动 物 学 研 究 2009，Apr. 30(2): 137−143 CN 53-1040/Q ISSN 0254-5853
Zoological Research DOI：10.3724/SP.J.1141.2009.02137

Developmental Expression of an Amphioxus (*Branchiostoma belcheri*) Gene Encoding a GATA Transcription Factor

ZHANG Yu-jun[1,2], MAO Bing-yu[1,*]

(1. *CAS-Max Planck Junior Scientist Group on Neural Patterning, State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, the Chinese Academy of Sciences, Kunming* 650223, *China*; 2. *Graduate School of the Chinese Academy of Sciences, Beijing* 100049, *China*)

Abstract: GATA factors are evolutionarily conserved and play crucial roles during embryonic development in both vertebrates and invertebrates. Vertebrate GATAs can be divided into two subgroups, the GATA1/2/3 and the GATA4/5/6 classes. Through genomic analysis, we have identified three GATA factors, representing the GATA1/2/3 and GATA4/5/6 subfamilies respectively, and one GATA like protein in the genome of the basal chordate amphioxus (*Branchiostoma floridae*, cephalochordata). Partial sequence of *GATA123* in the amphioxus *Branchiostoma belcheri* (*BbGATA123*) was cloned and its expression pattern during early embryonic development was studied. Expression of *BbGATA123* is first detected in the mesendoderm during gastrulation. Interestingly, in the late neurula and early larva stages, it is expressed strongly in the cerebral vesicle and the mid gut region. Its expression is compared to *Otx*, a gene known crucial for the development of anterior structures. Our observations suggest that GATA123, together with Otx, might play an important role in the development of amphioxus cerebral vesicle, the counterpart of the vertebrate brain.

Key words: GATA factors; Expression pattern; Amphioxus; *Branchiostoma belcheri*

文昌鱼一个 GATA 基因在胚胎发育中的表达图式

张煜珺[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：GATA基因在脊椎动物和非脊椎动物的发育中行使重要的功能，该家族的成员在进化上也是非常保守的。脊椎动物的GATA基因分为两个亚群：GATA1/2/3和GATA4/5/6。通过生物信息分析，在文昌鱼的基因组中找到了3个GATA基因：一个GATA1/2/3亚家族基因，两个GATA4/5/6亚家族基因；还找到一个类GATA基因。还克隆了白氏文昌鱼(*Branchiostoma belcheri*)GATA123的一段序列，并研究了它在早期胚胎发育中的表达图式。结果表明GATA123在原肠胚的中内胚层表达，而在神经胚晚期和幼体早期，GATA123在脑泡和消化道中部区域表达。这种表达模式与头部发育的重要基因Otx相类似。结果提示在文昌鱼脑泡的发育过程中GATA123和Otx很可能共同发挥着重要的作用。

关键词：GATA 因子；表达图式；文昌鱼；白氏文昌鱼
中图分类号：Q959.287; Q344; Q593.4; Q786 **文献标志码**：A **文章编号**：0254-5853(2009)02-0137-07

The GATA family transcription factors are named due to their ability to bind the consensus DNA sequence (A/T) GATA (A/G). Members of this group have been identified in organisms ranging from cellular slime mold, plants to vertebrates. Vertebrate GATAs and most of the metazoan GATA factors contain two distinctive zinc-finger domains followed by a conserved highly basic region. Several reports have demonstrated that the C-terminal zinc-finger and the adjacent basic domain are necessary for DNA binding *in vitro* (Molkentin, 2000) and only these DNA-binding domains are conserved throughout this protein family. In fact, many of protostome GATAs contain a single copy of the zinc-finger domain such

Received date: 2008-12-15; Accepted date: 2009-02-23
Foundation items: Supported by grants from the National Natural Science Foundation of China (30425011; 30530380)
*Corresponding author（通讯作者），E-mail: mao@mail.kiz.ac.cn
收稿日期：2008-12-15；接受日期：2009-02-23
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）

as fungal GATAs (Orkin, 1992).

Six GATA factors have been identified in vertebrates, five in *Drosophila* and eleven in nematode *Caenorhabditis elegans*. Through phylogenetic and functional analysis, the vertebrate GATA factors are divided into two classes: the GATA1/2/3 and the GATA4/5/6 classes (Lowry & Atchley, 2000; Molkentin, 2000; Patient & McGhee, 2002). The GATA1/2/3 factors are associated with erythroid and neural specification and the GATA4/5/6 factors play partially redundant roles in mesoderm and endoderm development (Patient & McGhee, 2002). In both vertebrates and invertebrates, the GATA1/2/3 genes are expressed in early ectodermal lineages (Nardelli et al, 1999; Tsarovina et al, 2004; Xu et al, 1997), whereas the GATA4/5/6 genes are expressed in mesendodermal lineages (Holtzinger & Evans, 2005; Molkentin, 2000; Patient & McGhee, 2002; Peterkin et al, 2005; Welch et al, 2004).

In addition to the GATAs, a GATA like protein (GATA like protein-1, GLP-1) has also been reported in mouse that contains only one GATA type zinc finger domain. The basic domain is not conserved in GLP-1 and it lacks the ability to bind the (A/T) GATA (A/G) sequence. In mouse, GLP-1 is required in somatic cells of the gonad for germ cell development (Li et al, 2007).

Two GATAs, representing GATA1/2/3 and GATA4/5/6 orthologs respectively, have been found or predicted in basal invertebrate deuterostomes, including urochordates and echinoderms (Gillis et al, 2007). The three vertebrate paralogs in each class have been suggested to arise from two whole genome duplication events that occurred during the evolution of vertebrates (Dehal & Boore, 2005). However, the GATA genes in amphioxus, the cephalochordates, are not yet studied. Through genomic analysis, we detected three GATA factors and one GATA like protein in the genome of amphioxus (*Branchiostoma floridae*). We have cloned partial sequence of *GATA123* gene of *Branchiostoma belcheri* and studied its expression pattern during early embryonic development.

1 Materials and Methods

1.1 Embryos

Adult amphioxus (*B. belcheri*) were collected during the breeding season from the South China Sea near Beihai (Guangxi Province, China) and maintained in the laboratory. Naturally fertilized eggs were collected and cultured at room temperature. The developing embryos and larvae at desired stages were fixed in fresh 4% paraformaldehyde at room temperature for 30 min or at 4℃ overnight, dehydrated in gradated methanol and stored in 70% methanol at −20℃. Some adults, embryos and larvae at different developmental stages were frozen in liquid nitrogen for DNA and RNA extraction.

1.2 Isolation of genes

The genome of amphioxus (*B. floridae*, http://genome.jgi-psf.org/Brafl1) was BLASTed against mouse GATA proteins and 4 hits were found. GenomeScan was used for predicting the GATA genes aided by EST database searches and manual correction.

Partial sequence of *B. belcheri* GATA123 was cloned by RT-PCR using cDNAs of 10-somites neurula as the template. Primers were designed according to an EST clone of *B. floridae*. Primers used were: 5'-CGACGTGTTCTTCCACCACCTC-3' and 5'-CTGCGACACTGACGAGGAAGAGA-3'. PCR products were cloned into PBS-T vector (Tiangen) and sequenced.

The predicted *BfGATA* genes and partial sequences of *BbGATA123* and *Otx* were submitted to the GenBank under accession numbers FJ615537-FJ615542.

1.3 *In situ* hybridization

Antisense digoxygenin RNA probes were prepared using properly linearized template and T7 RNA polymerase. Whole mount *in situ* hybridizations were performed in home-made baskets using the standard amphioxus protocol (Holland, 1999) with minor modifications. Embryos and larvae stored in 70% methanol were re-hydrated. Specimens were digested with 5 μg/mL proteinase K in PTW (1X PBS, 0.1% Tween 20) for 10 min. Digestion was stopped with 2 mg/mL glycine in PTW (5 min) and specimens were refixed for 1 h in 4% paraformaldehyde in PTW. After washing in PTW, specimens were acetylated with 0.25% and 0.5% acetic anhydride in 0.1 mol/L triethanolamine, washed in PTW followed by prehybridization in a hybridization buffer (50% deionized formamide (V/V), 0.01 g/mL Boehringer

Block, 1 mg/mL yeast RNA, 100 μg/mL heparin, 0.1% chaps, 5XSSC (saline sodium citrate), 0.1% Tween 20, 5 mmol/L EDTA) at 60℃ for 3 h. Hybridization was performed in the same hybridization buffer with 1 μg/mL antisense probe at 60℃ for 16 h. The embryos were then washed twice (30 minutes each time) at 55 ℃ with 50% formamide/5X SSC/0.1% Tween 20, twice with 50% formamide/2X SSC/0.1% Tween 20, and twice with 50% formamide/1X SSC/0.1% Tween 20. After 3 washes with PTW at room temperature, the embryos were incubated in the blocking solution (10 mg/mL Boehringer blocking reagent, 2 mg/mL BSA in PTW) for 3h. AP coupled Anti-DIG antibody (Roche, 1:1500 in blocking solution, preabsorbed with amphioxus powder) was then added and incubated at 4 ℃ overnight. After 3 washes with PTW and APT buffer (0.1 mol/L NaCl, 0.1 mol/L Tris PH9.5, 0.05 mol/L $MgCl_2$, 1% Tween), the embryos were transferred to stain solution in multi-well plates. The embryos were stored in dark and monitored for color reaction. The staining process may take from 2h to 5 days. The reaction was stopped by washing 2−3 times with PTW and the embryos were transferred into 50% glycerol for storage.

2 Results

2.1 Isolation of the amphioxus GATA genes

The genome of amphioxus (*B. floridae*) was BLASTed against mouse GATA proteins and 4 hits for GATA type of zinc fingers were found. Genomic analysis and EST database searches suggest that one of them corresponds to the GATA1/2/3 type factor, two of them correspond to two close homologs of the GATA4/5/6 type factors, and the fourth one corresponds to the amphioxus GATA like protein (hypothetical protein BRAFLDRAFT_103415, XP_002241081). They were named BfGAGA123, BfGAGA456a, BfGATA456b and BfGLP respectively. The predicted *BfGATA123* is supported over the whole open reading frame by ESTs. The open reading frame is encoded by 5 exons, same as its vertebrate GATA1/2/3 homologs. Through EST assembly, one alternative splicing isoform is also found, which lacks 236 amino acids in the N-terminal and was named *BfGAGA123* short (*BfGAGA123s*). The first 4 coding exons of *BfGAGA456a* and 3 of *BfGATA456b* can be predicted from the *B. floridae* genome. *BfGATA456a* and *b* show 98% identity over the predicted coding regions and only one amino acid difference at the protein level. We can not rule out the possibility that the two genes are actually the same one assembled into two scaffolds due to assembly mistake or polymorphism. If they represent two genes, they must have arisen from a recent duplication event. In vertebrates, the GATA4/5/6 factors are encoded by 6 exons. We predict that the last 2 and 3 exons of *BfGATA456a* and *b* respectively are missing due to sequence gaps. BfGATA123 and BfGATA456 both contain two conserved GATA type zinc finger domains as well as an adjacent basic domain (Fig.1). In both cases, the dual zinc finger domains are encoded by 3 exons with similar intron/exon boundaries, as in all deuterosterm GATA genes examined (Gillis et al, 2008). Most of the conserved class-specific motifs identified by Gillis et al (2007) were found in the predicted BfGATA factors, further supporting our predictions. BfGLP contains only one GATA type zinc finger domain and only a few adjacent basic amino acids, like its mouse homolog. No EST clone corresponding to the BfGAGA456a, BfGATA456b and BfGLP could be found in the NCBI EST database, suggesting that their expressions were quite low or very specific.

Phylogenetic analysis using the conserved zinc finger domains showed that the four proteins falled into the GATA1/2/3, GATA4/5/6 and the GLP branches respectively (Fig. 2). The amphioxus GATA123 and GATA456 locate at the root of the GATA1/2/3 and GATA4/5/6 subclasses respectively as expected.

2.2 Expression of BbGata123 and Otx during amphioxus embryogenesis

At gastrula stage, *BbGATA123* is expressed in the invaginating mesendoderm (Fig. 3A, B) but not in the ectoderm. At early neurula stage, it is detected in the forming somite region and the endoderm (Fig. 3C). Interestingly, at the late neurula stage (18 h), the expression of BbGATA123 becomes localized in the anterior tip of the mesendoderm, the cerebral vesicle, and the mid-gut region (Fig. 3D). In the 24-hour larva, it is strongly expressed in the cerebral vesicle, the floor plate of the anterior intestine and weakly in the tailbud region (Fig. 3E). However, its expression becomes very weak in the 45-hour larvae and no clear pattern could be detected (data not shown).

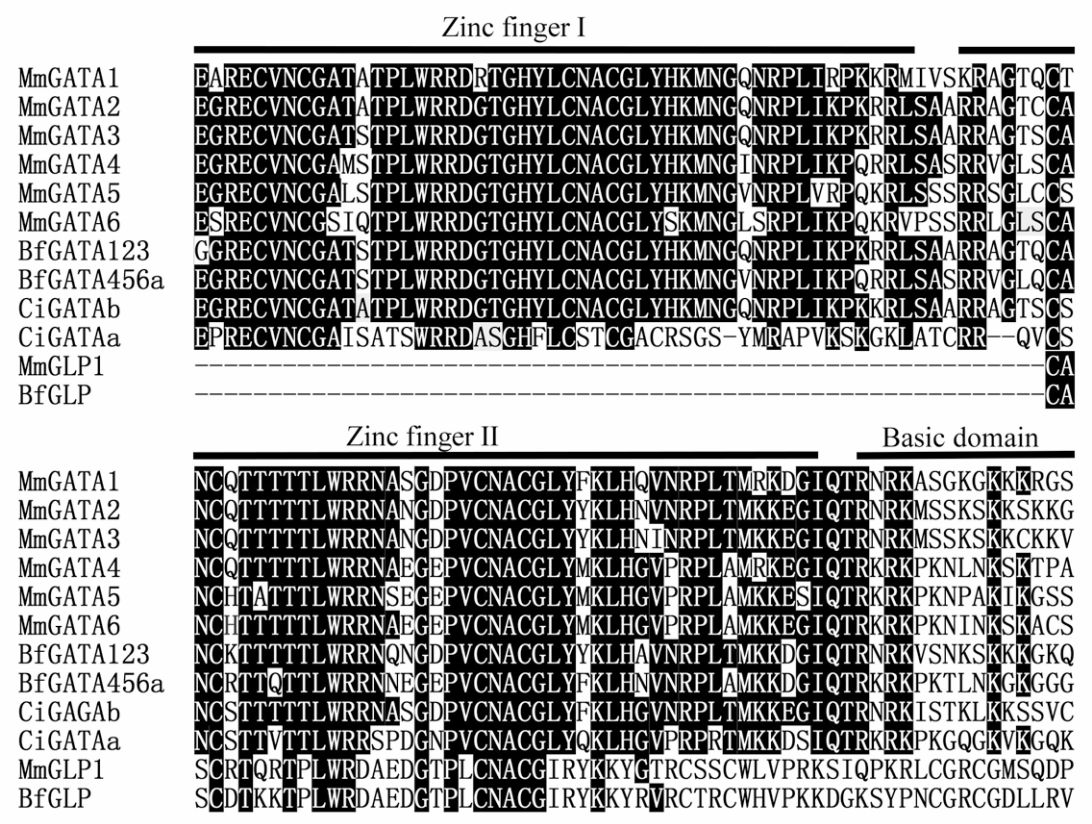

Fig. 1 Alignment of the zinc finger domains and the adjacent basic regions of the amphioxus, mouse and tunicate GATA factors

Black shading, identities; dashed lines, sequence gaps. Mm：*Mus musculus*；Ci：*Ciona intestinalis*；Bf：*Branchiostoma floridae*. The mouse and tunicate GATA protein sequences were used as in Gillis et al (2007).

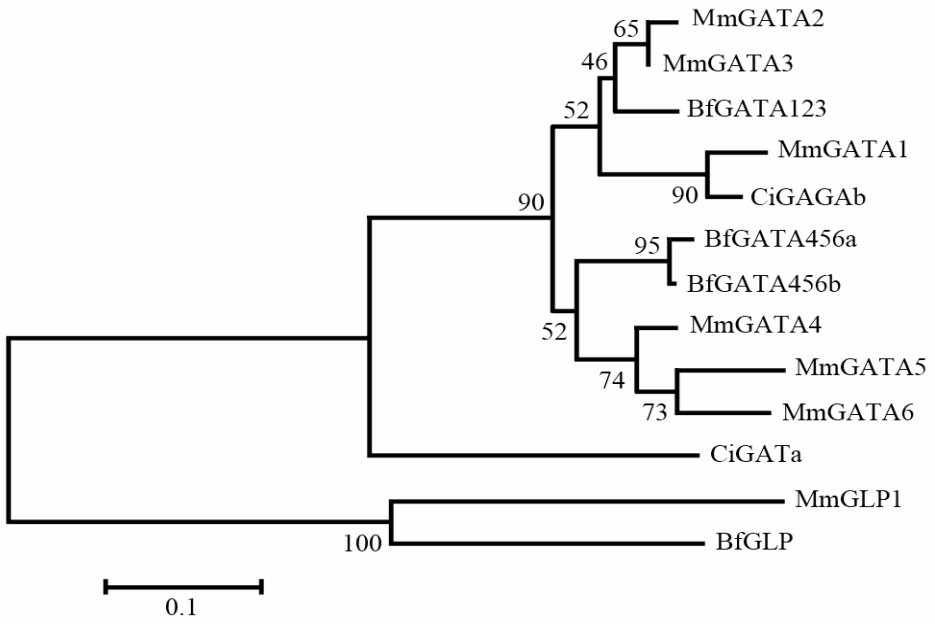

Fig. 2 Evolutionary relationships of GATA factors in amphioxus and other species

Tree topology was determined using neighbor joining method based on the alignment shown in Fig. 1. The four amphioxus GATA factors are classed into GATA1/2/3, GATA4/5/6 and GLP branches respectively. Mm：*Mus musculus*；Ci：*Ciona intestinalis*；Bf：*Branchiostoma floridae*.

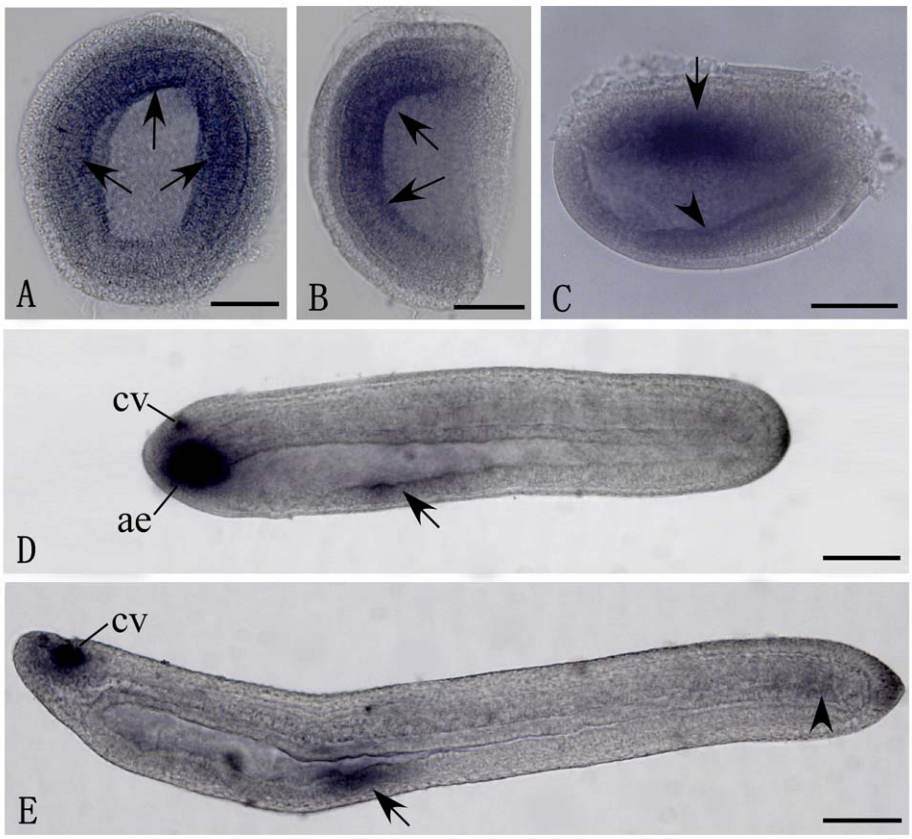

Fig. 3 Embryonic and larval expression of BbGATA123 in whole-mounts

Blastopore and lateral views (B) of a mid-gastrula embryo showing the expression in the invaginating mesendoderm (arrows). (C) Early neurula showing expression in the somite region (arrow) and endoderm (arrowhead). (D) Late neurula (18 hrs) with strong expression in the cerebral vesicle (cv), the anterior dorsal endomesoderm (ae) and weak expression in the floor plate of the anterior intestine (arrow). (E) Early larva (24 h) with strong expression in the cerebral vesicle (cv), the floor plate of the anterior intestine (arrow) and weakly in the tailbud region (arrowhead). (C)−(E), lateral views, anterior to the left. Scale bars, 50 μm.

The expression of BbGATA123 in the cerebral vesicle of early larvae suggests that it might have a role in the anterior-posterior patterning of the nervous system. Since Otx is well-known to play a role in this process, we compared their expression patterns during amphioxus embryonic development. In *B. floridae*, *Otx* was strongly expressed in anterior neural plate and mesendoderm at early neurula stage and later in the cerebral vesicle and anterior endoderm (Williams & Holland, 1996). A fragment of *B. belcheri Otx* was cloned by RT-PCR, which showed 92% identity to its homolog in *B. floridae*. In *B. belcheri*, *Otx* is expressed in the invaginating mesendoderm at gastrula stage (Fig. 4A, B). At early neurula stage (5 somites), its expression becomes localized in the anterior part of the embryo, including the neural plate and endoderm (Fig. 4C). In mid and late neurula stage embryos (14 and 18 hours), it is strongly expressed in the anterior dorsal endoderm and weakly in the forming cerebral vesicle and the anterior notochord (Fig. 4D, E). In the 45-hour larva, *Otx* is strongly expressed in the cerebral vesicle and the pharyngeal region (Fig. 4F).

3 Discussion

In this study, we identified the GATA factors from the amphioxus (*B. floridae*) genome. We have reconstructed the evolutionary relationships of these GATAs using molecular phylogenetic analysis. Our analysis indicates that amphioxus genome has a single GATA123 ortholog and possibly two close GATA456 paralogs. Expression analysis indicates that amphioxus GATA123 might be involved in the development of cerebral vesicle and anterior endoderm.

In vertebrates, GATA factors play various roles in different developmental processes (Patient &

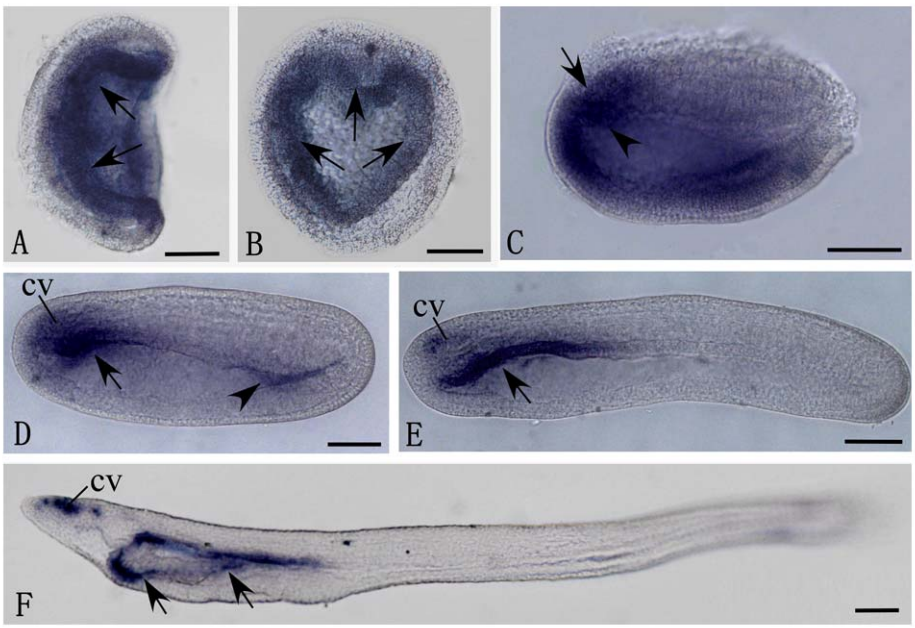

Fig. 4 Embryonic and larva expression of *Otx* in *Branchiostoma belcheri*

(A) Lateral and (B) blastopore views of a gastrula with expression in the invaginating inner layer (arrows). (C) Early neurula with expression in the anterior neural plate (arrow), the underlying mesendoderm (arrowhead) and broadly in the endoderm. (D) Mid neurula (14hrs) with strong expression in the cerebral vesicle (cv), anterior mesendoderm (arrow) and weak expression in the posterior endoderm (arrowhead). (E) Late neurula (18hrs) with expression in the cerebral vesicle and anterior dorsal endoderm (arrow). (F) Early larva (45hr) with expression in the cerebral vesicle (cv) and the pharyngeal region (arrows). (C)−(F), lateral views, anterior to the left. Scale bars, 50 μm.

McGhee, 2002). The GATA1/2/3 genes are expressed in the hematopoietic cell lineages and are essential for erythroid, megakaryocyte and T lymphocytes development (Orkin, 1992; Viger et al, 2008). In addition, GATA 2 and 3 are also expressed in the hindbrain, spinal cord and inner ear in mouse (Nardelli et al, 1999). The co-expression of amphioxus GATA123 with Otx in the cerebral vesicle, the counterpart of vertebrate brain, suggests that GATA factors might have a primitive role in brain development. In the endoderm, *Otx* is expressed in the pharyngeal region while *GATA123* in the anterior intestine region at larva stages, indicating their different roles.

Vertebrate GATA4/5/6 genes are mainly expressed in the mesodermal and endodermal tissues, such as heart, gut and gonads. They play important roles during the liver and pancreas development, which all derived from the anterior gut region (Molkentin, 2000). The roles of GATA factors in endoderm development have been evolutionarily conserved. In amphioxus, the anterior part of endoderm develops into the pharyngeal region and the hepatic diverticulum (the counterpart of vertebrate liver) develops from the anterior intestine region. The expression of amphioxus GATA123 in this region might be related to the specification of the hepatic diverticulum. However, in vertebrates, this function is fulfilled by the GATA4/5/6 class factors. The expression patterns of *BbGATA123* suggest that the functions of the two GATA subfamilies are not yet separated in amphioxus as in vertebrates. In vertebrates, GATA transcription factors are key regulators of hematopoiesis. However, we did not detect such expression patterns in amphioxus that could be indicative of its involvement in hematopoiesis. It will be of interest to check the expression of the amphioxus *GATA456* gene. The lack of its EST clones in the NCBI database suggests that its expression might be very low or specific.

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

amphioxus materials.

References:

Dehal P, Boore JL. 2005. Two rounds of whole genome duplication in the ancestral vertebrate[J]. *PLoS Biol*, **3**(10): e314.

Gillis WJ, Bowerman B, Schneider SQ. 2007. Ectoderm- and endomesoderm-specific GATA transcription factors in the marine annelid *Platynereis dumerilli*[J]. *Evol Dev*, **9**(1): 39-50.

Gillis WQ, Bowerman BA, Schneider SQ. 2008. The evolution of protostome GATA factors: molecular phylogenetics, synteny, and intron/exon structure reveal orthologous relationships[J]. *BMC Evol Biol*, **8**: 112.

Holland PW. 1999. Whole mount *in situ* hybridization to amphioxus embryos[J]. *Methods Mol Biol*, **97**: 641-644.

Holtzinger A, Evans T. 2005. Gata4 regulates the formation of multiple organs[J]. *Development*, **132**(17): 4005-4014.

Li S, Lu MM, Zhou D, Hammes SR, Morrisey EE. 2007. GLP-1: a novel zinc finger protein required in somatic cells of the gonad for germ cell development[J]. *Dev Biol*, **301**(1): 106-116.

Lowry JA, Atchley WR. 2000. Molecular evolution of the GATA family of transcription factors: conservation within the DNA-binding domain[J]. *J Mol Evol*, **50**(2): 103-115.

Molkentin JD. 2000. The zinc finger-containing transcription factors GATA-4, -5, and -6. Ubiquitously expressed regulators of tissue-specific gene expression[J]. *J Biol Chem*, **275**(50): 38949-38952.

Nardelli J, Thiesson D, Fujiwara Y, Tsai FY, Orkin SH. 1999. Expression and genetic interaction of transcription factors GATA-2 and GATA-3 during development of the mouse central nervous system[J]. *Dev Biol*, **210**(2): 305-321.

Orkin SH. 1992. GATA-binding transcription factors in hematopoietic cells[J]. *Blood*, **80**(3): 575-581.

Patient RK, McGhee JD. 2002. The GATA family (vertebrates and invertebrates)[J]. *Curr Opin Genet Dev*, **12**(4): 416-422.

Peterkin T, Gibson A, Loose M, Patient R. 2005. The roles of GATA-4, -5 and -6 in vertebrate heart development[J]. *Semin Cell Dev Biol*, **16**(1): 83-94.

Tsarovina K, Pattyn A, Stubbusch J, Muller F, van der Wees J, Schneider C, Brunet JF, Rohrer H. 2004. Essential role of Gata transcription factors in sympathetic neuron development[J]. *Development*, **131**(19): 4775-4786.

Viger RS, Guittot SM, Anttonen M, Wilson DB, Heikinheimo M. 2008. Role of the GATA family of transcription factors in endocrine development, function, and disease[J]. *Mol Endocrinol*, **22**(4): 781-798.

Welch JJ, Watts JA, Vakoc CR, Yao Y, Wang H, Hardison RC, Blobel GA, Chodosh LA, Weiss MJ. 2004. Global regulation of erythroid gene expression by transcription factor GATA-1[J]. *Blood*, **104**(10): 3136-3147.

Williams N, Holland P. 1996. Old head on young shoulders[J]. *Nature*, **383**(6600): 490.

Xu RH, Kim J, Taira M, Lin JJ, Zhang CH, Sredni D, Evans T, Kung HF. 1997. Differential regulation of neurogenesis by the two Xenopus GATA-1 genes[J]. *Mol Cell Biol*, **17**(1): 436-443.