

Lab Note

Evolutionary analysis of paired box gene family and biological function exploration of *Lr.Pax7* in lamprey (*Lethenteron reissneri*)

Ayqeqan Nurmamat^{1,2,†}, Zihao Yan^{1,2,†}, Yao Jiang^{1,†}, Haoran Guan^{1,2}, Ruyu Zhuang^{1,2}, Shuyuan Zhang^{1,2}, Yuesi Zhou^{1,2}, Min Xiu^{1,2}, Ya Pang^{1,2}, Ding Li^{1,2}, Liang Zhao^{1,2}, Xin Liu^{1,2,3,*}, and Yinglun Han^{1,2,3,*}

¹College of Life Sciences, Liaoning Normal University, Dalian 116081, China, ²Lamprey Research Center, Liaoning Normal University, Dalian 116081, China, and ³Collaborative Innovation Center of Seafood Deep Processing, Dalian Polytechnic University, Dalian 116034, China

[†]These authors contributed equally to this work.

*Correspondence address. Tel: +86-0411-85827098; E-mail: hanyinglun@163.com (Y.H.) / Tel: +86-18042673605; E-mail: liuxin@lnnu.edu.cn (X.L.)

Received 2 February 2024 Accepted 9 April 2024

Gene regulation refers to the precise regulation of gene expression in an organism, and transcription factors are proteins that bind to DNA and regulate gene expression by promoting or inhibiting the expressions of target genes. Since the late 1980s [1], scientists have studied special genes called *Pax* genes that control how genes function in organisms as they grow. There are nine *Pax* genes found in animals such as mice, zebrafish, and humans [2]. Based on the composition domain and homology of the sequence, the *Pax* family is divided into four subfamilies: *Pax1/9*, *Pax2/5/8*, *Pax4/6*, and *Pax3/7* [3]. *Pax7* plays a pivotal role in the implementation, protection, and repair of skeletal muscle. *Pax7* helps to control the balance between self-renewal and differentiation of satellite cells, ensuring that they can proliferate when needed to generate new muscle cells and differentiate into mature muscle fibers when necessary for muscle development and repair. The expression of the *Pax7* gene in nerve cells is critical for dorsal root and sensory ganglia development. The *Pax7* gene serves as a primary controlling factor for skeletal muscle development while influencing different biological processes; however, its exact role in jawless vertebrates such as lamprey remains unclear, and extensive research is needed to elucidate the intricate underlying mechanisms involved. Given the unique status of lamprey as an ancient jawless fish, possessing an ancient lineage and distinctive biological features, it is rare to explore gene function across hundreds of millions of years of vertebrate evolution. The use of lamprey as a model system for gene function research represents an innovative approach in the fields of evolutionary and comparative genomics. In this study, we investigated the regulatory mechanism of *Pax7* in lamprey via gene cloning, gene expression analysis, gene silencing and transcriptome data analysis. We also explored the interactions between genes with significant differences. Identification of *Lr.Pax7* in lamprey tissues

began with the retrieval of protein sequences that are similar to those of human *Pax* family members in sea lamprey (*Petromyzon marinus*) or zebrafish (*Danio rerio*) from the NCBI protein database (Supplementary Table S1) and the use of BLAST to identify corresponding homologs (Supplementary Table S2). Subsequently, we extracted the *Pax* sequences from our library. *Lethenteron reissneri* specimens were dissected to isolate various tissues. Primers targeting the *pax* domains were designed based on the *Pax7* nucleotide sequence in the Lampreys cDNA library, and the aim was to verify the effectiveness of the Lampreys cDNA as a template for validation (Supplementary Table S3). *Lr.Pax7* was successfully amplified via PCR in muscle tissue.

Here, a variety of methods were used for bioinformatics analysis. The results showed that the amino acid sequence of *Pax7* is highly similar among animals (Figure 1A), with a decreasing trend from higher to lower organisms, as revealed by sequence alignment. It can be observed from the evolutionary tree (Figure 1B) that *Pax* genes for each subfamily are present in ancestral chordate and that *Pax* genes are present in amphioxus. *Petromyzon marinus*, *Lethenteron camtschaticum*, and *Lethenteron reissneri* constitute a sister group and have become good models for the study of jawless vertebrates. *Pax9*, *Pax2*, *Pax6*, and *Pax7* show high similarity to those of other higher vertebrates. Therefore, these genes were named *Lr.Pax9*, *Lr.Pax2*, *Lr.Pax7*, and *Lr.Pax6*. The results indicate that the *Pax7* gene is significantly preserved across various species, from higher to lower. This suggests that the DNA sequence of the gene is remarkably similar among different species. *Lr.Pax7* is positioned between vertebrates and invertebrates and is most closely related to *P. marinus Pax7*. This finding provides more insight into the original evolutionary position of the lamprey. Crystal structure prediction analysis revealed that *Lr.Pax7* and *Hm*.

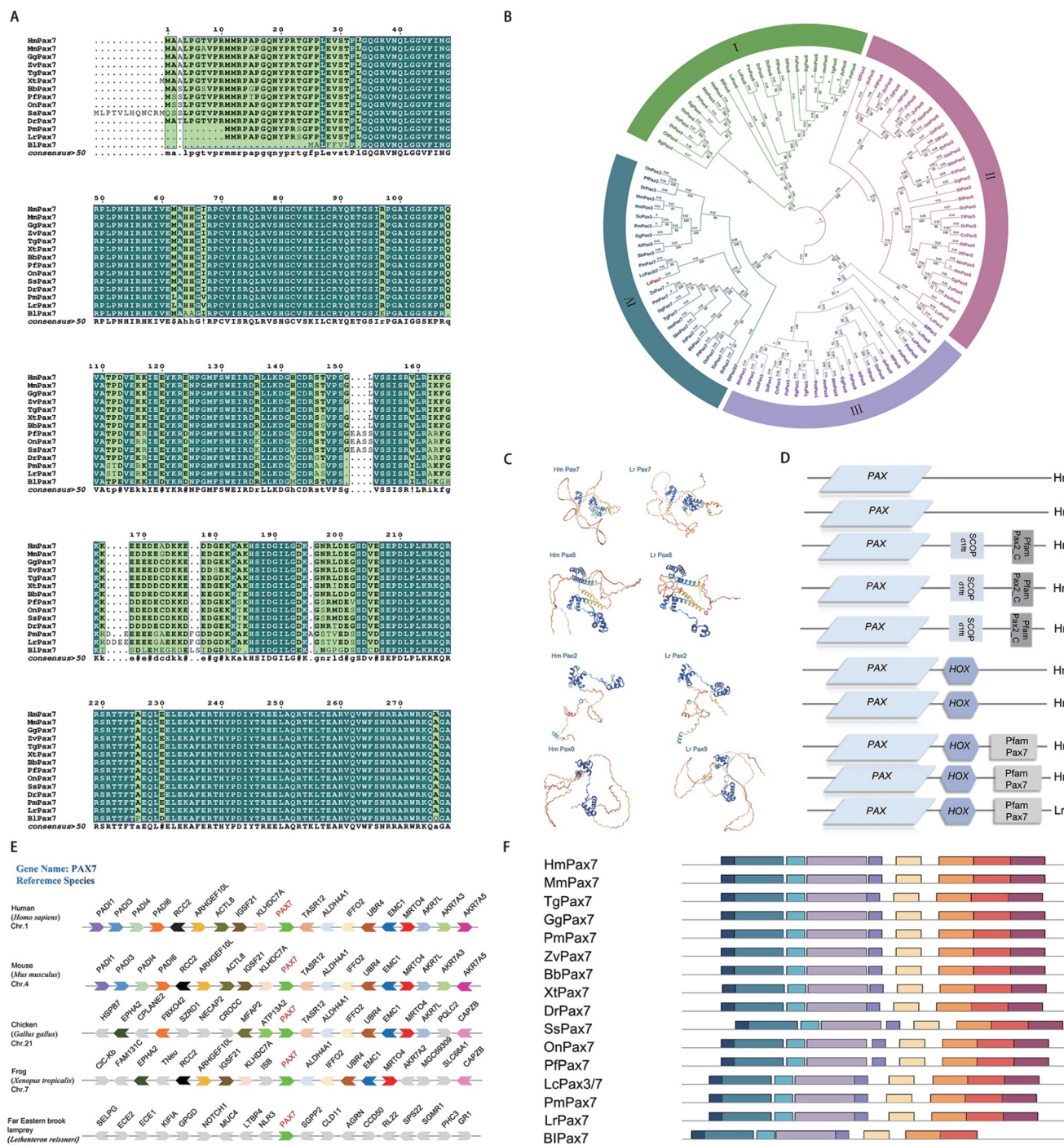


Figure 1. Bioinformatics analysis The identification and sequence analysis of Pax family genes. (A) Protein sequence alignment of the *Hm.Pax7-BI.Pax7*. (B) Phylogenetic tree for the *Pax* family based on the NJ method. (C) Prediction of the tertiary structures of *Pax7*, *Pax6*, *Pax2*, and *Pax9* in humans and lampreys. (D) Comparison of the domains between *Pax1-Pax9* from humans and lampreys. (E) Diverse genomic synteny analysis of the *Hm. Pax7* to *Lr.Pax7*. (F) Motif composition of the *Pax7* family proteins. The full motif composition of the *Pax* family and the full names of the species are listed in [Supplementary Table 1](#), and the matching amino acid sequences in the motifs are listed in [Supplementary Table 2](#).

*Pax*s have highly homologous structures (Figure 1C). The *Pax* gene has a similar structure (Figure 1D), including a conserved DNA-binding structure called the pair-box domain. This structure contains approximately 128 amino acids and is responsible for binding with specific DNA sequences, regulating gene expression,

and interacting with other proteins. To further investigate the evolutionary history of *Pax7* in vertebrates, we compared the genetic environment of *Pax7* with that of other vertebrates (Figure 1E). In addition, many *Pax* gene members also contain DNA-binding structures called homeodomains, which play important

roles in gene expression regulation and the development and establishment of various tissues and organs. *Pax* genes may also have domains that regulate transcriptional activation or inhibition, interacting with other proteins to enhance or inhibit the transcription of target genes. Some motifs were detected in all sequences, indicating that they are the most primitive and conserved functional motifs. This phenomenon also suggested that these motifs may perform similar biological functions in different species, even though they may have varied to some extent in different branches of evolution. Thus, the conservation of such motifs (Figure 1F) may mean that they are essential for specific biological functions that may be required across different species, leading to the retention of these motifs. Overall, the structure of *Pax* genes includes elements such as pairing box domains and homologous domains, which enable *Pax* genes to bind with specific DNA sequences and play important roles in embryonic development and tissue differentiation. The *Pax* gene family domain is also shown in the figure. The domain of *Lr.Pax7* is basically consistent with that of higher vertebrates, which further proves the conserved nature of the *Pax7* gene. *Lr.Pax7* serves as the ancestral form of *Pax7* in vertebrates, and our study has significant evolutionary implications.

We also used both qPCR and western blot (WB) analysis to detect the expression of a gene in various tissues to study the expression pattern and level of the gene in different tissues. The expression patterns of *Pax7* in adult *Lethenteron reissneri* were investigated via qPCR analysis of total RNA samples extracted from five normal tissues, namely, gill, kidney, spinal marrow, brain, and muscle tissues. The results revealed that *Lr.Pax7* exhibited predominant expression in muscle tissue, followed by the kidney, gill, spinal marrow and brain (Figure 2A). Western blot analysis revealed that the band density was significantly greater in the muscle sample than in the other tissues, indicating robust expression in the muscle tissue (Figure 2B). To better understanding of the biological function of *Lr.Pax7* which specifically targets the Pax domain of *Lr.Pax7*, the siRNA sequences were carefully selected to ensure efficient and specific knockdown of the *Lr.Pax7* gene. The siRNAs were administered to the lamprey specimens as part of an experimental study (Supplementary Table S4). Ten lampreys, each weighing approximately 4 g, were selected for the study. Lampreys were divided into two groups: the experimental group (siRNA) and the negative control group (NC). The results demonstrated a particularly effective silencing efficiency in muscle tissue at 72 h post-treatment (Figure 2C). Previous studies have indicated the presence of *Pax3* and *Pax7* genes in adult muscle tissues of higher vertebrates. These observations suggest that *Pax*-derived genes may retain ancestral functions. Subsequently, we selected muscle tissue for transcriptome analysis. The tissues exhibiting superior silencing efficiency (three biological replicates) were submitted to Oebiotech for transcriptome sequencing. We generated a volcano plot of the differentially expressed genes (Supplementary Figure 1A). A total of 630 genes were upregulated, while 839 genes were downregulated (Supplementary Figure 1). Then, enrichment analysis of these differentially expressed genes was conducted, and the results are as follows. Ten highly significant genes related to cancer pathways and the *PPAR* pathway were selected (Supplementary Table S5). Subsequently, qPCR experiments were performed with *GAPDH* as the reference gene to verify the authenticity of the transcriptome results. All the homogeneous transcriptome results were consistent

(Figure 2D), confirming the validity of the transcriptome data. According to the functional classification, KEGG is usually divided into three levels. The GO top 30 results showed that the changes in ribosomes and collagen in the cell components were obvious during gene interference. Significant changes were observed in oxidation-reduction, metabolic process and translation. The activity of oxidoreductases, the structural consistency of ribosomes, the composition of the extracellular structure matrix and the catalytic activity in the functional part of the molecule changed significantly (Figure 2E–G). KEGG enrichment analysis of the top 20 genes revealed significant enrichment of differential genes in *Th1* and *Th2* differentiation, *HIF-1*, and *IL-17*. Analysis of KEGG 1-3 revealed significant changes in the *Wnt*, *Notch*, and *PPAR* signaling pathways during *Pax7* silencing in lampreys. Some of the up-regulated genes have been linked to the activation of oncogenes or dysregulation of signaling pathways that lead to the onset and progression of cancer (Figure 2H–J).

The silencing of *Lr.Pax7* in lampreys led to a series of changes, affecting metabolism and immune processes, as evidenced by GO and KEGG pathway analyses. Suppression of *Lr.Pax7* influenced key metabolic enzymes, suggesting that it plays a role in vital metabolic pathways and triggers a stress response that impacts energy metabolism, hormone levels, and redox balance. Signaling molecules, such as *NFKBIA*, *TGFBF1*, *Fos*, *Bcl2*, *Bcl9*, and *Bcl10*, were activated, implicating *Lr.Pax7* in the lamprey's immune response and inflammation. Upregulated genes, especially *NFKBIA* and *TGFBF1*, indicate the modulation of the NF- κ B and TGF- β pathways. Moreover, the upregulation of *Fos* further underscores its influence on immune response genes. After observing the differentially expressed genes related to the function of *Pax7* in higher vertebrates, we detected changes in the expressions of *Pax7*-related genes (*Myog*, *Six4*, *Six1* and *MyoD*) in lamprey. *Pax7*, *Myog*, *Six4*, *Six1* and *MyoD* play pivotal roles in muscle growth and differentiation [4–6]. *Pax7* serves as a crucial regulator of muscle stem cells (satellite cells) and is indispensable for maintaining an undifferentiated myogenic precursor cell pool. Their collective actions form a regulatory network essential for muscle cell formation. This finding suggests that *Lr.Pax7* may be involved in the regulation of muscle development and differentiation in lamprey, but the function of *Pax7* in higher vertebrates may also be conserved in this ancient jawless fish. Interestingly, silencing of *Lr.Pax7* decreased *PDGFRa*, *PPAR γ* , and *Fabp4* expressions and hindered lipid droplet formation. These findings collectively suggest that *Lr.Pax7* plays a significant regulatory role in lamprey immunity, inflammation, and muscle development, paralleling its function in advanced vertebrates.

Taken together, in this study, a combination of diverse bioinformatics analyses, experimental investigations, transcriptome analyses, and other methodologies were used to further elucidate the evolution and biological functionality of *Pax7*. The phylogenetic tree revealed that the *Pax* family is classified into four distinct subfamilies, consistent with previous studies. *Lr.Pax7* is positioned in the outgroup of vertebrates within the phylogenetic tree. This observation suggests an evolutionary relationship from primitive to advanced forms in lampreys and implies that *Lr.Pax7* genes may represent ancestral *Pax7* genes in vertebrates. Furthermore, comprehensive domain similarity analysis indicated that *Lr.Pax7* likely shares similar functions with its counterparts in other vertebrate species, highlighting a high degree of conservation

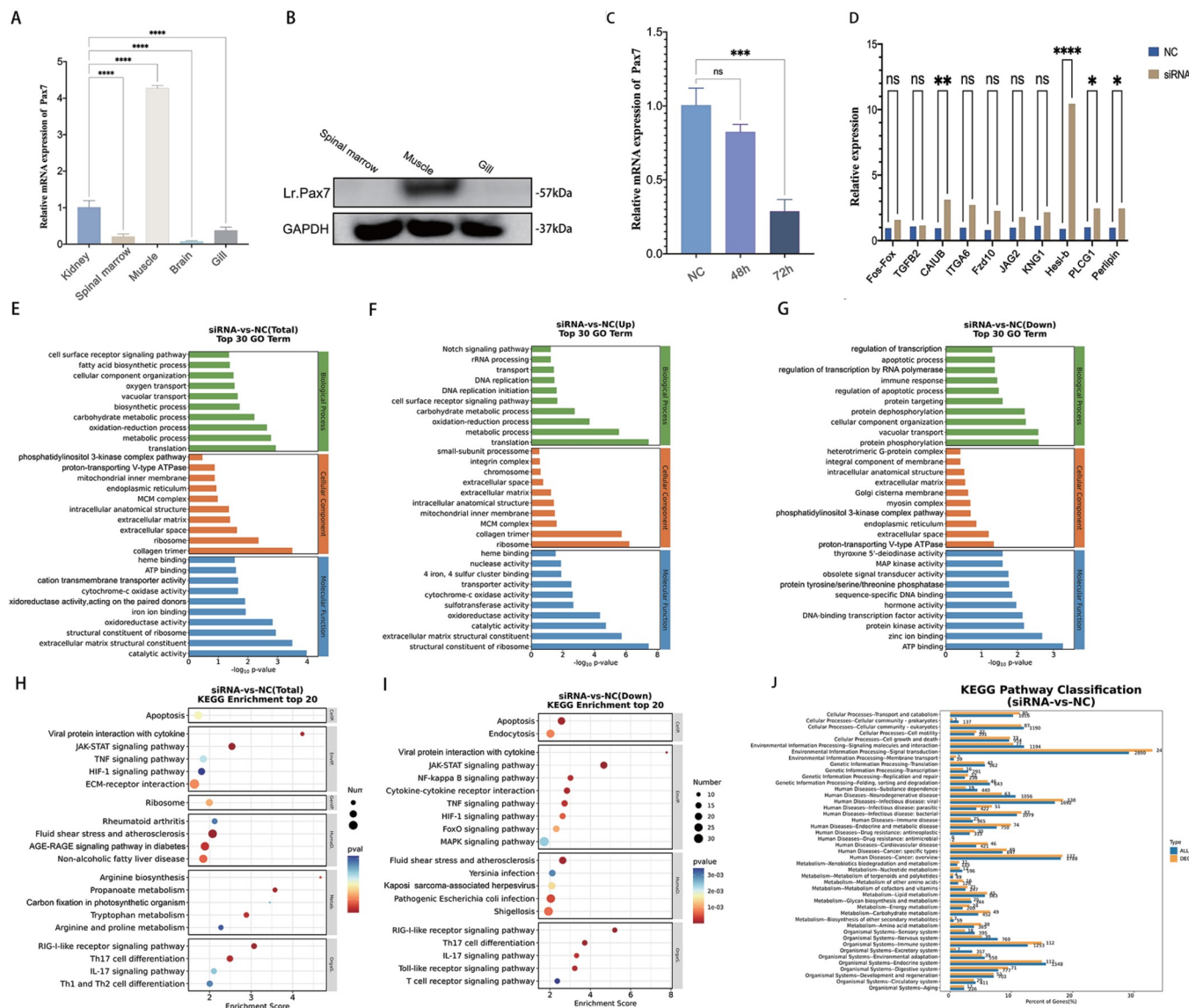


Figure 2. Expression of *Lr.Pax7* in tissues and analysis of *Lr.Pax7* gene silencing transcriptome data (A) Relative *Lr.Pax7* mRNA expression levels in various lamprey tissues. **** $P < 0.001$ (paired *t*-test). (B) Western blot analysis of the protein expression levels of *Lr.Pax7* in different tissues. (C) The expression of *Lr.Pax7* gene was quantified using qPCR following the silencing of *Lr.Pax7* gene, and an abundance map was generated. *Lr.GAPDH* served as the reference gene, and the results are expressed as the means $\pm$ SD. *Indicates significant difference compared with the control group ($n = 3$, ns: not significant, *** $P < 0.005$ (paired *t*-test). (D) Bar chart illustrating the relative expression levels of differentially expressed genes in the *Lr. Pax7* transcriptome using qPCR. Ns denotes non-significant, * $P < 0.05$, ** $P < 0.01$, and **** $P < 0.001$ (paired *t*-test). (E–G) GO enrichment analysis results for the gene silencing expression in the transcriptome. (E) Total genes regulated by *Lr.Pax7* gene silencing. (F) Up-regulated genes by *Lr.Pax7* gene silencing. (G) Down-regulated genes by *Lr.Pax7* gene silencing. (H–J) KEGG enrichment analysis results for the gene silencing expression in the transcriptome. (H) *Lr.Pax7* gene silencing: total regulated genes. (I) Down-regulated genes by *Lr.Pax7* gene silencing. (J) KEGG pathway classification.

across different organisms. Additionally, this study explored the potential conserved regulatory mechanisms underlying the evolution of *Pax7*. The conservation of molecular pathways between lamprey and humans suggests that fundamental genetic regulatory mechanisms were established early in vertebrate evolution and have remained conserved throughout the evolutionary process. This similarity provides valuable insights into the molecular and cellular processes involved in stem cell regulation, tissue regeneration, as well as underlying pathways associated with diseases. These findings suggest that *Pax7* plays a critical regulatory role in

various biological processes in lampreys, including immune-related processes and metabolism. The dysregulation of *Pax7* may contribute to the development and progression of various diseases, including cancer. Further studies are needed to fully elucidate the molecular mechanisms underlying the role of *Pax7* in these processes.

Supplementary Data

Supplementary data is available at *Acta Biochimica et Biophysica Sinica* online.

Funding

This work was supported by the grants from the National Natural Science Foundation of China (Nos. 32270557 and 31601865), the Liaoning Revitalization Talents Program (No. XLYC2007189), the High-level Talent Innovation Support Program of Dalian (No. 2023RJ012), and the Project of the Education Department of Liaoning Province (No. LJKZ0990).

Conflict of Interest

The authors declare that they have no conflict of interest.

References

1. Dambly-Chaudière C, Jamet E, Burri M, Bopp D, Basler K, Hafen E, Dumont N, *et al.* The paired box gene *pox neuro*: a determinant of poly-innervated sense organs in drosophila. *Cell* 1992, 69: 159–172
2. Dressler GR. Patterning and early cell lineage decisions in the developing kidney: the role of pax genes. *Pediatr Nephrol* 2011, 26: 1387–1394
3. Lang D, Powell SK, Plummer RS, Young KP, Ruggeri BA. PAX genes: roles in development, pathophysiology, and cancer. *Biochem Pharmacol* 2007, 73: 1–14
4. Olguín HC, Pisconti A. Marking the tempo for myogenesis: Pax7 and the regulation of muscle stem cell fate decisions. *J Cell Mol Med* 2012, 16: 1013–1025
5. Stewart CE, Rittweger J. Adaptive processes in skeletal muscle: molecular regulators and genetic influences. *J Musculoskelet Neuronal Interact* 2006, 6: 73–86
6. Kielbówna L, Jędrzejowska I. How is myogenesis initiated in chordates? *Folia Biol (Krakow)* 2012, 60: 107–119