

Expansion of the Actin Gene Family in Amphioxus

WEI Yun-hu^{1,2}, ZHANG Yu-jun^{1,2}, CHEN Yuan^{1,2}, MAO Bing-yu^{1,*}

(1. State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, the Chinese Academy of Sciences, Kunming 650223, China; 2. Graduate University of the Chinese Academy of Sciences, Beijing 100049, China)

Abstract: Actins are a small family of ubiquitous proteins that are essential cytoskeletal components and are highly conserved during evolution. Actins are usually divided into two classes, the cytoplasmic and muscle actins, which have different functional roles. Here we systematically analyzed the actin genes in the genome of the primitive chordate amphioxus (*Branchiostoma floridae*). We found that amphioxus contains more than 30 actin genes, many of which are linked. Phylogenetic analysis suggests the amphioxus actin genes have clearly undergone extensive expansion through tandem duplications. The actin genes' structure also varies a lot, containing 2 to 7 exons. We also cloned two muscle type of actin genes from the amphioxus (*B. belcheri*) and compared their expression patterns during early development. The slight difference in their expression suggests functional diversification of these actin genes. Our results shed light on the evolution both of actin genes themselves and their functional roles in chordate development.

Key words: Amphioxus; Actin; Gene family; Evolution; Expression pattern

文昌鱼肌动蛋白基因家族的扩增

魏云虎^{1,2}, 张煜珺^{1,2}, 陈源^{1,2}, 毛炳宇^{1,*}

(1. 中国科学院昆明动物研究所 遗传资源与进化国家重点实验室, 云南 昆明 650223; 2. 中国科学院研究生院, 北京 100049)

摘要: 肌动蛋白是一种分布广泛而且在进化上十分保守的蛋白, 是构成细胞骨架的关键组分。肌动蛋白通常被分成肌肉型和胞质型两种类型, 各自行使着不同的功能。在此, 作者对佛罗里达文昌鱼基因组中的肌动蛋白基因家族进行了系统分析, 发现文昌鱼中该基因家族成员多达 30 多个, 其中很多都是连锁分布的。基因结构趋于多样, 分别包含 2~7 个外显子。进化分析的结果显示, 文昌鱼的肌动蛋白基因家族可能通过串联重复而发生了扩增。作者还克隆了厦门文昌鱼两个不同的肌肉型肌动蛋白基因, 并比较了它们在文昌鱼早期胚胎中的表达图式。结果显示, 这两个基因在表达上有着细微的差别, 提示文昌鱼肌动蛋白基因家族成员在功能上的分化。上述结果将有助于阐明肌动蛋白基因家族及其功能在脊索动物中的演化。

关键词: 文昌鱼; 肌动蛋白; 基因家族; 进化; 表达图式

中图分类号: Q959.287; Q245; Q344; Q593.4; Q786 **文献标志码:** A **文章编号:** 0254-5853-(2009)05-0473-07

Actins are a small family of ubiquitous proteins that are essential cytoskeletal components and are highly conserved during evolution. Actins are usually divided into two classes, the cytoplasmic and muscle actins, which have different functional roles (Pollard & Cooper, 1986). Vertebrates have at least six actin isoforms

(Vandekerckhove & Weber 1984; Miwa et al, 1991), two cytoplasmic actins (β -cytoplasmic, γ -cytoplasmic) and four muscle actins (α -skeletal, α -cardiac, α -vascular, and γ -enteric). The cytoplasmic actin is widely expressed in many tissues while the muscle actins are specifically expressed in different types of muscles. In chordates, the

Received date: 2009-05-31; Accepted date: 2009-06-10

Foundation items: This work was supported by grants from the National Natural Science Foundation of China (30425011; 30530380); the Innovation Project of Chinese Academy of Sciences (KSCX2-YW-R-090)

收稿日期: 2009-05-31; 接受日期: 2009-06-10

基金项目: 国家杰出青年科学基金 (30425011); 国家自然科学基金重点项目 (30530380); 中国科学院知识创新工程重要方向项目 (KSCX2-YW-R-090)

*通讯作者(Corresponding author), 博士生导师, E-mail: mao@mail.kiz.ac.cn

cytoplasmic and muscle actin are different at about 20 amino acid positions, which serve to distinguish the two types of actins (Vandekerckhove & Weber 1979, 1984; Kovilur et al, 1993). Also, the cytoplasmic and muscle actin genes have different intron-exon structures in mammals.

The actin genes in the primitive chordates, urochordates and cephalochordates, have been subjected to several studies (Bovenshulte & Weber, 1997; Kusakabe 1995, 1997; Kusakabe et al, 1997a, b, 1999; Suzuki & Satoh, 2000). The amphioxus, which is thought to be the closest living relative of the vertebrates, has been a model system to study the evolution and mechanisms of development in vertebrates (Holland et al, 2004). The Florida amphioxus (*Branchiostoma floridae*) cytoplasmic and muscle actins are clearly distinguishable and closely related to vertebrate cytoplasmic and muscle actins, respectively (Kusakabe et al, 1997b). The cytoplasmic actin gene BfCA1 is widely expressed while the muscle actin BfMA1 is expressed strongly in somatic muscle and weakly in smooth muscle associated with gill slits. Also, southern blot analysis of genomic DNA has suggested that the amphioxus genome contains multiple muscle and cytoplasmic actin genes that might have arisen from recent duplication and gene conversion. Recent results indicate that vertebrate actin isoforms evolved after the separation of the vertebrates and urochordates (Kusakabe, 1995; Kusakabe et al, 1997a). It has also been suggested that the actin genes have arose independently in each of the three chordate lineages.

With the nearly complete genome sequencing of the Florida amphioxus (*B. floridae*, Putnam et al, 2008), we can systematically study the actin genes in this primitive chordate species. We found that the amphioxus contains more than 30 actin genes, many of which are linked. Phylogenetic analysis suggests that the amphioxus actin genes have clearly undergone expansion through tandem and clusteral duplications. The comparative analyses of the actin genes' structure suggest a model of the evolution of these genes. We also cloned two muscle type of actin genes from the amphioxus and compared their expression patterns during early development. The slight difference in their expression suggests functional

diversification of these actin genes. Our results shed light on the evolution both of actin genes themselves and their functional roles in chordate development.

1 Materials and Methods

1.1 Blast searches and phylogenetic analysis

The coding region of *Branchiostoma floridae* muscle actin gene (BfMA1, GenBank accession number D87407) was used to blast the *B. floridae* genome (<http://genome.jgi-psf.org/Braf11>). The corresponding regions were extracted and the gene structures were analyzed using Genomescan (<http://genes.mit.edu/genomescan.html>) and manual corrections. A phylogenetic tree was constructed by the neighbor-joining method using the MEGA4 program (Zhao et al, 2007).

1.2 Cloning of *Branchiostoma belcheri* muscle actin genes and in situ hybridization

Adult amphioxus (*B. belcheri*) were collected during the breeding season from the South China Sea and maintained in the laboratory. Naturally fertilized eggs were collected and cultured at room temperature. The 3' partial sequences plus the 3' untranslated regions of two *B. belcheri* muscle type actin genes were cloned using 3' RACE. One corresponds to the published BfMA1 (Kusakabe et al, 1999), and a second one corresponds to a new muscle actin gene (corresponding to the 23-9 gene, see below) whose corresponding genomic sequence is available in *B. belcheri*. The gene specific primers used were: 5'-ACATCCGTAAGGACCTGTACGCCAAC-3' (BfMA1) and 5'-GATCGTCAGCAGAAGGAGATCA CCTT-3' (23-9). Total mRNA of 36-hour larvae was used as template for the reverse transcription. The PCR products were then cloned into pGEM-T vector, sequenced and used to prepare probed for *in situ* hybridization. *In situ* hybridizations were carried out as described (Zhang & Mao, 2009).

2 Results

2.1 The actin genes in amphioxus genome

Taking advantage of the nearly completed sequenced genome of the Florida amphioxus (*Branchiostoma floridae*), we searched systematically for

actin homologous genes. We made blast searches using the published *B. floridae* muscle actin (Kusakabe et al, 1997b) coding region as the query. In total 11 scaffolds were found to contain actin homologous sequences. The corresponding regions were extracted and analyzed by Genomescan and manual corrections. Surprisingly, 37 genes were predicted, 27 of which were full length and the rest contained gaps in the sequences (Fig.1, Tab. 1). Most of them (31 out of 37) are organized into 5 gene clusters (Fig.1). The genes are referred to according to their locations on the scaffolds (Fig.1, Tab. 1). We also found that scaffold 264 (16kb) and partial sequence of scaffold 226 (8 kb, containing genes 226-1 to 226-3) show a high similarity with sequences in scaffold 231. We can not rule out the possibility that they might represent different alleles of the same cluster, or they can result from recent cluster duplications. Taking into account also the possibility of miss-assembly of the sequences, the confirmation of the exact number of actin genes in amphioxus still awaits the improvement of the

sequence data. Phylogenetic analysis (Fig.2) of the 27 predicted full length actin proteins showed that they fell into three different groups: the muscle type, the cytoplasmic type and those that showed similar homology to both muscle and cytoplasmic types. In addition to the reported cytoplasmic actin (Kusakabe et al, 1997b, corresponding to gene 68), a cluster of 4 genes (on scaffold 207) was also clustered into the cytoplasmic type. They likely represent a new type of cytoplasmic actin.

2.2 The structure and evolution of the actin genes

The amphioxus muscle and cytoplasmic actin genes have been reported to contain 4 and 5 exons respectively. Indeed, in the genes with confirmed exon-intron structures, 14/31 contain 5 exons, 9/31 contain 4 exons. Interestingly, genes with different structures also exist. 5/31 contain 6 exons, two of them have only two exons, and one gene even contains 7 exons (Tab. 2). Most of genes locate in the same scaffold have the same structure (Fig.1, Tab. 2), strongly suggesting that tandem

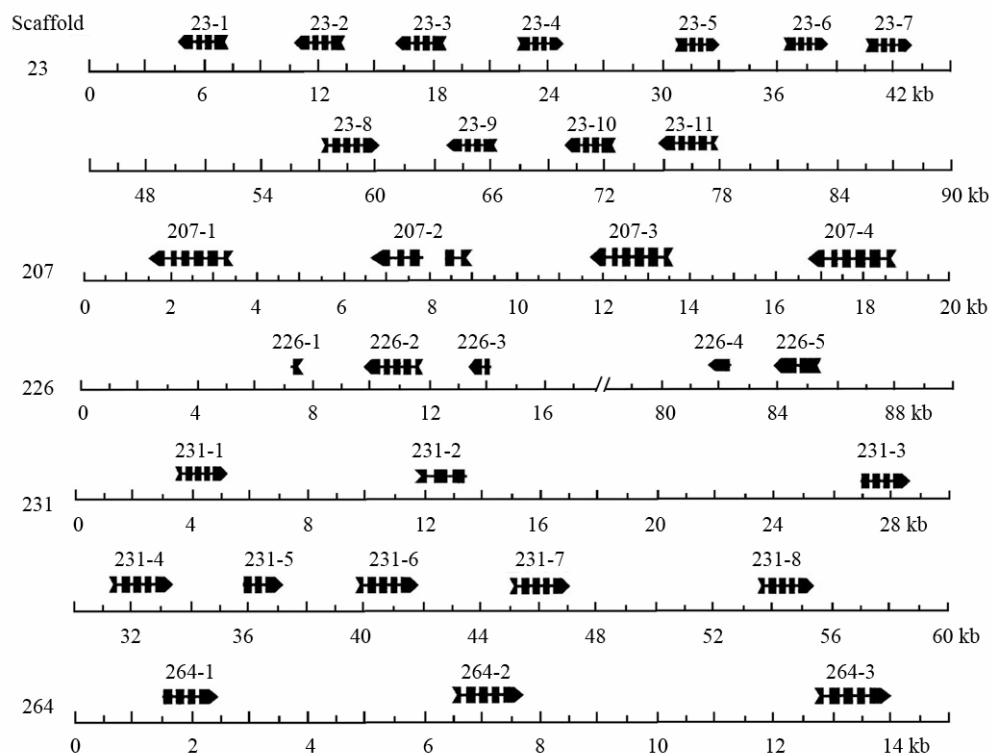


Fig. 1 The actin gene clusters in the genome of *Branchiostoma floridae*

The genes are referred to according to their locations on the scaffolds. The scaffolds containing single actin genes (25, 68, 107, 121, 164 and 227) were not shown here. The directions and gene structures were illustrated. The scaffold numbers are according to the second version of the genomic sequence of *B. floridae* (<http://genome.jgi-psf.org/Braf11>).

Tab. 1 The predicted actin genes in the *Branchiostoma floridae* genome

Gene name	Hypothetical protein No. BRAFLDRAFT_	Gene ID	Exons	Scaffold (Bf_V2_)	Note	
23-1	91591	7240061	4	23	Identical with 23-2 at protein level.	
23-2	91591	7240061	4			
23-3	91592	7219456	4			
23-4	91593	7213556	4			
23-5	91594	7219447	4			
23-6	127488	7247450	4			
23-7	115566	7247449	4			
23-8	91599	7213557	5			Gap
23-9	115345	7240054	4			
23-10	126328	7213558	4			
23-11	126329	7219458	5			
25	117851	7214412	2	25		
68	56687	7213170	5	68		
107	116948	7207692	6	107		
121	114976	7242166	6	121		
164	58093	7235147	5	164	Missing first 19aa ¹	
207-1	114535	7218086	6	207		
207-2			6?		Gap	
207-3	104919	7218293	6			
207-4	104920	7219752	6		1 amino acid difference	
226-1			5?	226	Gaps	
226-2	86184	7243434	5			
226-3			?		Gaps	
226-4			2?		Gaps	
226-5	117851	7214412	2			
227	201011	7216746	7	227	13 additional aa ¹	
231-1	115291	7219299	5	231		
231-2			5?		Gaps	
231-3			5		Gaps	
231-4	275467	7213337	5			
231-5			5?		Gaps	
231-6	126063	7211552	5		Fusion ²	
231-7	126063	7211552	5		Fusion ²	
231-8			5		Gaps	
264-1			5	264	Gaps	
264-2	88990	7246371	5			
264-3	125401	7246689	5			

Annotated hypothetical proteins and genes in the GenBank identical or close to our predicted genes were listed.

1, The hypothetical proteins in the database contain less or more amino acids than our predicted version.

2, The hypothetical protein 126063 is a fusion version of 231-6 and -7.

duplication was the major means of expansion of the actin gene family in amphioxus. Since most of the intron positions (42, 123, 206, 270, Tab. 2) are conserved across species (Kusakabe et al, 1999), even in invertebrate actin genes, we prefer the model that the ancestor actin gene to be multi-exon. The actin genes contain only two exons might arise due to intron loss during evolution. Since intron gain and loss happen heavily during evolution, it is not easy to reconstruct the evolutionary history of the amphioxus actin genes according only to their structures.

2.3 Expression of two muscle type actin genes in

amphioxus (*B. belcheri*)

In vertebrates, the muscle type actin genes were differentially expressed muscle types, e.g., the skeletal, smooth and cardiac muscles. Since multiple muscle actin genes were present in amphioxus, we cloned partial sequences of two different muscle type actin genes and compared their expression during amphioxus early development. The two actin genes showed similar expression in the somatic muscles in neurula and early larvae of amphioxus (Fig.3: A–C, F–H), but their expression showed slightly differences at later stages. The expression of 23-1 was much stronger than that of

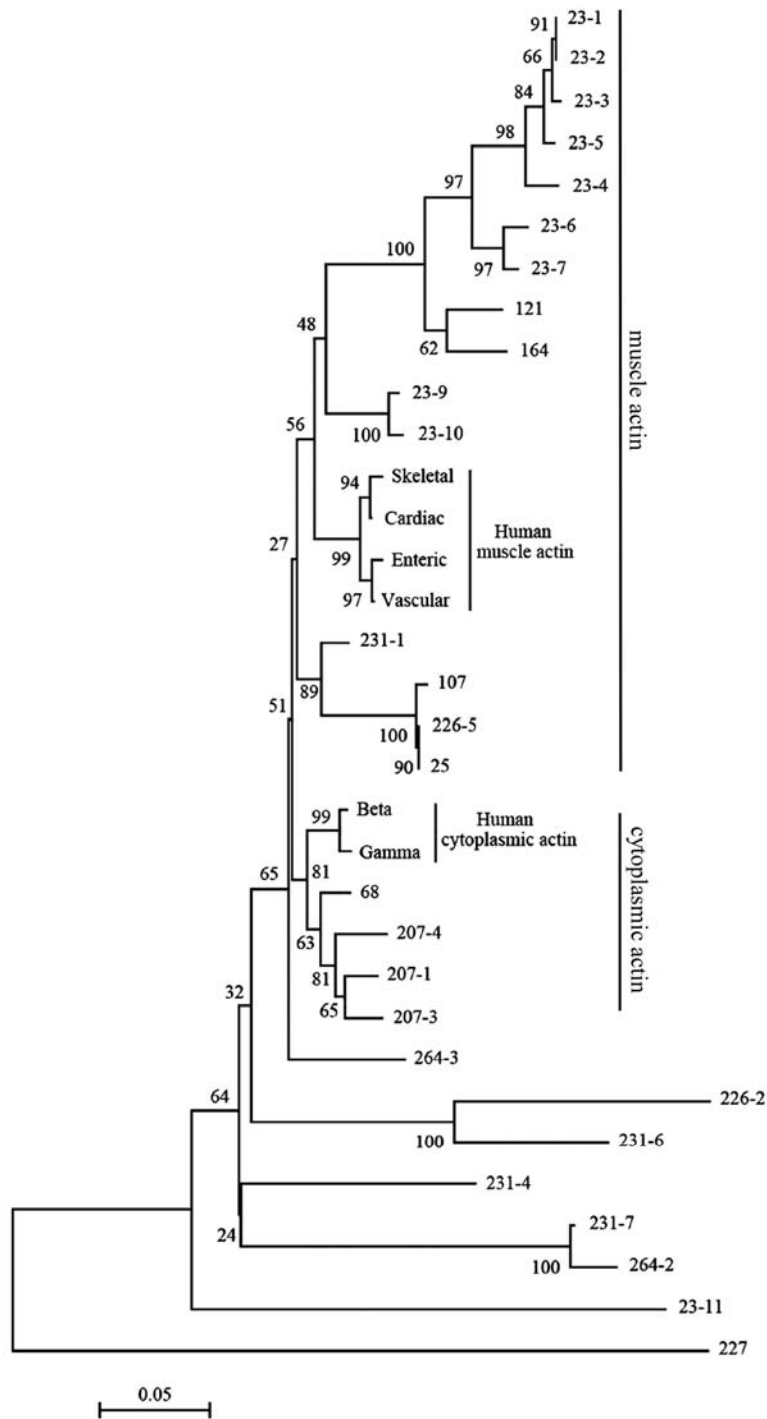


Fig. 2 Phylogenetic analysis of the amphioxus and human actin proteins

The phylogenetic tree was constructed by the neighbor-joining method using the MEGA4 program. The predicted *Branchiostoma floridae* actins were referred to according to their locations on the scaffolds (Fig.1 and Tab.1). The muscle and cytoplasmic clusters of actin genes are indicated. The rest genes do not show a closer relationship to either cytoplasmic or muscle type actins. The accession numbers of the human actin protein sequences used were: β -cytoplasmic: M10277; γ -cytoplasmic: NP_001605; α -skeletal: M20543; α -cardiac: J00073; α -vascular: BAA00492; γ -enteric: X26940.

Tab. 2 Comparison of intron positions in amphioxus actin genes

Genes	Intron position									
	42	64	83	114	123	154	169	206	270	309
23-1-7, -9, -10					+			+	+	
23-8, -11; 68; 164; 226-2; 231-1-8; 264-1-3	+				+			+	+	
207-1-4	+				+			+	+	+
25; 226-5								+		
107		+		+			+	+	+	
121		+		+		+		+	+	
227	+		+		+			+	+	+

Nomenclature for intron positions is according to the number of the corresponding amino acids in the protein (Bovenshulte & Weber, 1997). Note that due to slight length variations, the exact position of each intron in different actins might vary slightly from listed here.

23-9 at larval stages. At early larva stage (24 h), the expression of 23-1 became weak in the middle region of the body (Fig.3D) and the expression of 23-9 was only clear in the anterior part of the body (Fig.3I). At the 45-hour larva stage, the expression of 23-1 was still strong in the muscle in the anterior trunk and the tail bud (Fig.3E) while the expression of 23-9 became very weak and could be detected only weakly in the pharyngeal region (Fig.3J).

3 Discussion

Actins are very well conserved at protein level during evolution. Here we showed that the amphioxus actin genes have clearly undergone rapid lineage specific expansion and contain more than 30 actin genes. Some of them fall clearly into the muscle or cytoplasmic types,

while the rest show similar homology to both muscle and cytoplasmic actins. We also extracted the 1kb sequences upstream of the start codon and 300 bp 3' untranslated region (UTR) sequences of the actin genes and tried to construct phylogenetic trees using these sequences. However, the result showed that the 3' UTR sequences have become quite divergent and the tree was no longer reliable (data not shown). The 5' 1kb regions contain the 5' UTR sequences and the potential promoter regions. Comparisons of these sequences revealed partial conservation of some of the sequences (data not shown). Detailed analysis of the potential promoter regions of these genes might help to resolve their transcriptional regulation mechanisms. It will be of interest to check systematically the expression patterns of the different actin genes to follow their potential functional

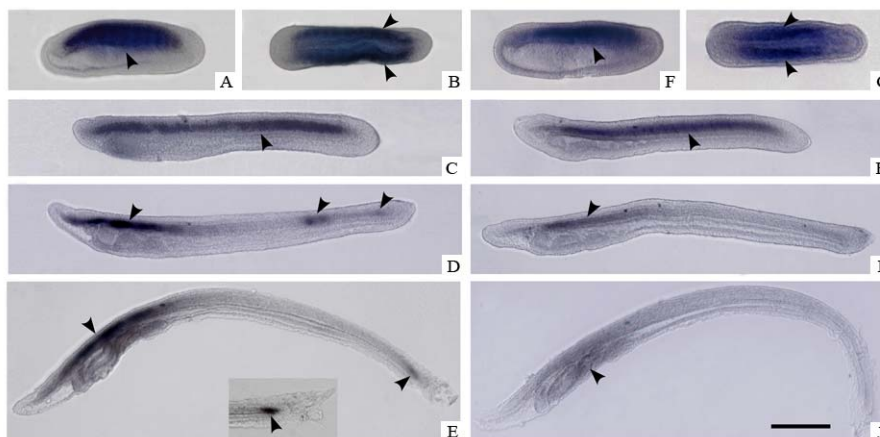


Fig. 3 Embryonic and larval expression of the muscle actin genes 23-1 (A-E) and 23-9 (F-J) in *Branchiostoma belcheri* (A,B) Mid neurula (14h) with strong expression in the somatic muscle (arrowheads). (C,H) Late neurula (18h) with strong expression in the somatic muscle (arrowheads). (D, I) Early larva (24h) with expression in the muscle (arrowheads). Note that the expression of 23-1 became weak in the middle region of the body (D) and the expression of 23-9 was only clear in the anterior part of the body (I). (E, J), 45-hour larva. At this stage, the expression of 23-1 was still strong in the muscle in the anterior trunk and the tail bud (arrowheads in E and insert) while the expression of 23-9 became very weak and could be detected weakly in the pharyngeal region (arrowhead in J). (B), (G), dorsal views, anterior to the left. All other panels are lateral views, anterior to the left. Scale bar, 100 μ m.

divergences. In sea urchins, the actin gene family has been also reported to be rapidly evolving whose members are expressed in distinct cell lineages in a developmentally regulated fashion (Kissinger et al, 1997). Since they are usually clustered, the amphioxus actin gene family is also a good system to study gene evolution and the evolutionary relationships between available amphioxus species. It will be of interest to

check the gene number, site of expression, and gene linkages of the actin gene family in different amphioxus species.

Acknowledgements: We thank Profs. CHEN Jun-yuan (Nanjing University) and ZHANG Hong-wei (Shandong University) for help with the amphioxus materials.

References:

- Bovenschulte M, Weber K. 1997. Deuterostomic actin genes and the definition of the chordates: cDNA cloning and gene organization for cephalochordates and hemichordates[J]. *J Mol Evol*, **45**(6): 653-660.
- Holland LZ, Laudet V, Schubert M. 2004. The chordate amphioxus: an emerging model organism for developmental biology[J]. *Cell Mol Life Sci*, **61**(18): 2290-2308.
- Kissinger JC, Hahn JH, Raff RA. 1997. Rapid evolution in a conserved gene family. Evolution of the actin gene family in the sea urchin genus *Heliocidaris* and related genera[J]. *Mol Biol Evol*, **14**(6): 654-665.
- Kovilur S, Jacobson JW, Beach RL, Jeffery WR, Tomlinson CR. 1993. Evolution of the chordate muscle actin gene[J]. *J Mol Evol*, **36** (4): 361-368.
- Kusakabe R, Satoh N, Holland LZ, Kusakabe T. 1999. Genomic organization and evolution of actin genes in the amphioxus *Branchiostoma belcheri* and *Branchiostoma floridae*[J]. *Gene*, **227** (1): 1-10.
- Kusakabe R, Kusakabe T, Satoh N, Holland ND, Holland LZ. 1997a. Differential gene expression and intracellular mRNA localization of amphioxus actin isoforms throughout development: implications for conserved mechanisms of chordate development[J]. *Dev Genes Evol*, **207**(4): 203-215.
- Kusakabe T, Araki I, Satoh N, Jeffery WR. 1997b. Evolution of chordate actin genes: evidence from genomic organization and amino acid sequences[J]. *J Mol Evol*, **44**(3): 289-298.
- Kusakabe T, Makabe KW, Satoh N. 1992. Tunicate muscle actin genes: structure and organization as a gene cluster[J]. *J Mol Biol*, **227**(3): 955-960.
- Kusakabe T. 1995. Expression of larval-type muscle actin encoding genes in the ascidian *Halocynthia roretzi*[J]. *Gene*, **153**(2): 215-218.
- Kusakabe T. 1997. Ascidian actin genes: developmental regulation of gene expression and molecular evolution[J]. *Zool Sci*, **14**(5): 707-718.
- Miwa T, Manabe Y, Kurokawa K, Kamada S, Kanda N, Bruns G, Ueyama H, Kakunaga T. 1991. Structure, chromosome location, and expression of the human smooth muscle (enteric type) -actin gene: evolution of six human actin genes[J]. *Mol Cell Biol*, **11** (6): 3296-3306.
- Pollard, TD, Cooper, JA. 1986. Actin and actin-binding proteins. A critical evaluation of mechanisms and functions[J]. *Annu Rev Biochem*, **55**: 987-1035.
- Putnam NH, Butts T, Ferrier DE, Furlong RF, Hellsten U, Kawashima T, Robinson-Rechavi M, Shoguchi E, Terry A, Yu JK, Benito-Gutiérrez EL, Dubchak I, Garcia-Fernández J, Gibson-Brown JJ, Grigoriev IV, Horton AC, de Jong PJ, Jurka J, Kapitonov VV, Kohara Y, Kuroki Y, Lindquist E, Lucas S, Osoegawa K, Pennacchio LA, Salamov AA, Satou Y, Sauka-Spengler T, Schmutz J, Shin-I T, Toyoda A, Bronner-Fraser M, Fujiyama A, Holland LZ, Holland PW, Satoh N, Rokhsar DS. 2008. The amphioxus genome and the evolution of the chordate karyotype[J]. *Nature*, **453**(7198): 1064-1071.
- Suzuki MM, Satoh N. 2000. Genes expressed in the amphioxus notochord revealed by EST analysis[J]. *Dev Biol*, **224**(2): 168-177.
- Vandekerckhove J, Weber K. 1984. Chordate muscle actins differ distinctly from invertebrate muscle actins: the evolution of the different vertebrate muscle actins[J]. *J Mol Biol*, **179**(3): 391-413.
- Zhang Y, Mao B. 2009. Developmental expression of an amphioxus (*Branchiostoma belcheri*) gene encoding a GATA transcription factor[J]. *Zool Res*, **30**(2): 137-143.
- Zhao S, Jiang H, Wang W, Mao B. 2007. Cloning and developmental expression of the *Xenopus* Nkx6 genes[J]. *Dev Genes Evol*, **217**(6): 477-483.