms. *Genes Dev.* **1994**, *8*, 133–146. [[CrossRef](#)]
21. Mowbray, C.; Hammerschmidt, M.; Whitfield, T.T. Expression of BMP signalling pathway members in the developing zebrafish inner ear and lateral line. *Mech. Dev.* **2001**, *108*, 179–184. [[CrossRef](#)]
22. Marques, C.; Fernández Monzón, I.; Rosa, J.; Viegas, M.; Cancela, M.L.; Laizé, V. Spatiotemporal expression and retinoic acid regulation of bone morphogenetic proteins 2, 4 and 16 in Senegalese sole. *J. Appl. Ichthyol.* **2014**, *30*, 713–720. [[CrossRef](#)]
23. Marques, C.L.; Fernandez, I.; Viegas, M.N.; Cox, C.J.; Martel, P.; Rosa, J.; Cancela, M.L.; Laize, V. Comparative analysis of zebrafish bone morphogenetic proteins 2, 4 and 16: Molecular and evolutionary perspectives. *Cell. Mol. Life Sci. CMLS* **2016**, *73*, 841–857. [[CrossRef](#)] [[PubMed](#)]
24. Crotwell, P.L.; Sommervold, A.R.; Mabee, P.M. Expression of bmp2a and bmp2b in late-stage zebrafish median fin development. *Gene Expr. Patterns GEP* **2004**, *5*, 291–296. [[CrossRef](#)] [[PubMed](#)]
25. Quint, E.; Smith, A.; Avaron, F.; Laforest, L.; Miles, J.; Gaffield, W.; Akimenko, M.A. Bone patterning is altered in the regenerating zebrafish caudal fin after ectopic expression of sonic hedgehog and bmp2b or exposure to cyclopamine. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 8713–8718. [[CrossRef](#)] [[PubMed](#)]
26. Kishimoto, Y.; Lee, K.; Li, Z.; Hammerschmidt, M.; Schulte-Merker, S. The molecular nature of zebrafish swirl: BMP2 function is essential during early dorsoventral patterning. *Development* **1997**, *124*, 4457–4466. [[PubMed](#)]
27. Zhang, W.Z.; Lan, T.; Nie, C.H.; Guan, N.N.; Gao, Z.X. Characterization and spatiotemporal expression analysis of nine bone morphogenetic protein family genes during intermuscular bone development in blunt snout bream. *Gene* **2018**, *642*, 116–124. [[CrossRef](#)]
28. Su, S.Y.; Dong, Z.J.; Qu, J.Q.; Liang, Z.Y.; Zhang, J.Q.; Ma, L.X.; Liu, W.; Xu, P.; Yuan, X.H. Molecular cloning and single nucleotide polymorphism analysis of IGF2a genes in the common carp (*Cyprinus carpio*). *Genet. Mol. Res. GMR* **2012**, *11*, 1327–1340. [[CrossRef](#)]
29. Rafael, M.S.; Laize, V.; Cancela, M.L. Identification of *Sparus aurata* bone morphogenetic protein 2: Molecular cloning, gene expression and in silico analysis of protein conserved features in vertebrates. *Bone* **2006**, *39*, 1373–1381. [[CrossRef](#)]
30. Clement, J.H.; Fettes, P.; Knochel, S.; Lef, J.; Knochel, W. Bone morphogenetic protein 2 in the early development of *Xenopus laevis*. *Mech. Dev.* **1995**, *52*, 357–370. [[CrossRef](#)]
31. Guyomard, R.; Boussaha, M.; Krieg, F.; Hervet, C.; Quillet, E. A synthetic rainbow trout linkage map provides new insights into the salmonid whole genome duplication and the conservation of synteny among teleosts. *BMC Genet.* **2012**, *13*, 1–12. [[CrossRef](#)] [[PubMed](#)]
32. Glasauer, S.M.K.; Neuhauss, S.C.F. Whole-genome duplication in teleosts fishes and its evolutionary consequences. *Mol. Genet. Genom.* **2014**, *289*, 1045. [[CrossRef](#)] [[PubMed](#)]
33. Meyer, A.; Van de Peer, Y. From 2R to 3R: Evidence for a fish-specific genome duplication (FSGD). *Bioessays* **2005**, *27*, 937–945. [[CrossRef](#)] [[PubMed](#)]
34. Kasahara, M.; Naruse, K.; Sasaki, S.; Nakatani, Y.; Qu, W.; Ahsan, B.; Yamada, T.; Nagayasu, Y.; Doi, K.; Kasai, Y. The medaka draft genome and insights into vertebrate genome evolution. *Nature* **2007**, *447*, 714–719. [[CrossRef](#)]
35. Jaillon, O.; Aury, J.M.; Brunet, F.; Petit, J.L.; Stange-Thomann, N.; Mauceli, E.; Bouneau, L.; Fischer, C.; Ozouf-Costaz, C.; Bernot, A. Genome duplication in the teleosts fish *tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. *Nature* **2004**, *431*, 946–957. [[CrossRef](#)]
36. Lv, Y.P.; Yao, W.J.; Chen, J.; Bao, B.L. Newly identified gene muscle segment homeobox C may play a role in intermuscular bone development of *Hemibarbus labeo*. *Genet. Mol. Res.* **2015**, *14*, 11224–11234. [[CrossRef](#)]

37. Nie, C.H.; Wan, S.M.; Tomljanovic, T.; Treer, T.; Hsiao, C.D.; Wang, W.M.; Gao, Z.X. Comparative proteomics analysis of teleost IBs and ribs provides insight into their development. *BMC Genom.* **2017**, *18*, 147. [[CrossRef](#)]
38. Danos, N.; Ward, A.B. The homology and origins of IBs in fishes: Phylogenetic or biomechanical determinants? *Biol. J. Linn. Soc.* **2012**, *106*, 607–622. [[CrossRef](#)]
39. Karsenty, G.; Wagner, E.F. Reaching a Genetic and Molecular Understanding of Skeletal Development. *Dev. Cell* **2002**, *2*, 389–406. [[CrossRef](#)]
40. Carrassón, M.; Matallanas, J. Morphometric characteristics of the alimentary tract of deep-sea Mediterranean teleosts in relation to their feeding habits. *Mar. Biol.* **1994**, *118*, 319–322. [[CrossRef](#)]
41. Sheikh, I.H.M. Scanning Electron Microscopic Studies of Gill Arches and Rakers in Relation to Feeding Habits of Some Fresh Water Fishes. *J. Basic Appl. Zool.* **2013**, *66*, 121–130.
42. Verma, S.R.; Tyagi, M.P.; Dalela, R.C. Morphological variations in the stomach of few teleosts in relation to food and feeding habits. *Gegenbaurs Morphol. Jahrb.* **1974**, *120*, 367. [[PubMed](#)]
43. Laforest, L.; Brown, C.W.; Poleo, G.; Géraudie, J.; Tada, M.; Ekker, M.; Akimenko, M.A. Involvement of the sonic hedgehog, patched 1 and bmp2 genes in patterning of the zebrafish dermal fin rays. *Dev. (Camb. Engl.)* **1998**, *125*, 4175–4184.
44. Wan, S.M.; Yi, S.K.; Zhong, J.; Nie, C.H.; Guan, N.N.; Zhang, W.Z.; Gao, Z.X. Dynamic mRNA and miRNA expression analysis in response to intermuscular bone development of blunt snout bream (*Megalobrama amblycephala*). *Sci. Rep.* **2016**, *6*, 31050. [[CrossRef](#)] [[PubMed](#)]
45. Skolnick, J.; Fetrow, J.S. From genes to protein structure and function: Novel applications of computational approaches in the genomic era. *Trends Biotechnol.* **2000**, *18*, 34–39. [[CrossRef](#)]
46. Pevsner, J. Basic Local Alignment Search Tool (BLAST). In *Bioinformatics and Functional Genomics*, 2nd ed.; Wiley: Hoboken, NJ, USA, 2009; pp. 100–138.
47. Birney, E.; Clamp, M.; Durbin, R. GeneWise and Genomewise. *Genom. Res.* **2004**, *14*, 988–995. [[CrossRef](#)]
48. Yamada, K.D.; Tomii, K.; Katoh, K. Application of the MAFFT sequence alignment program to large data-reexamination of the usefulness of chained guide trees. *Bioinformatics* **2016**, *32*, 3246–3251. [[CrossRef](#)]
49. Stamatakis, A. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **2006**, *22*, 2688–2690. [[CrossRef](#)]
50. Stamatakis, A.; Hoover, P.; Rougemont, J. A rapid bootstrap algorithm for the RAxML Web servers. *Syst. Biol.* **2008**, *57*, 758–771. [[CrossRef](#)]
51. Price, M.N.; Dehal, P.S.; Arkin, A.P. FastTree: Computing large minimum evolution trees with profiles instead of a distance matrix. *Mol. Biol. Evol.* **2009**, *26*, 1641–1650. [[CrossRef](#)]
52. Livak, K.J.; Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **2001**, *25*, 402–408. [[CrossRef](#)] [[PubMed](#)]
53. Yang, J.; Yan, R.; Roy, A.; Xu, D.; Poisson, J.; Zhang, Y. The I-TASSER Suite: Protein structure and function prediction. *Nat. Methods* **2015**, *12*, 7–8. [[CrossRef](#)] [[PubMed](#)]

