

RESEARCH

Open Access



Genome-wide identification and structural analysis of the *BMP* gene family in *Triplophysa dalaica*

Yizheng Zhang¹, Jinhui Yu², Rui Han¹, Zhigang Ma¹, Meng Zhang², Yikai Li², Yongtao Tang¹, Guoxing Nie^{1*} and Chuanjiang Zhou^{2*}

Abstract

Background Bone morphogenetic proteins (*BMPs*) are part of the transforming growth factor beta (TGF- β) superfamily and play crucial roles in bone development, as well as in the formation and maintenance of various organs. *Triplophysa dalaica*, a small loach fish that primarily inhabits relatively high elevations and cooler water bodies, was the focus of this study. Understanding the function of *BMP* genes during the morphogenesis of *T. dalaica* helps to clarify the mechanisms of its evolution and serves as a reference for the study of *BMP* genes in other bony fishes. The data for the *T. dalaica* transcriptome and genome used in this investigation were derived from the outcomes of our laboratory sequencing.

Results This study identified a total of 26 *BMP* genes, all of which, except for *BMP1*, possess similar TGF- β structural domains. We conducted an analysis of these 26 *BMP* genes, examining their physicochemical properties, subcellular localization, phylogenetic relationships, covariance within and among species, chromosomal localization, gene structure, conserved motifs, conserved structural domains, and expression patterns. Our findings indicated that three *BMP* genes were associated with unstable proteins, while 11 *BMP* genes were located within the extracellular matrix. Furthermore, some *BMP* genes were duplicated, with the majority being enriched in the GO:0008083 pathway, which is related to growth factor activity. It was hypothesized that genes within the *BMP1/3/11/15* subgroup (Group I) play a significant role in the growth and development of *T. dalaica*. By analyzing the expression patterns of proteins in nine tissues (gonad, kidney, gill, spleen, brain, liver, fin, heart, and muscle), we found that *BMP* genes play diverse regulatory roles during different stages of growth and development and exhibit characteristics of division of labor.

Conclusions This study contributes to a deeper understanding of *BMP* gene family member expression patterns in high-altitude, high-salinity environments and provides valuable insights for future research on the *BMP* gene family in bony fishes.

Keywords *Triplophysa dalaica*, Bone morphogenetic protein, Gene family, Gene expression

*Correspondence:

Guoxing Nie
niegx@htu.cn
Chuanjiang Zhou
chuanjiang88@163.com

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Introduction

Bone morphogenetic proteins (*BMPs*) constitute a highly potent class of growth factors within the larger TGF- β superfamily. The biological actions of *BMPs* were initially described in the 1960s when they were found to stimulate the formation of ectopic bone [1]. Since the successful cloning and identification of the first *BMP* genes in the 1980s [2], more than 40 *BMP* genes have been identified. *BMPs* are crucial for osteoblast differentiation and bone formation and have been implicated in cancer development by promoting the growth and invasion of cancer cells [3, 4]. These genes also regulate various physiological systems, including the circulatory, gastrointestinal, pulmonary, urinary, reproductive, and neurological systems [5], playing critical roles in embryonic development, growth, and differentiation [6]. Based on sequence homology, *BMP* family genes are divided into groups according to their roles in cell differentiation and growth: *BMP1/3/11/15* (Group I), *BMP12/13/14* (Group II), *BMP2/4/16* (Group III), *BMP9/10* (Group IV), and *BMP5/6/7/8* (Group V) [7, 8]. *BMP1* encodes a bone morphogenetic protein, which is a metalloproteinase responsible for regulating the deposition of extracellular matrix in vertebrate fibers. It also plays a vital role in various extracellular matrix metabolism and regulatory processes, contributing to essential biological functions [9]. *BMP2* and *BMP4* have well-established roles in critical processes such as embryonic development and the creation and differentiation of adipose precursor cells [10]. *BMP4* influences the differentiation and proliferation of lipogenic progenitor cells, with differentiated adipocytes secreting *BMP4* [11]. *BMP9* is involved in chronic liver disease (CLD), where its elevated expression promotes fibrosis in fibrotic livers [12]. *BMP7* promotes the neuronal differentiation of bone marrow mesenchymal stem cells (BMSCs), while *BMP8* plays a crucial role in preventing premature luteinization of granulosa cells for ovarian follicle development [13, 14]. *BMP* genes exhibit variable expression in various carp tissues, including the gills, gut, liver, spleen, skin, heart, gonads, muscle, kidneys, cephalic kidneys, brain, and blood [8].

T. dalaica is found primarily in northern China in the Yellow River tributaries and in artesian water basins such as Inner Mongolia's Dali Lake [15]. *T. dalaica* thrives in high-altitude environments characterized by low water temperatures, especially in high-salt alkaline waters such as still lakes and river slopes. This species exhibits remarkable adaptations to plateau conditions, including its ability to acclimate to low-oxygen aquatic habitats and endure cold temperatures [15]. This harsh environment requires specific physiological adaptations for survival and reproduction, and the *BMP* gene family is involved in many physiological and metabolic processes, including

bone and cartilage formation. Therefore, investigating the role of *BMP* genes in the adaptation of this species to extreme conditions will help to unravel the underlying mechanisms of its molecular adaptation and evolution and provide some assistance for subsequent studies of the *BMP* gene family in scleractinian fishes.

In this study, we identified 26 *BMP* genes in *T. dalaica* based on its whole genome. We also explored the potential roles and regulatory mechanisms of these *BMP* genes in *T. dalaica* through bioinformatics analysis and comparative studies with other vertebrates. These findings lay the foundation for further exploration of *BMP* gene functions in *T. dalaica* and contribute to the understanding of its physiological characteristics at the molecular level, providing support for the protection of germplasm resources in *T. dalaica*.

Materials and methods

Materials

The data for the *T. dalaica* transcriptome (Table S1) and genome (PRJNA624716) used in this investigation were derived from the outcomes of our laboratory sequencing. The *T. dalaica* specimens utilized in this study were collected from Dalinuoer Lake (43°22'43"N, 116°39'24"E), Inner Mongolia [16]. For the sequencing of *T. dalaica*, we employed PacBio sequencing and Hi-C technology, which resulted in chromosome-level genomic and transcriptomic data [15]. The de novo assembled genome had a total size of 607.91 Mb, with a contig N50 of 9.27 Mb. The total data volume collected was 126.5 Gb and 106 Gb.

BMP gene identification and sequence analysis

The CDSs and amino acid sequences of four species, *Danio rerio*, *Homo sapiens*, *Cyprinus carpio*, and *Xenopus laevis*, were obtained from publicly available data in the Ensembl database (<http://asia.ensembl.org/>) and the National Center for Biotechnology Information (NCBI) website (<http://www.ncbi.nlm.nih.gov/genbank/>). We conducted a BLASTP alignment search against the *T. dalaica* genome database utilizing *BMP* gene protein sequences from these species as query sequences, with an e-value cutoff set at 1e-5. To confirm the accuracy of the identification of the candidate genes in *T. dalaica*, a reverse blast was carried out using the NCBI database. The protein sequences of *T. dalaica*'s *BMP* genes were established by analyzing the intersection of the findings. We assessed the physical and chemical properties of *BMP* family members, including amino acid count, theoretical isoelectric point, molecular weight, and instability index, using TBtools software [17]. Subcellular localization was investigated using the WoLF PSORT website (<https://www.genscript.com/wolf-psort.html>), while the

identification of open reading frames (ORFs) for *BMP* genes was executed via the utilization of ORF Finder (<https://www.ncbi.nlm.nih.gov/gorf/gorf.html>) [18].

Multiple sequence comparisons and phylogenetic analysis of *BMP* genes

The amino acid sequences of conserved domains in *BMP* genes belonging to the TGF superfamily were extracted using SAMRT [19]. Multiple sequence alignment was performed using BioEdit [20], aligning the protein sequences of *BMP* genes from *D. rerio*, *H. sapiens*, *C. carpio*, and *X. laevis* obtained from the NCBI and Ensembl databases (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>) [21]. A phylogenetic tree was constructed in MEGA11 using the matched sequences and the neighbor-joining technique [22], with a bootstrap repeat count of 1000. The generated phylogenetic tree was further refined using the ChiPlot online tool (<https://www.chiplot.online/>) [23]. The *BMP* protein sequences of the four

species *D. rerio*, *H. sapiens*, *C. carpio*, and *X. laevis* are given in Table S3.

Gene nomenclature

The *BMP* genes in *T. dalaica* were identified through a combination of multiple sequence alignment and phylogenetic analyses. *BMP* genes in *T. dalaica* were named based on *BMP* gene sequences from *D. rerio*, *H. sapiens*, *C. carpio*, and *X. laevis*. In instances where multiple *T. dalaica* genes were grouped with the other four species, a Latin numeral suffix was appended to each gene name. The final nomenclature of the *BMP* genes in *T. dalaica*, along with that of the selected species, can be found in Table 1.

Synten analysis

To explore gene duplication events within *T. dalaica*, we conducted an intraspecific collinearity analysis. Additionally, a comparative collinearity analysis map was constructed between *T. dalaica* and two additional

Table 1 *BMP* gene families in the genomes of the five vertebrates

<i>Triplophysa dalaica</i>	<i>Danio rerio</i>	<i>Xenopus laevis</i>	<i>Cyprinus carpio</i>	<i>Homo sapiens</i>
26	22	16	44	17
<i>BMP1a</i>	<i>BMP1a</i>	<i>BMP1</i>	<i>BMP1a-1/2</i>	<i>BMP1</i>
<i>BMP1b-1</i>	<i>BMP1b</i>		<i>BMP1b-1/2</i>	
<i>BMP1b-2</i>				
<i>BMP2a</i>	<i>BMP2a</i>	<i>BMP2</i>	<i>BMP2a-1/2</i>	<i>BMP2</i>
<i>BMP2b</i>	<i>BMP2b</i>		<i>BMP2b-1/2</i>	
<i>BMP3a</i>	<i>BMP3a</i>	<i>BMP3a</i>	<i>BMP3a-1/2</i>	<i>BMP3a</i>
<i>BMP3b-1</i>	<i>BMP3b</i>	<i>BMP3b</i>	<i>BMP3b-1/2</i>	<i>BMP3b</i>
<i>BMP3b-2</i>				
<i>BMP4</i>	<i>BMP4</i>	<i>BMP4</i>	<i>BMP4-1/2</i>	<i>BMP4</i>
<i>BMP5</i>	<i>BMP5</i>	<i>BMP5</i>	<i>BMP5-1/2</i>	<i>BMP5</i>
<i>BMP6a</i>	<i>BMP6</i>	<i>BMP6</i>	<i>BMP6-1/2</i>	<i>BMP6</i>
<i>BMP7a</i>	<i>BMP7a</i>	<i>BMP7a</i>	<i>BMP7a-1/2</i>	<i>BMP7</i>
<i>BMP7b-1</i>	<i>BMP7b</i>	<i>BMP7b</i>	<i>BMP7b-1/2</i>	
<i>BMP7b-2</i>				
	<i>BMP8a</i>	<i>BMP8a</i>	<i>BMP8a-1/2</i>	<i>BMP8a</i>
<i>BMP9</i>	<i>BMP9</i>	<i>BMP9</i>	<i>BMP9-1/2</i>	<i>BMP9</i>
<i>BMP10a</i>	<i>BMP10a</i>	<i>BMP10</i>	<i>BMP10a-1/2</i>	<i>BMP10</i>
<i>BMP10b</i>	<i>BMP10b</i>		<i>BMP10b-1/2</i>	
<i>BMP11-1</i>	<i>BMP11</i>	<i>BMP11</i>	<i>BMP11-1/2</i>	<i>BMP11</i>
<i>BMP11-2</i>				
<i>BMP12</i>	<i>BMP12</i>		<i>BMP12-1/2</i>	<i>BMP12</i>
<i>BMP13a</i>	<i>BMP13a</i>	<i>BMP13</i>	<i>BMP13a-1/2</i>	<i>BMP13</i>
<i>BMP13b</i>	<i>BMP13b</i>		<i>BMP13b-1/2</i>	
<i>BMP14</i>	<i>BMP14</i>	<i>BMP14</i>	<i>BMP14-1/2</i>	<i>BMP14</i>
<i>BMP15-1</i>	<i>BMP15</i>	<i>BMP15</i>	<i>BMP15-1/2</i>	<i>BMP15</i>
<i>BMP15-2</i>				
<i>BMP16</i>	<i>BMP16</i>		<i>BMP16-1/2</i>	

osteichthyan species, *D. rerio*, and *C. carpio*. Genomic data and annotation information for *D. rerio* and *C. carpio* were obtained from the NCBI and Ensembl databases, respectively. For both intraspecific and interspecific collinearity analyses, we used TBtools software for analysis and diagramming [17].

Chromosomal localization and gene structure analysis of the *BMP* genes

The chromosomal localization of *T. dalaica* *BMP* genes was ascertained by importing the complete genome annotation file of *T. dalaica* into TBtools software in conjunction with the *BMP* gene sequence data [17]. Subsequently, the physical positions of the chromosomes corresponding to the *BMP* genes were visualized.

Predictions of the secondary structure elements of *BMP* protein sequences, including α -helices, β -turns, extended strands, and disordered coils, were made using SOPMA (<http://pbil.ibcp.fr/>) [24]. Subcellular localization prediction of *BMP* protein sequences was performed using Cell-PLoc2.0 <http://www.csbio.sjtu.edu.cn/bioinf/Cell-PLoc-2/> [25]. Information on *BMP* gene exons and introns was extracted from the *T. dalaica* whole-genome annotation file, and visualization was carried out using TBtools (<http://tbtools.bioinfodata.com/>) [17].

Analysis of conserved motifs and conserved structural domains of the *BMP* genes

MEME (<http://meme-suite.org/>) was used to predict protein motif patterns in the *T. dalaica* *BMP* gene family. The maximum motif value was set to 10, with a width range of 6 to 50. The results were visualized using TBtools. Conserved domain structures of *BMP* genes were predicted using the NCBI-Batch-CD-search tool (<https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi>) and the SMART tool (<http://smart.embl-heidelberg.de>). TBtools and the *BMP* protein phylogenetic tree were used for visualization of conserved *BMP* protein motifs and domains [17].

GO enrichment analysis of *BMP* genes

GO enrichment analysis of the *BMP* genes was conducted using assembled transcripts obtained from whole-genome sequencing of *T. dalaica*. The basic unit of GO enrichment was term, and all differently expressed genes were mapped to each term in the Gene Ontology database (<http://www.geneontology.org/>) to obtain the number of genes in each term. Apply chi-square test or hypergeometric test to identify GO terms that are significantly enriched in differentially expressed genes compared to the entire genome background. Determine the main biological functions performed by the differentially

expressed genes through significant GO functional enrichment analysis.

Expression patterns of *BMP* genes

The original image data derived from sequencing were converted into sequence data via base calling, yielding raw reads, which were saved in fastq file format. We used Trimmomatic software (<http://www.usadellab.org/cms/index.php?page=trimmomatic>) to filter the raw data, remove reads with adaptors, discard reads with more than 10% 'N' bases, and eliminate reads where more than 50% of bases had a quality score less than 20 [26]. Afterward, we applied FastQC software (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc>) for quality control of the clean data, and ultimately, Trinity software was used for the assembly of these clean data. The transcriptome assembled by Trinity was utilized as the reference sequence (ref), onto which clean reads of each sample were mapped. We implemented the RSEM software for this process, with the Bowtie2 parameter set to a mismatch of 0 (the default Bowtie2 parameter) [27]. RSEM provided statistical analysis of the bowtie alignment results, furnishing the read count numbers for each gene in every sample [28]. These counts were subsequently subjected to TPM (transcripts per million) conversion—a normalized value indicating the quantity of transcripts for a specific gene or transcript per million total transcripts within the sample—which facilitated the analysis of gene expression levels. Ultimately, this analysis resulted in the quantified abundances of *BMP* gene transcripts. A heatmap of gene expression was generated using the online tool ChiPlot (<https://www.chipplot.online/>) [23, 29].

We also performed real-time fluorescence quantitative PCR (qRT-PCR) analysis of selected tissues and genes to verify the accurate transcriptome abundance analysis of the gene expression trends. Three specimens of *T. dalaica* from the saline-alkali waters of Dali Lake (Inner Mongolia) were collected, and RNA was extracted from six tissues—brain, liver, spleen, gonad, kidney, and gill—followed by cDNA synthesis and RT-PCR. Specific primers for certain members of the *BMP* gene family in *T. dalaica* were designed using Primer 5.0 (Table S2). The qRT-PCR experiments were conducted using the TransStart Top Green qPCR SuperMix kit. The amplification system consisted of a 10 μ L reaction mixture with 1 μ L cDNA, 0.2 μ L forward primer, 0.2 μ L reverse primer, 5 μ L TransStart Top Green qPCR SuperMix, and 3.6 μ L ddH₂O. The experiments were performed on a LightCycler 96 Real-Time PCR System with three biological and three technical replicates. β -actin was used as the internal control gene, and the data were processed using the $2^{-\Delta\Delta Ct}$ method, where $\Delta\Delta Ct$ = (average Ct value of

the target gene in treated samples—average Ct value of the internal control gene in treated samples)—(average Ct value of the target gene in control samples—average Ct value of the internal control gene in control samples) [30]. With the kidney group set as 1 for the same gene, the relative expression levels were calculated and graphed in GraphPad Prism 9. The relative expression levels of all the genes were analyzed via one-way ANOVA followed by Duncan's multiple range test, with $P < 0.05$ indicating statistical significance. The final results are presented in Fig. 1.

Results

BMP gene identification and characterization

Utilizing the whole-genome data of *T. dalaica*, we conducted a comparison to eliminate duplicate sequences, considering other species such as *D. rerio*, *H. sapiens*, *C. carpio*, and *X. laevis*. Furthermore, by integrating information from relevant literature [31, 32], we successfully identified a total of 26 BMP gene sequences (*BMP1a*–*BMP16*) in *T. dalaica*, which were subsequently categorized into five groups: *BMP1/3/11/15* (Group I), *BMP12/13/14* (Group II), *BMP2/4/16* (Group III), *BMP9/10* (Group IV), and *BMP5/6/7* (Group V).

These 26 genes were discovered within the *T. dalaica* genome and are described in Table 1. We also analyzed the physicochemical characteristics of the 26 protein sequences, which are presented in Table 2. The study revealed that the open reading frame (ORF) lengths of the 26 BMP genes varied from 738 (*BMP15-1*) to 3069 (*BMP1b-1*) base pairs (bp), with the number of exons ranging from 1 to 21. The predicted isoelectric points ranged from 4.97 to 9.94, with 11 sequences exhibiting acidity and 15 showing alkalinity. Interestingly, most BMP proteins in *T. dalaica* tend to precipitate under highly alkaline conditions, possibly due to the alkaline environment of the species. The predicted code sequence of the BMP gene family varies in the number of amino acids, ranging from 245 (*BMP15-1*) to 1022, with relative molecular masses varying from 28.07 kDa to 115.35 kDa. Notably, the average affinity coefficients for *T. dalaica* BMP gene family proteins were consistently negative, indicating affinities for each of these proteins. Additionally, a majority of the BMP proteins displayed a genetic instability index above 40, indicating stability, whereas those considered unstable had an index less than 30.

Subcellular localization findings, as presented in Table 2, revealed that 11 BMP proteins were found in the extracellular plasma, while 9 BMP proteins were located in the plasma membrane. The remaining BMP proteins were detected in various locations, including the cell nucleus, fibromyalgia, endothelial plasma, and cellulose.

BMP Genetic System Development Analysis

To explore the evolutionary correlation of the BMP gene family in *T. dalaica*, we generated a comprehensive phylogenetic tree using protein sequences derived from *D. rerio*, *H. sapiens*, *C. carpio*, and *X. laevis* (Fig. 2). The phylogenetic tree analysis revealed distinct clustering of the BMP gene family in *T. dalaica* and other species into five groups: *BMP1/3/11/15* (Group I), *BMP12/13/14* (Group II), *BMP2/4/16* (Group III), *BMP9/10* (Group IV), and *BMP5/6/7* (Group V). Notably, the *BMP8a* protein sequences of *D. rerio*, *H. sapiens*, *C. carpio*, and *X. laevis* formed a distinct branch. The majority of the BMP genes in *T. dalaica* were categorized into Group I. Phylogenetic analysis revealed a notable level of similarity between *T. dalaica* and two other species, *D. rerio* and *C. carpio*, both of which belong to the Cypriniformes order, in contrast to humans and the amphibian *X. laevis* (Fig. 2).

BMP gene collinearity analysis

We identified three occurrences of gene segment duplication in *T. dalaica*, all resulting from gene fragment duplications; these duplications were located on chromosomes 3 and 21, 8 and 10, and 18 and 22 (Fig. 3). *BMP1a* and *BMP1b-2* underwent gene segment duplications and belonged to the same gene group, with their encoded proteins exhibiting activities similar to those of *BMP11-1* and *BMP11-2*. Interestingly, segmental duplication of *BMP7a* occurred alongside the non-BMP gene. The *BMP7* gene belongs to the TGF- β superfamily, and Fig. 4 shows that *BMP7a* is abundantly expressed in fins, muscles and gonads, which suggests that *BMP7a* may be involved in the realization of locomotor function in *T. dalaica* and influence sex differentiation in *T. dalaica*. Figure 3 shows that *BMP7a* is duplicated with the *GDF15* gene, possibly because it belongs to the same TGF- β superfamily as the *BMP7a* gene is, and the *GDF15* gene has an assisting role in the physiological function of *BMP7a*. Since *T. dalaica*, *D. rerio*, and *C. carpio* are all Cypriniformes, we explored interspecies collinearity among the three species and revealed a significant collinearity link. Most of the *T. dalaica* BMP genes exhibited collinearity with *D. rerio* and *C. carpio* genes (Fig. 5).

Chromosomal localization and gene structure analysis of the BMP genes

We determined the chromosomal location of the BMP gene family in *T. dalaica* using TBtools software [17], and the results are presented in Fig. 6. This investigation revealed that the 26 BMP genes in *T. dalaica* were dispersed among 19 chromosomes (1, 2, 4, 6, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 20, 21, 22, 23, 25). Intriguingly, an overrepresentation of BMP genes was observed on

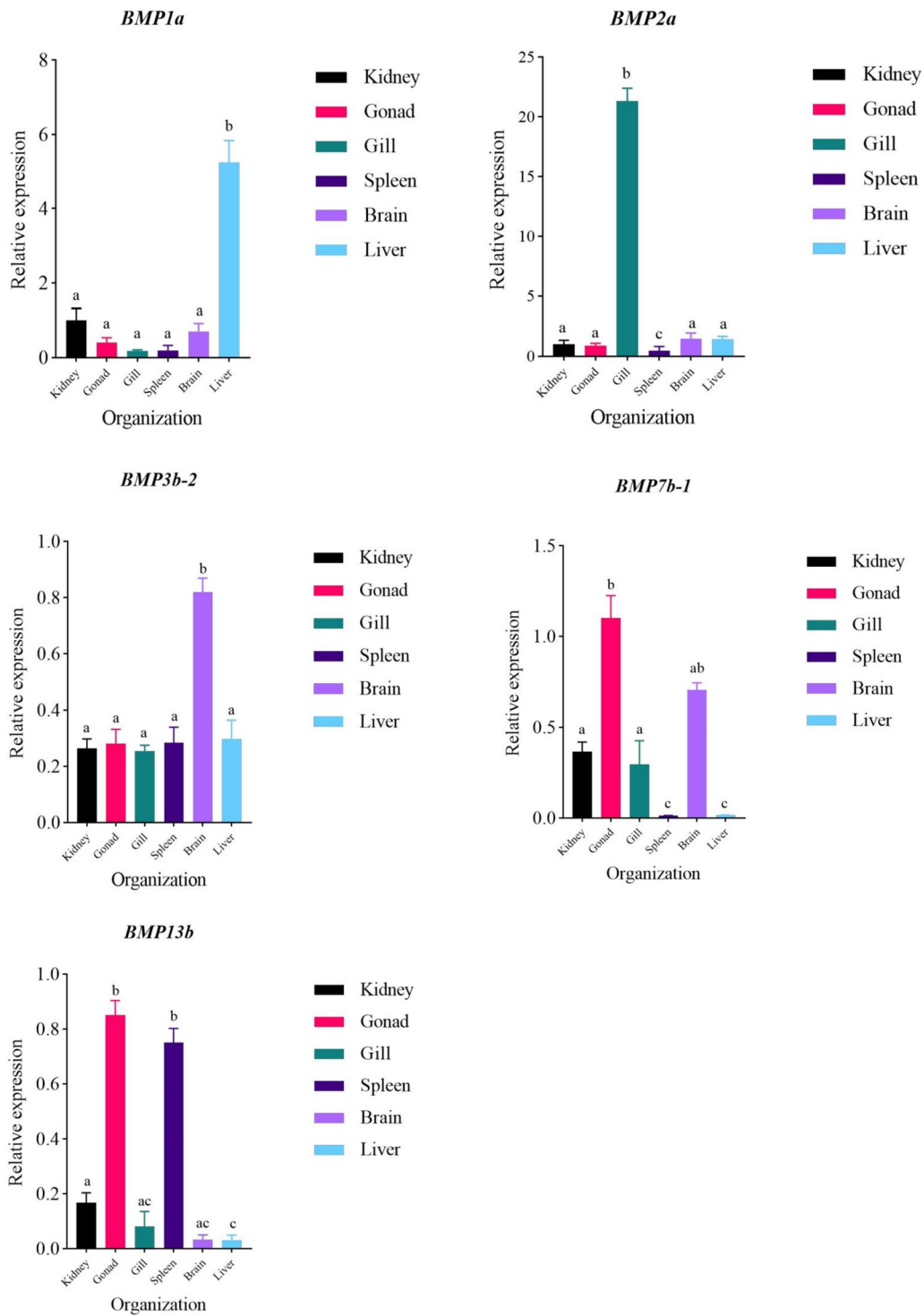


Fig. 1 Expression of different genes in each tissue. **a,b,c** Three letters and the same letter indicate no significant difference, and different letters indicate significant differences

chromosomes 4, 11, and 15, suggesting a non-uniform and stochastic distribution pattern within the species. Additionally, we observed close proximity of *BMP9* and

BMP3b-2 on chromosome 17, suggesting the potential formation of a gene cluster resulting from early gene duplication events (Fig. 6) [33, 34].

Table 2 Characteristics of BMP Gene Family Members in the Genome of *Triplophysa dalaica*

Name	ORF (bp)	Exon numbers	Isoelectric Point	Molecular weight (kDa)	Amino Acid	Instability Index	Aliphatic index	Average coefficient of hydrophobicity	Subcellular localization	NCBI accession number
<i>BMP1a</i>	2931	20	6.22	109.77	976	40.68	63.44	-0.547	plas	OR7333219
<i>BMP1b-1</i>	3069	21	5.62	115.35	1022	42.42	62.67	-0.589	extr	OR7333217
<i>BMP1b-2</i>	2928	20	5.57	108.73	975	44.43	67.6	-0.426	plas	OR7333218
<i>BMP2a</i>	1176	2	8.58	43.93	391	57.04	80.43	-0.427	plas	OR7333221
<i>BMP2b</i>	1248	2	8.61	46.97	415	53.64	78.94	-0.524	plas	OR7333220
<i>BMP3a</i>	1356	3	9.36	51.76	451	53.45	77.38	-0.566	extr	OR7333224
<i>BMP3b-1</i>	1341	3	9.08	51.27	446	51.25	80	-0.462	extr	OR7333222
<i>BMP3b-2</i>	1314	3	9.07	49.72	437	54.51	79.43	-0.397	plas	OR7333223
<i>BMP4</i>	1206	2	6.68	45.8	401	49.64	84.56	-0.528	plas	OR7333225
<i>BMP5</i>	1338	7	9.22	50.41	445	46.86	74.76	-0.461	mito	OR7333226
<i>BMP6a</i>	1254	7	8.07	48.09	417	53.05	77.19	-0.481	extr	OR7333227
<i>BMP7a</i>	1278	7	7.25	48.49	425	49.41	79.13	-0.376	extr	OR7333230
<i>BMP7b-1</i>	1284	7	6.19	48.7	427	48.51	84.05	-0.292	plas	OR7333228
<i>BMP7b-2</i>	1284	7	6.19	48.7	427	48.51	84.05	-0.292	plas	OR7333229
<i>BMP9</i>	1197	2	5.57	45.77	398	38.65	84.42	-0.44	cyto	OR7333231
<i>BMP10a</i>	1317	3	6.26	50.15	438	47.32	75.64	-0.58	mito	OR7333233
<i>BMP10b</i>	1422	2	4.97	52.66	473	43.7	79.34	-0.454	extr	OR7333232
<i>BMP11-1</i>	1113	3	5.64	41.44	370	62.39	79.32	-0.419	extr	OR7333234
<i>BMP11-2</i>	1128	3	6.22	42.06	375	56.11	80.61	-0.373	extr	OR7333235
<i>BMP12</i>	1290	2	9.54	48.03	429	56.84	79.28	-0.457	nucl	OR7333236
<i>BMP13a</i>	1224	2	9.15	46.39	407	53.38	76.93	-0.545	extr	OR7333238
<i>BMP13b</i>	1248	2	9.42	47.39	415	56.8	76.39	-0.543	extr	OR7333237
<i>BMP14</i>	1530	2	9.94	57.19	509	38.97	80.47	-0.44	E.R	OR7333239
<i>BMP15-1</i>	738	1	9.05	28.07	245	53.49	71.14	-0.56	nucl	OR7333240
<i>BMP15-2</i>	1287	2	8.45	49.87	428	57.05	79.67	-0.57	plas	OR7333241
<i>BMP16</i>	1275	2	9.43	47.34	424	52.12	79.98	-0.45	extr	OR7333242

Abbreviations used in this table: *plas* Plasma membrane, *extr* Extracell, *mito* Mitochondria, *cyto* Cytoplasm, *nucl* Nucleus, *E.R.* Endoplasmic reticulum

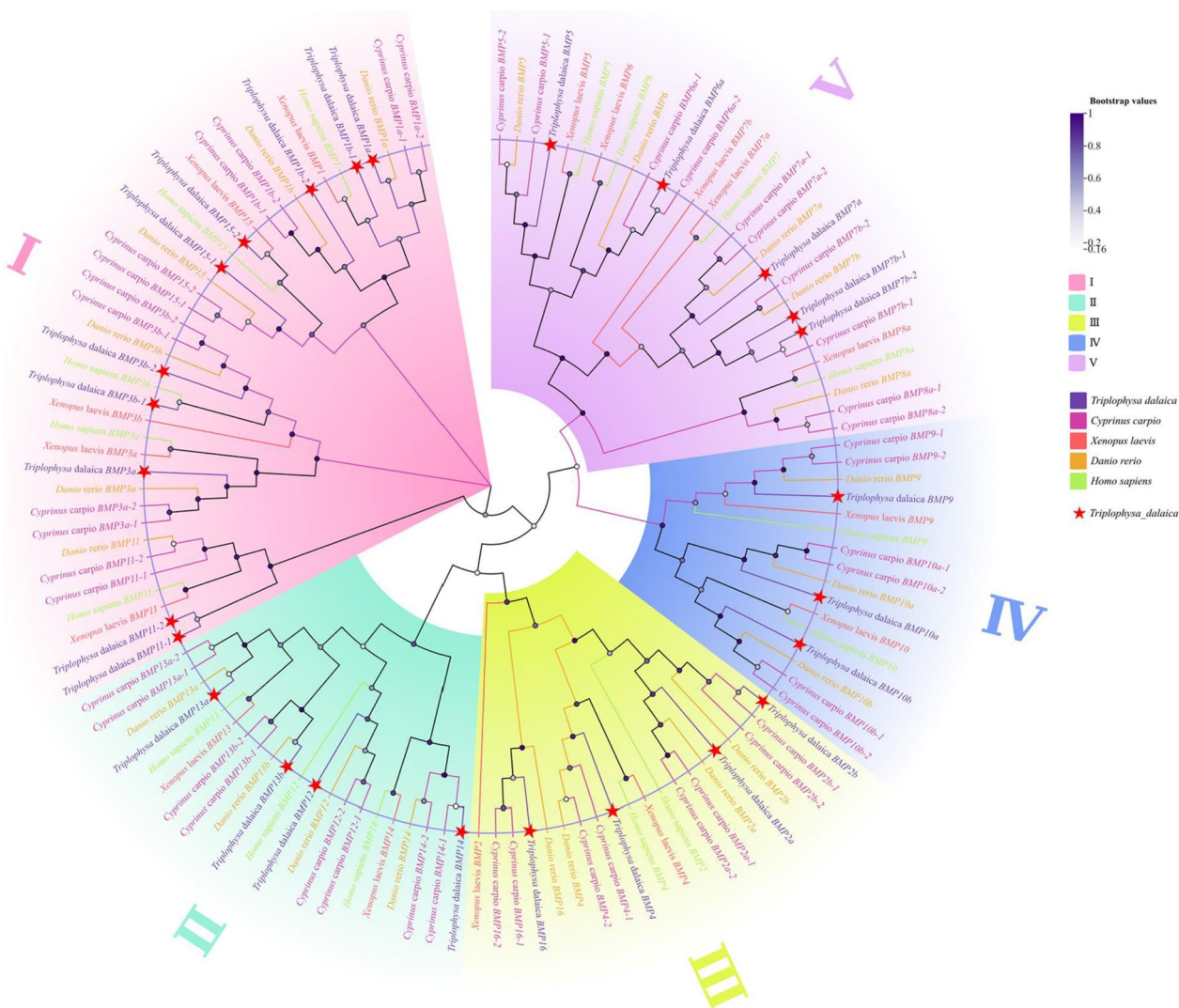


Fig. 2 Phylogenetic analysis of the *BMP* genes of *T. dalaica* and selected species. The different colors on the outer ring in the figure represent the different groups, the different species are represented by different colors, and the *BMP* gene of *T. dalaica* is marked by a red star

To investigate the structural variation of the *BMP* gene family, we used TBtools to visualize the exon–intron architectures of the 26 *BMP* genes (Fig. 7) [17]. The results showed that *BMP1b-1* had the highest number of exons (21), while the number of other exons ranged from 2 to 21. We also generated an NJ phylogenetic tree to examine *T. dalaica* *BMP* gene protein sequences and assess gene arrangement. *BMP* genes with similar structural properties were grouped in the evolutionary tree. These findings shed light on the diversity and evolution of the *BMP* gene family. Furthermore, the secondary structure prediction indicated that *T. dalaica* *BMP* gene proteins contain α -helices, β -turns, extended strands, and irregular coils (Table 3). The percentage of α -helices ranged from 11.17% (*BMP1a*) to 32.41% (*BMP9*), β -turns

from 1.23% (*BMP13a*) to 5.84% (*BMP1a*), extended strands from 13.05% (*BMP12*) to 27.77% (*BMP1a*), and irregular coils from 46.73% (*BMP9*) to 62.45% (*BMP15-1*). The secondary structure prediction suggested that irregular coils are the predominant structural component of *T. dalaica* *BMP* genes.

Analysis of conserved motifs and structural domains of the *BMP* genes

We utilized the MEME online program to investigate the conserved protein motifs of the *BMP* gene family in *T. dalaica*, resulting in the identification of ten conserved protein motifs (Fig. 8(b)). Notably, motif 1 featured a conserved sequence of LYVD-FKDJGWDDWIIAPEGYEAYCEGEC, motif 2 was

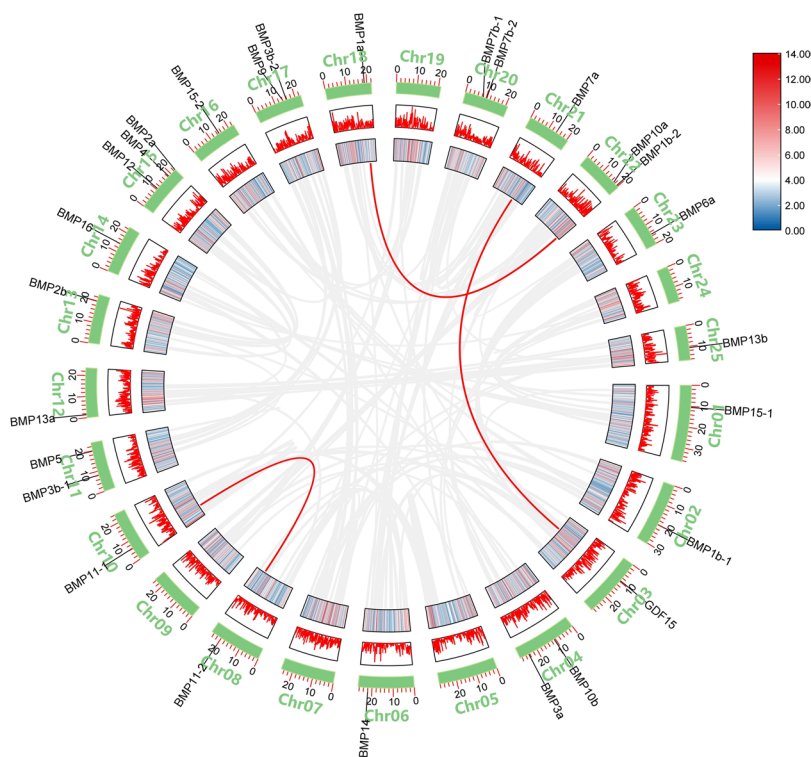


Fig. 3 Analysis of *BMP* gene covariance in *T. dalaica*. The gray line indicates the genome covariance of *T. dalaica*, and the colored line connecting the *BMP* genes indicates the duplication of *BMP* genes in *T. dalaica*. The position of the *BMP* genes on the chromosomes is indicated by the short black line, and the density of the genes on each chromosome is shown at the same time

characterized by a conserved sequence of CCVPT-KLSPISVLYLDDSENVVLKKY, and motif 4 exhibited a conserved sequence of EBMVVEESGCR, all with a length of 11 amino acids. The analysis revealed that genes with closer evolutionary relationships displayed more similar motif architectures. While motif 9 was found to be exclusively conserved in *BMP1a*, *BMP1b-1*, and *BMP1b-2*, motifs 1, 2, and 4 were detected in all the other family members, underscoring their notable conservation within the core functional domains of the *BMP* genes (Fig. 8). The distribution of motifs aligns with the phylogenetic tree, demonstrating that genes within the same subgroup exhibited similar motif patterns (Fig. 8(a)).

Moreover, as depicted in Fig. 9, the *BMP* gene family members in *T. dalaica* that clustered together under the same branch exhibited notable similarities in terms of conserved protein domains. These observations are consistent with the observed distribution patterns and the results of the phylogenetic tree analysis. Notably, the *BMP1a*, *BMP1b-1*, and *BMP1b-2* genes produced proteins with distinct conserved domains ("ZnMc," "CUB," "Fxa"), arising from their unique gene structures. *BMP1*, classified as a zinc-dependent metalloproteinase, does not belong to the TGF- β superfamily [35, 36].

GO enrichment analysis of the *BMP* genes

We conducted a GO enrichment analysis, and the results presented in Table 4 revealed the involvement of the *BMP* gene family in *T. dalaica* in a total of 15 distinct GO pathways. Among these pathways, eight were associated with molecular functions (F), six were related to biological processes (P), and one was linked to cellular components (C). Notably, the GO pathways related to molecular functions were primarily associated with growth and development processes, including activities such as metalloproteinase activity, calcium ion binding, metalloendopeptidase activity, zinc ion binding, growth factor activity, receptor binding activity, and protein domain-specific binding. The GO pathways linked to biological processes predominantly included protein hydrolysis, cellular cytoskeleton organization, cell migration, and cell adhesion. The examination of the *T. dalaica* *BMP* gene phylogenetic tree (Fig. 7) suggested that the *BMP1/3/11/15* (Group I) subfamily is involved primarily in these biological processes. Most *BMP* genes were enriched in the GO:0008083 pathway, associated with growth factor activity, aligning with the characteristic function of the *BMP* gene family, suggesting that the *BMP1/3/11/15* (Group I) subfamily predominantly plays a role in *T. dalaica*'s growth and development processes.

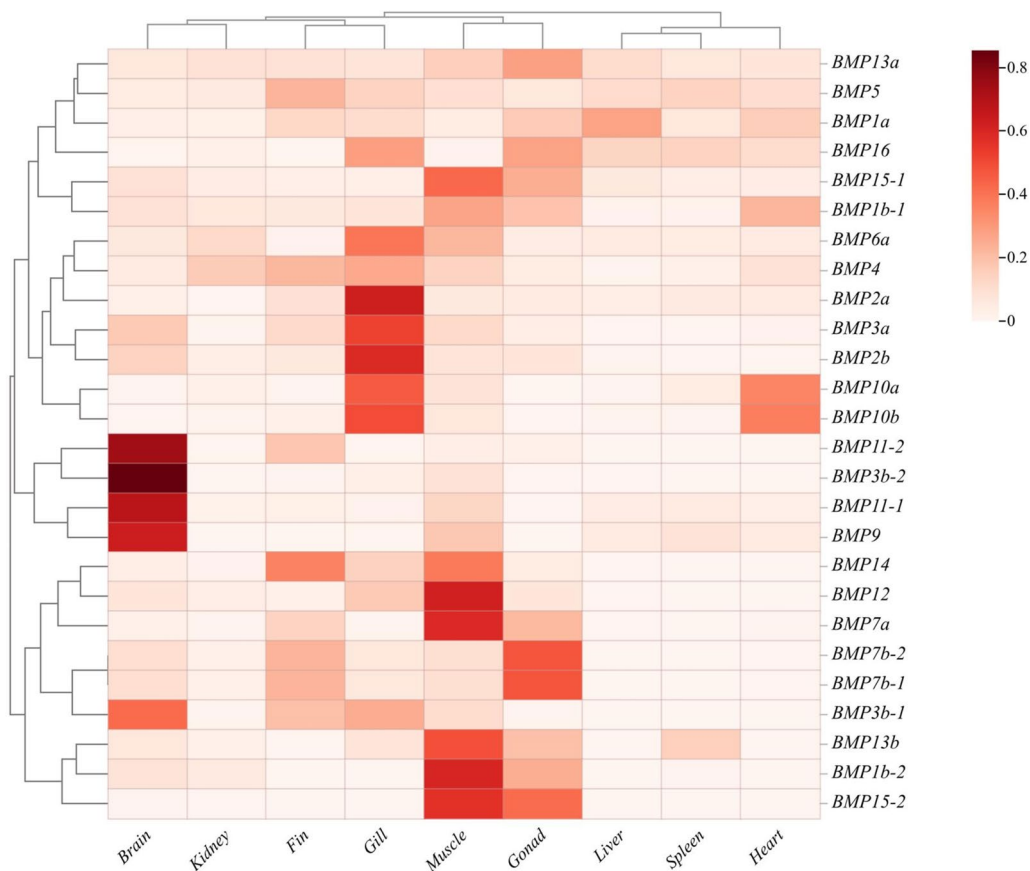


Fig. 4 Analysis of the *BMP* gene expression pattern in *T. dalaica*. The square color scale of the heatmap indicates the TPM values after row normalization; the specific values are marked in the color block, and the maximum value is indicated in red

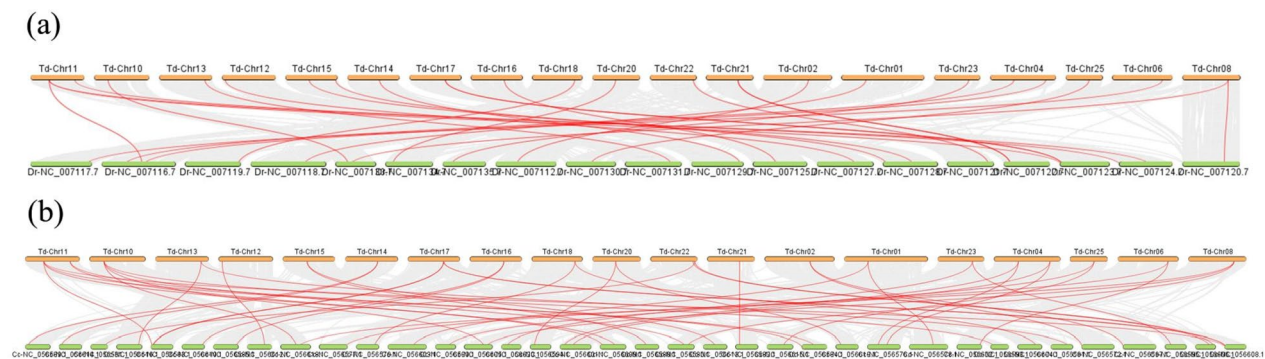


Fig. 5 Interspecies covariance analysis of *T. dalaica* with *D. rerio* and *C. carpio*. **a** represents the covariance results of *T. dalaica* with *D. rerio*, **b** represents the covariance results of *T. dalaica* with *C. carpio*, gray lines represent the covariance blocks of *T. dalaica* with *D. rerio* and *C. carpio*, and red lines represent the highlighted *BMP* gene blocks. The abbreviations "Td", "Dr" and "Cc" denote *T. dalaica*, *D. rerio* and *C. carpio*, respectively

Expression analysis

Comprehending gene expression patterns plays a vital role in revealing gene functionality. In our study, we conducted an analysis of *BMP* gene expression data from various tissues of *T. dalaica*, thus shedding light on the

expression patterns of members of the *BMP* gene family. The expression of several *BMP* genes in six tissues, *T. dalaica*, brain, liver, spleen, gonad, kidney and gill, was analyzed via real-time fluorescence quantitative PCR to verify the trend in the abundance of the *BMP* gene

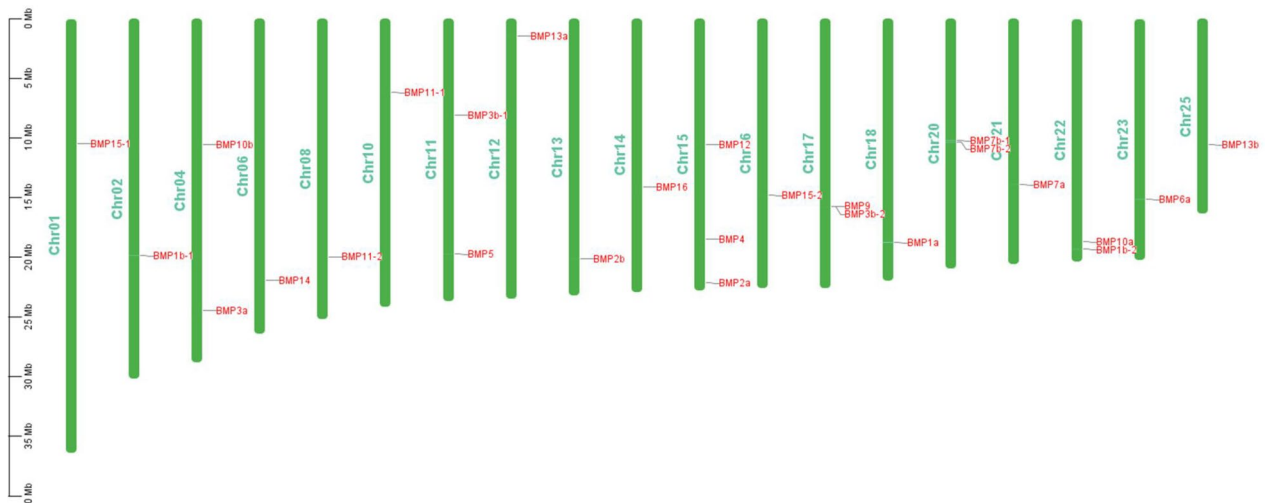


Fig. 6 Chromosomal locations of *BMP* genes in the *T. dalaica* genome

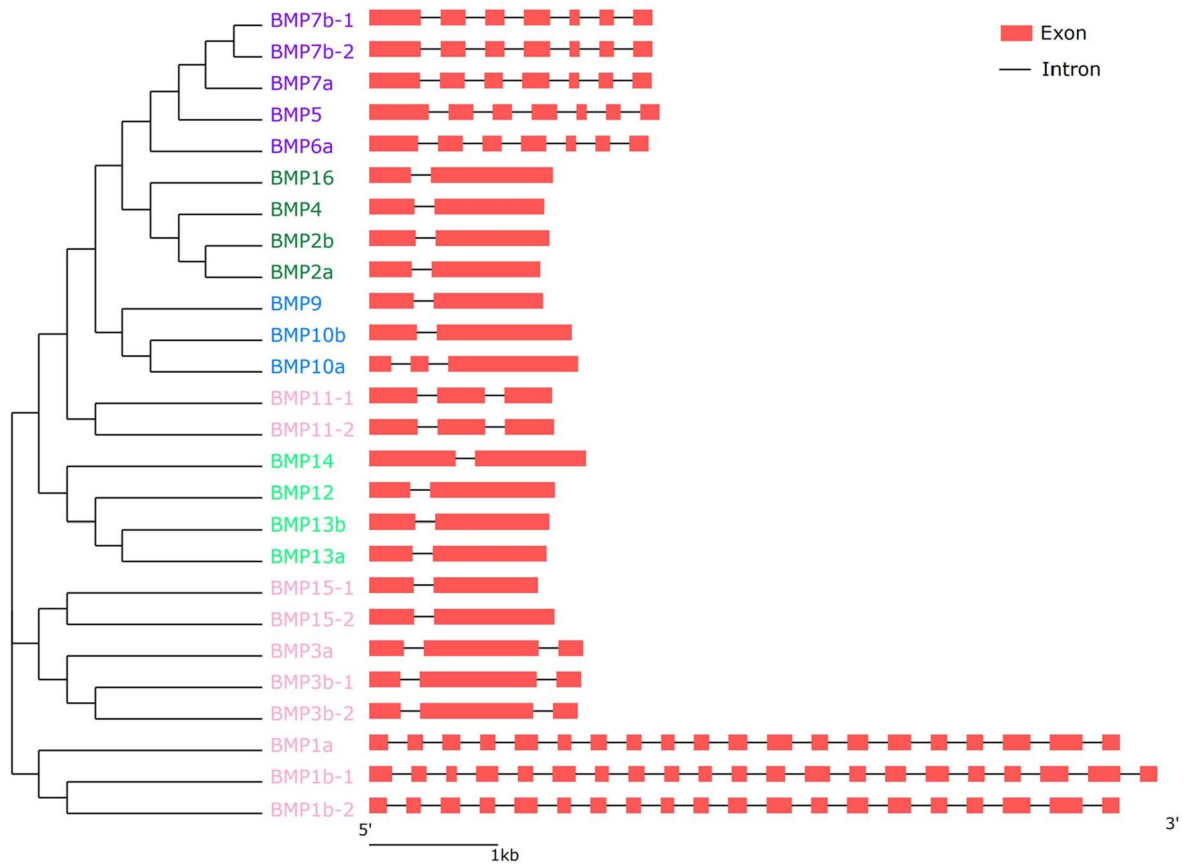


Fig. 7 Gene structure of the *BMP* gene of *T. dalaica*. The phylogenetic tree on the left was constructed in MEGA11 based on the neighbor-joining method, with a bootstrap value of 1000, and different colors represent different groups

Table 3 Secondary structure of the *BMP* gene protein from *Triplophysa dalaica*

Name	Alpha helix	Beta turn	Extended strand	Random coil
<i>BMP1a</i>	11.17%	5.84%	27.77%	55.23%
<i>BMP1b-1</i>	12.43%	5.38%	27.40%	54.79%
<i>BMP1b-2</i>	13.03%	5.33%	27.59%	54.05%
<i>BMP2a</i>	24.04%	1.28%	19.18%	55.50%
<i>BMP2b</i>	22.65%	3.13%	18.80%	55.42%
<i>BMP3a</i>	26.61%	5.32%	15.74%	52.33%
<i>BMP3b-1</i>	25.34%	2.69%	16.37%	55.61%
<i>BMP3b-2</i>	29.98%	4.58%	15.33%	50.11%
<i>BMP4</i>	20.95%	1.75%	21.45%	55.86%
<i>BMP5</i>	23.37%	2.47%	18.88%	55.28%
<i>BMP6a</i>	26.62%	3.12%	17.27%	53.00%
<i>BMP7a</i>	24.47%	4.71%	18.82%	52.00%
<i>BMP7b-1</i>	22.72%	4.22%	21.55%	51.52%
<i>BMP7b-2</i>	22.72%	4.22%	21.55%	51.52%
<i>BMP9</i>	32.41%	2.76%	18.09%	46.73%
<i>BMP10a</i>	24.89%	2.51%	18.26%	54.34%
<i>BMP10b</i>	30.66%	3.59%	14.16%	51.59%
<i>BMP11-1</i>	23.51%	2.70%	21.35%	52.43%
<i>BMP11-2</i>	21.07%	2.93%	22.67%	53.33%
<i>BMP12</i>	25.64%	3.73%	13.05%	57.58%
<i>BMP13a</i>	24.57%	1.23%	14.00%	60.20%
<i>BMP13b</i>	25.30%	2.41%	13.73%	58.55%
<i>BMP14</i>	23.77%	2.55%	14.93%	58.74%
<i>BMP15-1</i>	16.33%	2.04%	19.18%	62.45%
<i>BMP15-2</i>	27.80%	1.87%	13.32%	57.01%
<i>BMP16</i>	20.99%	2.59%	17.45%	58.96%

transcriptome. As shown in Fig. 1, the expression of *BMP1a* was greater in the liver than in the other tissues. *BMP2a* had the highest expression in the gill, *BMP3b-2* had a much greater expression in the brain than in the other tissues, and *BMP7b-1* had the highest expression in the gonad and the second highest expression in the brain. Similarly, the expression of *BMP13b* was greater in the gonad and spleen than in the other tissues. Figure 1 shows that the change in expression of the *BMP* gene in different tissues was the same as the change in expression in terms of transcriptome abundance (Fig. 4), which also indicated that the change in gene expression in the transcriptome abundance analysis in this study was correct. The heatmap revealed that the majority of the *BMP* genes in *T. dalaica* were expressed in the brain, gonads, gills, muscles, and fins, while exhibiting comparatively lower expression levels in the kidney, liver, spleen, and heart tissues. In particular, the brain presented noteworthy expression levels of *BMP11-2*, *BMP3b-1*, *BMP3b-2*, *BMP11-1*, and *BMP9*, with *BMP3b-2* demonstrating the highest expression. This discovery indicates that *BMP11s*

are likely secreted by neurons, suggesting their preference for expression in the brain [33]. In muscle tissues, *BMP* genes are universally expressed, with *BMP12* and *BMP1b-2* exhibiting the highest expression levels. In the gills, *BMP2a*, *BMP3a*, *BMP2b*, *BMP10a*, and *BMP10b* exhibited increased expression levels (Fig. 4). *BMP1a* is highly expressed in the gonads, liver, and heart; *BMP1b-1*, in muscles and gonads; and *BMP1b-2*, in muscles and gonads, suggesting that *BMP1* is broadly involved in the development of gonads and metabolism in *T. dalaica* (Fig. 4). *BMP3a*, *BMP3b-1*, and *BMP3b-2* are highly expressed in the brain, fins, gills, and muscles, suggesting the potential impact of *BMP3* on skeletal plasticity. *BMP11* is highly expressed in the brain, fins, and muscles, suggesting its potential role in the development, differentiation, and tissue formation of the nervous system in *T. dalaica*. *BMP15-1* and *BMP15-2* are highly expressed in muscles and gonads, suggesting that they play a role in reproductive system development. GO enrichment analysis revealed that the *BMP1/3/11/15* (Group I) genes primarily contributed to the growth and development of *T. dalaica*.

BMP7a, *BMP7b-1*, and *BMP7b-2* are expressed primarily in the fins, muscles, and gonads. *BMP7a* is almost not expressed in the spleen, and the expression of *BMP7b-1* and *BMP7b-2* in the spleen and liver of *T. dalaica* is also negligible. *BMP12* expression is almost nonexistent in the spleen and heart, suggesting that it plays a minor role in the physiological processes of the spleen and heart in *T. dalaica*. Similarly, *BMP13b* is hardly expressed in the liver, fins, or heart. *BMP14* was also almost not expressed in the spleen or heart. It is hypothesized that these genes do not participate in the physiological functions of the corresponding tissues in *T. dalaica* under normal physiological conditions. However, further research may be needed to determine the specific roles of these genes.

Discussion

The primary functions of *BMP* genes include the control of osteoblast and chondrocyte development and differentiation and the promotion of bone healing, all of which are crucial for bone production [32, 37, 38]. The *BMP* gene family comprises essential signaling molecules involved in regulating cell division, proliferation, function, and tissue morphogenesis [39]. Furthermore, due to their significance as growth factors, they have a wide range of potential applications in various industries, including skeletal disease therapy and regenerative medicine [40, 41].

Leveraging whole-genome sequencing data collected by our research team, we successfully identified a comprehensive set of 26 *BMP* gene family members in the *T. dalaica* genome. Within this study, we conducted an

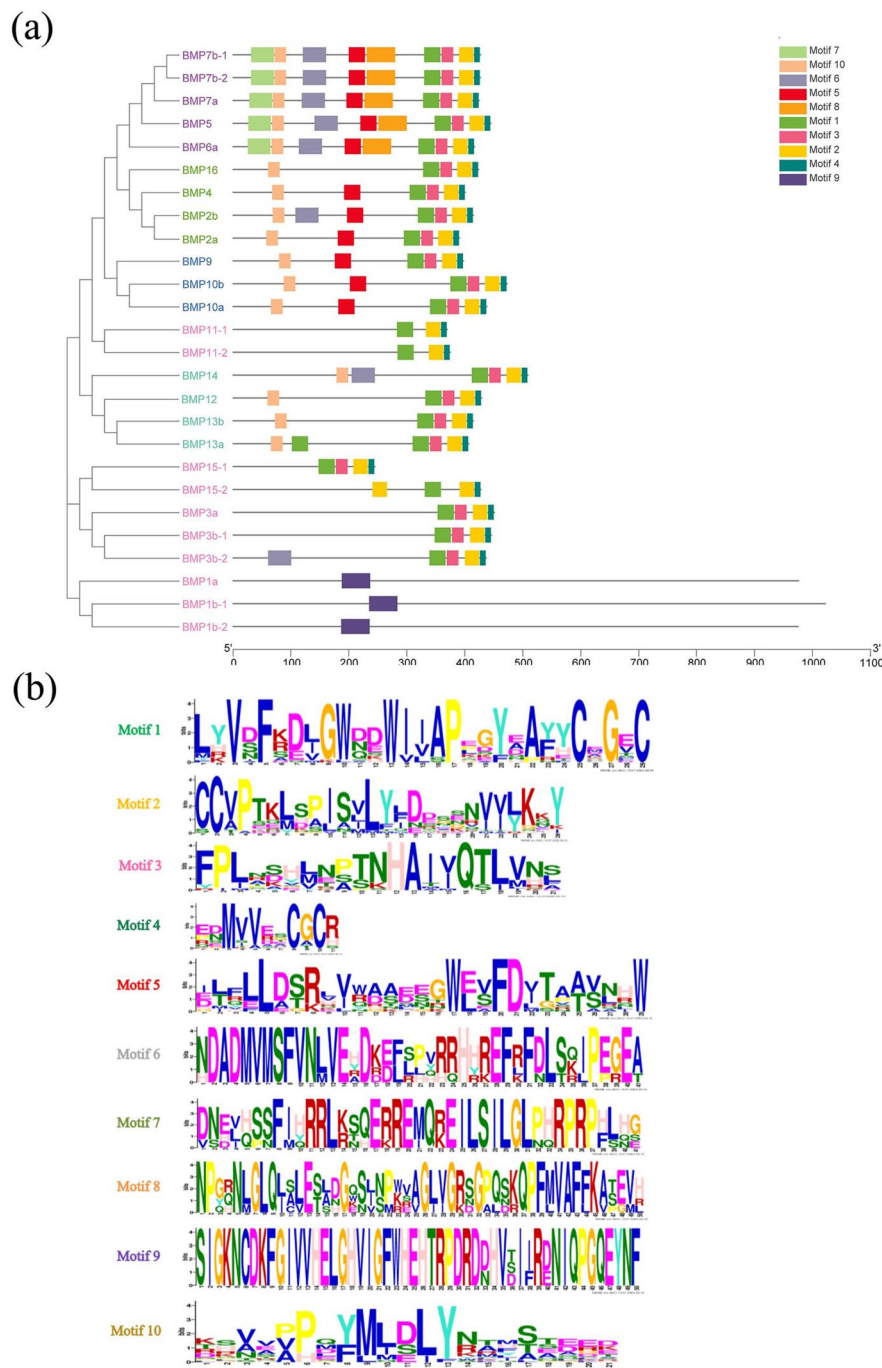


Fig. 8 Conserved motifs of the *BMP* gene of *T. dalaica*. **a** The phylogenetic tree on the left was constructed in MEGA11 based on the neighbor-joining method, with a bootstrap value of 1000, and different colors represent different groups. **b** All motifs were identified by the MEME database, and different colored blocks represent different motifs

in-depth analysis encompassing the physicochemical characteristics, subcellular localization, phylogenetic associations, intraspecies and interspecies collinearity, chromosomal mapping, and gene structure of these 26 *BMP* gene family members in *T. dalaica*, as well as

the identification of conserved motifs and domains. The classification of these *BMP* genes within the *T. dalaica* genome was determined through the evaluation of sequence homology and evolutionary relationships. We identified five different subgroups, namely, the

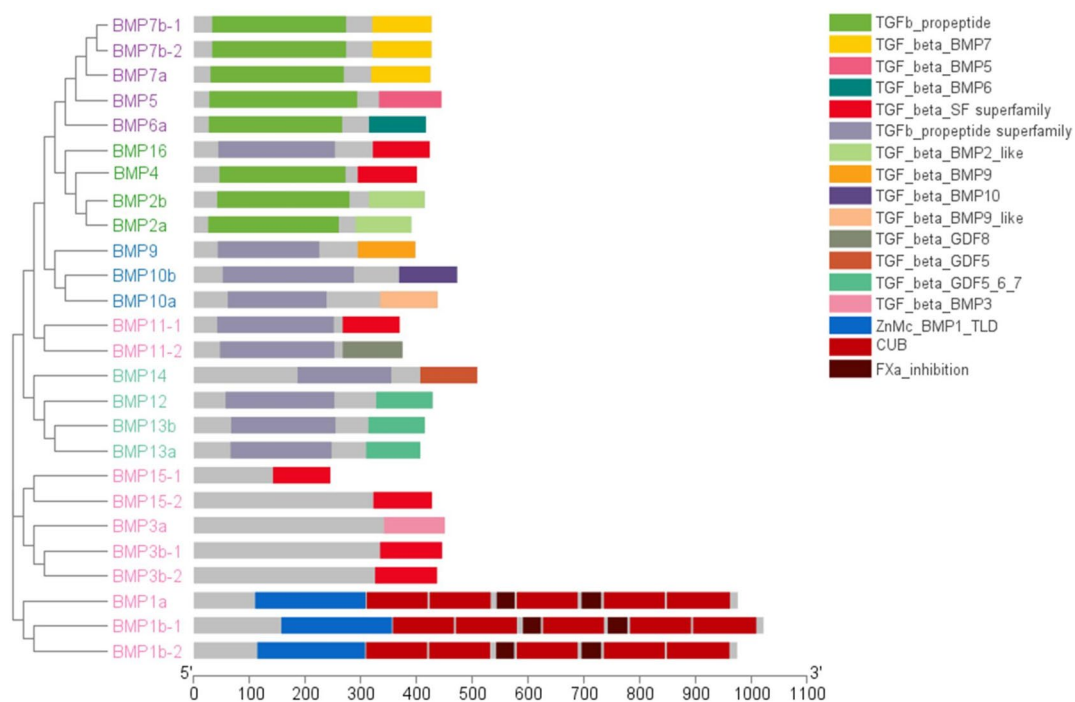


Fig. 9 Structural domains of the *BMP* gene of *T. dalaica*. The phylogenetic tree on the left was constructed in MEGA11 based on the neighbor-joining method, with a bootstrap value of 1000, and different colors represent different groups

Table 4 GO enrichment analysis of the *BMP* gene family

GO ID	GO Term	Biological Function	genes
GO:0004222	metalloendopeptidase activity	F	<i>BMP1b-1, BMP1b-2, BMP1a</i>
GO:0005509	calcium ion binding	F	<i>BMP1b-1, BMP1b-2, BMP1a</i>
GO:0006508	proteolysis	P	<i>BMP1b-1, BMP1b-2, BMP1a</i>
GO:0008237	metallopeptidase activity	F	<i>BMP1b-1, BMP1b-2, BMP1a</i>
GO:0008270	zinc ion binding	F	<i>BMP1b-2, BMP1a,</i>
GO:0008083	growth factor activity	F	<i>BMP2b, BMP2a, BMP3b-1, BMP3b-2, BMP3a, BMP4, BMP5, BMP6a, BMP7b-1, BMP7b-2, BMP7a, BMP9, BMP10b, BMP10a, BMP11-1, BMP11-2, BMP12, BMP13b, BMP13a, BMP14, BMP15-2, BMP16</i>
GO:0005102	signaling receptor binding	F	<i>BMP11-1, BMP11-2</i>
GO:0004064	arylesterase activity	F	<i>BMP13a</i>
GO:0007010	cytoskeleton organization	P	<i>BMP13a</i>
GO:0016477	cell migration	P	<i>BMP13a</i>
GO:0019904	protein domain specific binding	F	<i>BMP13a</i>
GO:0098609	cell–cell adhesion	P	<i>BMP13a</i>
GO:0005615	extracellular space	C	<i>BMP15-1</i>
GO:0001541	ovarian follicle development	P	<i>BMP15-1</i>
GO:0060016	granulosa cell development	P	<i>BMP15-1</i>

BMP 1/3/11/15 (Group I), *BMP 12/13/14* (Group II), *BMP 2/4/16* (Group III), *BM 9/10* (Group IV), and *BMP 5/6/7/8* (Group V) subgroups, which differed from the findings of previous studies of cobia [32]. These findings

align with the research outcomes documented in previous studies on carp [33]. Compared to the other species studied (*D. rerio* and *X. laevis*), *T. dalaica* has a relatively complete *BMP* gene family, which includes three *BMP1*

genes, three *BMP3* genes, three *BMP7* genes, two *BMP11* genes, and two *BMP15* genes. This difference may be the result of the scleractinian WGD event [34]. This gene duplication results in new copies of genes with the potential for differentiated functions or new functions adapted to new environments, likely due to the specific survival environment of the *T. dalaica* species, which allows these genes to be expressed in large numbers [42, 43]. This phenomenon aligns with the ecological adaptations of scleractinian fishes [44, 45]. Notably, *T. Dalaica* may have undergone a gene deletion event during evolution, as the *BMP8* gene is missing in *T. dalaica* compared to other scleractinian fishes (e.g., zebrafish, carp, etc.) and vertebrates [45].

In *T. dalaica*, the *BMP15* gene is highly expressed in muscles and gonads and plays a significant role in the development of the reproductive system [46, 47]. The *BMP11* gene is produced by neurons and is significantly expressed in the brain (Fig. 4) [33]. Most of the *BMP* genes belong to *BMP1/3/11/15* (Group I) and are likely influenced by the specific environmental conditions in which *T. dalaica* thrives [48]. The results of our phylogenetic analysis revealed that *T. dalaica*, *D. rerio*, and *C. carpio*, all of which belong to the Cypriniformes order, share a close evolutionary relationship and form a distinct cluster. This separation distinguishes them from amphibians, such as *X. laevis*, and mammals, such as *H. sapiens*. These findings align with the taxonomic criteria for species classification, indicating that *BMP* genes in *T. dalaica* have maintained a high degree of conservation throughout their evolutionary history. The conserved base sequence, analysis of structurally conserved domains, and evolutionary relationships among *T. dalaica* *BMP* genes further suggest that *BMP* genes within the same subgroup share comparable base sequences, types of structural domains, and distribution patterns [49, 50]. This observation aligns with the common amino acid sequences shared among members of the TGF- β superfamily. Notably, there was repetition and fragmentation of the genes *BMP1a* and *BMP1b-2* and of *BMP11-1* and *BMP-11-2*, which belong to the same gene group. Notably, this phenomenon was also observed for genes outside the *BMP* group, including *BMP7a*. Furthermore, a Gene Ontology (GO) enrichment analysis of the *BMP* genes revealed that, except for *BMP1a*, *BMP1b-1*, and *BMP1b-2*, the remaining *BMP* genes were associated with the GO pathway related to growth factor activity. By analyzing the expression patterns of *BMP* genes, we observed that genes within the *BMP1/3/11/15* subgroup (Group I) exhibited relatively high expression levels in most tissues. This finding suggested the significant role of these genes in the growth and development of *T. dalaica*. These findings are consistent with prior studies, such as

those conducted in cobia and carp, where *BMP15* exhibited high expression across multiple tissues and is generally regarded as crucial for female fertility [32]. *T. dalaica* lives in high-altitude and saline-alkaline environments, leading to the development of unique morphological and physiological characteristics adapted to extreme conditions. The genes *BMP1*, *BMP13*, *BMP15*, and *BMP16* are highly expressed in the gonads and likely play key roles in reproductive development and regulating sex ratios to adapt to harsh environments. *BMP2*, *BMP3*, *BMP4*, *BMP6*, *BMP10*, and *BMP16* are highly expressed in the gills, which is speculated to be a result of the environmental adaptation of *T. dalaica*, which involves the evolution of more efficient gills to obtain more oxygen.

Conclusion

In this research, we identified a total of 26 bone morphogenetic protein (*BMP*) genes within the genome of *T. dalaica*. Our phylogenetic and covariance analyses revealed that *T. dalaica* exhibited the closest genetic relationship to *D. rerio* and *C. carpio*, revealing strong covariance. We further examined the *BMP* genes in *T. dalaica* and identified instances of gene duplication, which can likely be attributed to whole-genome duplication (WGD) events. Upon scrutinizing the gene expression patterns, we noticed that the duplicated *BMP* genes in *T. dalaica* exhibited elevated expression levels. Furthermore, through extensive Gene Ontology (GO) enrichment analysis, it became evident that the *BMP1/3/11/15* genes (Group I) played a primary role in the growth and development of *T. dalaica*. This study contributes to a deeper understanding of *BMP* gene family member expression patterns in high-altitude, high-salinity environments and provides valuable insights for future research on the *BMP* gene family in bony fishes.

Abbreviations

GO	Gene Ontology
CDS	Coding sequence
F	Molecular Functions
P	Biological Processes
C	Cellular Components
WGD	Whole Genome Duplication
TPM	Transcripts Per Kilobase of exon model per Million mapped reads
NJ	Neighbor-Joining Algorithm

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-024-10049-z>.

Additional file 1: Table S1. *Triplophysa dalaica* transcriptome data used in this study.

Additional file 2: Table S2. Primers used for qRT-PCR of the *BMP* gene family.

Additional file 3: Table S3. NCBI accession numbers for the *BMP* protein sequences of the four species.

Acknowledgements

This study was supported by the High-Performance Computing Center of Henan Normal University.

Authors' contributions

Yizheng Zhang: Methodology, Formal analysis, Visualization, Writing original draft. Jinhui Yu: Methodology, Formal analysis, Visualization. Rui Han, Meng Zhang and Yikai Li: Investigation, Methodology, Formal analysis, Visualization. Zhigang Ma: Methodology, Formal analysis, Visualization. Yongtao Tang: Methodology, Formal analysis, Visualization. Guoxing Nie: Investigation, Conceptualization, Project administration, Supervision. Chuanjiang Zhou: Investigation, Methodology, Writing original draft, Conceptualization, Project administration, Funding acquisition, Supervision, Writing—review & editing. The authors read and approved the final manuscript.

Funding

This work was supported by the following funding bodies: the National Natural Science Foundation of China (U2004146, 31872199), the Henan Province Department of Science and Technology (182102110046, 222102110294), and the Training Plan of Young Excellent Teachers in Colleges and Universities of Henan Province (2019GGJ5063).

Availability of data and materials

The datasets supporting the results of this article are included within manuscript and available on request (Dr. Chuanjiang Zhou). The *Triplophysa dalaica* genome DNA sequencing data have been deposited into the NCBI Sequence Read Archive under BioProject: PRJNA624716 (<https://academic.oup.com/gbe/article/13/8/evab153/6311268>). The datasets generated and analysed during the current study are available in the NCBI GenBank: OR733217-OR733242 (Data will be released when the manuscript published).

Declarations

Ethics approval and consent to participate

The authors declare that all methods were carried out in accordance with relevant guidelines and regulations. All experiments involving animals and all experimental protocols were approved by the Animal Research and Ethics Committee of Henan Normal University and the Laboratory Animal Guidelines for the Ethical Review of Animal welfare (GB/T 35892–2018) and were conducted in accordance with the research standards, all methods were carried out in accordance with relevant guidelines and regulations, all methods are reported in accordance with ARRIVE guidelines for the reporting of animal experiments.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹College of Fisheries, Engineering Technology Research Center of Henan Province for Aquatic Animal Cultivation, Henan Normal University, Xinxiang 453007, People's Republic of China. ²College of Life Sciences, Henan Normal University, Xinxiang 453007, People's Republic of China.

Received: 27 October 2023 Accepted: 24 January 2024

Published online: 19 February 2024

References

- Urist MR. Bone: formation by autoinduction. *Science* (New York, NY). 1965;150(3698):893–9. <https://doi.org/10.1126/science.150.3698.893>.
- Wozney JM, Rosen V, Celeste AJ, Mitscock LM, Wang EA. Novel Regulators of Bone Formation: Molecular Clones and Activities. *Science* (New York, NY). 1989;242(4885):1528–34.
- Ehata S, Miyazono K. Bone Morphogenetic Protein Signaling in Cancer; Some Topics in the Recent 10 Years. *Frontiers in cell and developmental biology*. 2022;10:883523. <https://doi.org/10.3389/fcell.2022.883523>.
- Yang D, Yang X, Dai F, Wang Y, Yang Y, Hu M, Cheng Y. The Role of Bone Morphogenetic Protein 4 in Ovarian Function and Diseases. *Reproductive sciences* (Thousand Oaks, Calif). 2021;28(12):3316–30. <https://doi.org/10.1007/s43032-021-00600-8>.
- Wang RN, Green J, Wang Z, Deng Y, Qiao M, Peabody M, Zhang Q, Ye J, Yan Z, Denduluri S, et al. Bone Morphogenetic Protein (BMP) signaling in development and human diseases - ScienceDirect. *Genes Dis*. 2014;1(1):87–105.
- Cheng H, Jiang W, Phillips FM, Haydon RC, Peng Y, Zhou L, Luu HH, An N, Breyer B, Vanichakarn P, et al. Osteogenic Activity of the Fourteen Types of Human Bone Morphogenetic Proteins (BMPs). *J Bone Joint Surg Am*. 2004;85(8):1544–52. <https://doi.org/10.2106/00004623-200308000-00017>.
- Katagiri T, Watabe T. Bone Morphogenetic Proteins. *Cold Spring Harb Perspect Biol*. 2016;8(6):a021899.
- Chen L, Dong C, Kong S, Zhang J, Li X, Xu P. Genome wide identification, phylogeny, and expression of bone morphogenetic protein genes in tetraploidized common carp (*Cyprinus carpio*). *Gene*. 2017;627:157–63.
- Scott IC, Imamura Y, Pappano WN, Troedel JM, Recklies AD, Roughley PJ, Greenspan DS. Bone Morphogenetic Protein-1 Processes Probiglycan. *J Biol Chem*. 2000;275(39):30504–11.
- Yang L, Hao W, Wang H, Ren W, Yan P, Wei S. BMP2 increases hyperplasia and hypertrophy of bovine subcutaneous preadipocytes via BMP/SMAD signaling. *In Vitro Cellular & Developmental Biology-Animal*. 2022;58(3):210–9. <https://doi.org/10.1007/s11626-022-00661-2>.
- Gustafson B, Hammarstedt A, Hedjazifar S, Hoffmann JM, Svensson PA, Grimby J, Rondinone C, Smith U. BMP4 and BMP antagonists regulate human white and beige adipogenesis. *Diabetes*. 2015;64(5):1670–81.
- Addante A, González-Corrales J, Roncero C, Lazcanoiturburu N, García-Sáez J, Herrera B, Sánchez A. BMP9 Promotes an Epithelial Phenotype and a Hepatocyte-like Gene Expression Profile in Adult Hepatic Progenitor Cells. *Cells*. 2022;11(3):365.
- Zhang H, Zhang W, Bai G, Gao L, Li K. Bone Morphogenetic Protein-7 (BMP-7) Promotes Neuronal Differentiation of Bone Marrow Mesenchymal Stem Cells (BMSCs) In Vitro. *Biomed Res Int*. 2021;2021:7239783. <https://doi.org/10.1155/2021/7239783>.
- Wu FJ, Wang YW, Luo CW. Human BMP8A suppresses luteinization of rat granulosa cells via the SMAD1/5/8 pathway. *Reproduction*. 2020;159(3):315–24. <https://doi.org/10.1530/rep-19-0305>.
- Zhou C, Hu B, Tang Y, Yang C, Ma W, Wang X, Liu R, Yan X, Dong J, Wang X, et al. The chromosome-level genome of *Triplophysa dalaica* (Cypriniformes: Cobitidae) provides insights into its survival in extremely alkaline environment. *Genome Biol Evol*. 2021;13(8):evab153. <https://doi.org/10.1093/gbe/evab153>.
- Zhou C, Hu B, Tang Y, Chen X, Ma Z, Ding Q, Nie G. Genome-wide characterization of the *Triplophysa dalaica* slc4 gene family and expression profiles in response to salinity changes. *BMC Genomics*. 2022;23(1):824. <https://doi.org/10.1186/s12864-022-09057-8>.
- Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R. TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Mol Plant*. 2020;13(8):1194–202. <https://doi.org/10.1016/j.molp.2020.06.009>.
- Horton P, Park KJ, Obayashi T, Fujita N, Harada H, Adams-Collier CJ, Nakai K. WoLF PSORT: protein localization predictor. *Nucleic acids res*. 2007;35(suppl_2):W585–587. <https://doi.org/10.1093/nar/gkm259>.
- Letunic I, Khedkar S, Bork P. SMART: recent updates, new developments and status in 2020. *Nucleic Acids Res*. 2021;49(D1):D458–d460. <https://doi.org/10.1093/nar/gkaa937>.
- Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: a sequence logo generator. *Genome Res*. 2004;14(6):1188–90. <https://doi.org/10.1101/gr.849004>.
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res*. 1997;25(24):4876–82. <https://doi.org/10.1093/nar/25.24.4876>.

22. Tamura K, Stecher G, Kumar S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Mol Biol Evol.* 2021;38(7):3022–7. <https://doi.org/10.1093/molbev/msab120>.
23. Xie J, Chen Y, Cai G, Cai R, Hu Z, Wang H. Tree Visualization By One Table (tvBOT): a web application for visualizing, modifying and annotating phylogenetic trees. *Nucleic Acids Res.* 2023;51(W1):W587–92. <https://doi.org/10.1093/nar/gkad359>.
24. Geourjon C, Deléage G. SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Comput Appl Biosci.* 1995;11(6):681–4. <https://doi.org/10.1093/bioinformatics/11.6.681>.
25. Chou KC, Shen HB. Cell-PLOC 2.0: an improved package of web-servers for predicting subcellular localization of proteins in various organisms. *Nat Sci.* 2010;2(10):1090–103.
26. Ba M, Marc L, Bjoern U. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics (Oxford, England).* 2014;30(15):2114–20.
27. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods.* 2012;9(4):357–9.
28. Dc N, Bo L: RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics.* 2011;12(1):323.
29. Corchete LA, Rojas EA, Alonso-López D, De Las RJ, Gutiérrez NC, Burguillos FJ. Systematic comparison and assessment of RNA-seq procedures for gene expression quantitative analysis. *Sci Rep.* 2020;10(1):19737. <https://doi.org/10.1038/s41598-020-76881-x>.
30. Lk J, St D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods (San Diego, Calif).* 2001;25(4):402–8.
31. Lei L, Zhu J, Chen C, Wang Y, Wu C, Qi M, Wang Y, Liu X, Hong X, Yu L, et al. Genome-wide identification, evolution and expression analysis of bone morphogenetic protein (BMP) gene family in chinese soft-shell turtle (*Pelodiscus sinensis*). *Front Genet.* 2023;14:1109478. <https://doi.org/10.3389/fgene.2023.1109478>.
32. Ma Q, Yang Y, Mao F, Zhou Q, Wang L, Chen G. Genome-wide identification, phylogeny and expression analysis of the bmp gene family associated with development and skeleton deformity in cobia (*Rachycentron canadum*). *Aquaculture Reports.* 2023;31:101644. <https://doi.org/10.1016/j.aqrep.2023.101644>.
33. Chen L, Dong C, Kong S, Zhang J, Li X, Xu P. Genome wide identification, phylogeny, and expression of bone morphogenetic protein genes in tetraploidized common carp (*Cyprinus carpio*). *Gene.* 2017;627:157–63. <https://doi.org/10.1016/j.gene.2017.06.020>.
34. Xu P, Zhang X, Wang X, Li J, Liu G, Kuang Y, Xu J, Zheng X, Ren L, Wang G, et al. Genome sequence and genetic diversity of the common carp. *Cyprinus carpio* *Nat Genet.* 2014;46(11):1212–9. <https://doi.org/10.1038/ng.3098>.
35. Muir AM, Massoudi D, Nguyen N, Keene DR, Lee SJ, Birk DE, Davidson JM, Marinkovich MP, Greenspan DS. BMP1-like proteinases are essential to the structure and wound healing of skin. *Matrix biology : journal of the International Society for Matrix Biology.* 2016;56:114–31. <https://doi.org/10.1016/j.matbio.2016.06.004>.
36. Ge G, Greenspan DS. Developmental roles of the BMP1/TLD metalloproteinases. *Birth Defects Res C Embryo Today.* 2006;78(1):47–68. <https://doi.org/10.1002/bdrc.20060>.
37. Marques CL, Fernández I, Rosa J, Viegas MN, Cancela ML, Laizé V. Spatiotemporal expression and retinoic acid regulation of bone morphogenetic proteins 2, 4 and 16 in Senegalese sole. *J Appl Ichthyol.* 2014;30(4):713–20.
38. Zhang WZ, Lan T, Nie CH, Guan NN, Gao ZX. Characterization and spatiotemporal expression analysis of nine bone morphogenetic protein family genes during intermuscular bone development in blunt snout bream. *Gene.* 2018;642:116–24.
39. Yun S, Zhang X. Genome-wide identification, characterization and expression analysis of AGO, DCL, and RDR families in *Chenopodium quinoa*. *Sci Rep.* 2023;13(1):3647. <https://doi.org/10.1038/s41598-023-30827-1>.
40. Ahmad HI, Ijaz N, Afzal G, Asif AR, Ur Rehman A, Rahman A, Ahmed I, Yousaf M, Elokil A, Muhammad SA, et al. Computational Insights into the Structural and Functional Impacts of nsSNPs of Bone Morphogenetic Proteins. *Biomed Res Int.* 2022;2022:4013729. <https://doi.org/10.1155/2022/4013729>.
41. Salazar VS, Gamer LW, Rosen V. BMP signalling in skeletal development, disease and repair. *Nat Rev Endocrinol.* 2016;12(4):203–21. <https://doi.org/10.1038/nrendo.2016.12>.
42. Kong S, Dong C, Lv H, Chen L, Zhang J, Pu F, Li X, Xu P. Genome wide identification of taste receptor genes in common carp (*Cyprinus carpio*) and phylogenetic analysis in teleost. *Gene.* 2018;678:65–72. <https://doi.org/10.1016/j.gene.2018.07.078>.
43. Schoenebeck J, Hutchinson S, Mullen A, Beale H, Carrington B, Faden D, Rimbault M, Decker B, Kidd J, Sood R, et al. Variation of BMP3 Contributes to Dog Breed Skull Diversity. *PLoS Genet.* 2012;8:e1002849. <https://doi.org/10.1371/journal.pgen.1002849>.
44. Fan H, Zhou Y, Wen H, Zhang X, Zhang K, Qi X, Xu P, Li Y. Genome-wide identification and characterization of glucose transporter (glut) genes in spotted sea bass (*Lateolabrax maculatus*) and their regulated hepatic expression during short-term starvation. *Comp Biochem Physiol D: Genomics Proteomics.* 2019;30:217–29.
45. Li Y, Liu S, Qin Z, Yao J, Jiang C, Song L, Dunham R, Liu Z. The serpin superfamily in channel catfish: identification, phylogenetic analysis and expression profiling in mucosal tissues after bacterial infections. *Dev Comp Immunol.* 2015;49(2):267–77. <https://doi.org/10.1016/j.dci.2014.12.006>.
46. Galloway SM, McNatty KP, Cambridge LM, Laitinen MPE, Juengel JL, Jokiranta TS, McLaren RJ, Luiro K, Dodds KG, Montgomery GW, et al. Mutations in an oocyte-derived growth factor gene (BMP15) cause increased ovulation rate and infertility in a dosage-sensitive manner. *Nat Genet.* 2000;25(3):279–83. <https://doi.org/10.1038/77033>.
47. Wei LN, Liang XY. Abnormal expression of growth differentiation factor 9 (GDF9) and bone morphogenetic protein 15 (BMP15) in ovarian tissues from unstimulated polycystic ovaries. *Fertil Steril.* 2011;96(3):S130–1.
48. Wang Y, Yang L, Wu B, Song Z, He S. Transcriptome analysis of the plateau fish (*Triplophysa dalaica*): Implications for adaptation to hypoxia in fishes. *Gene.* 2015;565(2):211–20. <https://doi.org/10.1016/j.gene.2015.04.023>.
49. Mulloy B, Rider CC. The Bone Morphogenetic Proteins and Their Antagonists. *Vitam Horm.* 2015;99:63–90. <https://doi.org/10.1016/bs.vh.2015.06.004>.
50. Morikawa M, Derynck R, Miyazono K. TGF- β and the TGF- β Family: Context-Dependent Roles in Cell and Tissue Physiology. *Cold Spring Harb Perspect Biol.* 2016;8(5):a021873. <https://doi.org/10.1101/cshperspect.a021873>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.