Genes Genet. Syst. (2005) **80**, p. 233–236

Construction of a BAC Library for Chinese Amphioxus *Branchiostoma belcheri* and Identification of Clones Containing *Amphi-Pax* Genes

Wei Wang[1,2], Huai-Liang Xu[2], Lu-Ping Lin[1], Bing Su[2] and Yi-Quan Wang[1]*

[1]*Key Laboratory of the Ministry of Education for Cell Biology and Tumor Cell Engineering, School of Life Sciences, Xiamen University, Xiamen, 361005, China*
[2]*Key Laboratory of Cellular and Molecular Evolution, Kunming Institute of Zoology, The Chinese Academy of Sciences, Kunming, 650223, China*

(Received 21 May 2005, accepted 16 July 2005)

Amphioxus is a crucial organism for the study of vertebrate evolution. Although a genomic BAC library of *Branchiostoma floridae* has been constructed, we report here another BAC library construction of its distant relative species *Branchiostoma belcheri*. The amphioxus BAC library established in present study consists of 45,312 clones arrayed in one hundred and eighteen 384-well plates. The average insert fragment size was 120 kb estimated by Pulsed Field Gel Electrophoresis (PFGE) analysis of 318 randomly selected clones. The representation of the library is about 12 equivalent to the genome, allowing a 99.9995% probability of recovering any specific sequence of interest. We further screened the library with 4 single copied *Amphi-Pax* genes and identified total of 26 positive clones with average of 6.5 clones for each gene. The result indicates this library is well suited for many applications and should also serve as a useful complemental resource for the scientific community.

Key words: Amphioxus, *Amphi-Pax* genes, Bacterial artifcial chromosome (BAC) library, *Branchiostoma belcheri*

Amphioxus, at the boundary between invertebrates and vertebrates, occupies an extremely important phylogenetic position as a sister group to vertebrate within the phylum Chordata (Gee, 1994). Its relative small genome does not undergo the large-scale gene duplications took place during the lineage divergence of vertebrates and contains around only half the number of genes as that of mammals, probably also lacks the extensive repeat elements in vertebrate genomes (Gibson-Brown et al., 2003). This small genome should be relatively easy to characterize, thus can provide ideal out group references for understanding the evolutionary origin of vertebrate genome. So far extensive efforts devoted toward the amphioxus genes (Takahashi and Holland, 2004; Oda et al., 2004; Zhang et al., 2004), particularly toward the large scale cDNA sequencing (Mou et al., 2002; Panopoulou et al., 2003; Suzuki and Satoh, 2000), have created a strong demand for deep coverage large-insert genomic library, especially for three widely studied species *Branchiostoma floridae*, *B. lanceolatum* and *B. belcheri*. The first BAC library constructed on *B. floridae* was reported (http://bacpac.chori.org) already and its whole genome sequencing is under way (http://www.jgi.doe.gov/). In addition, NHGRI (National Human Genome Research Institute) is considering sequencing species (*B. lanceolatum*) in the future (http://www.genome.com). Although above three species are now placed in a same genus taxonomically, the genetic distance among those amphioxus far surpasses that between primate and rodent (Minguillon et al., 2002; Wang et al., 2004). Considering of interspecies divergence during the independent evolution, the genomic information derived from single species only is insufficient to explore the last common ancestor of vertebrate and cephalochordate (Minguillon et al., 2002). In addition, recent analysises based on mitochondrial genes revealed that the West Pacific amphioxus is highly divergent from other two species (Nohara et al., 2004; Xu et al., 2005). Therefore we constructed a BAC library of Chinese amphioxus *B. belcheri* and characterized the library using *Amphi-Pax* genes in present study. The library provides a very useful complemental resource for comparative genomic and evolutionary researches.

High-molecular-weight DNA (HMW DNA) was extracted from *B. belcheri* freshly collected from Xiamen beach. Samples were homogenized in our modified extraction

Edited by Takashi Endo
* Corresponding author. E-mail: wangyq@xmu.edu.cn

buffer (137 mM NaCl, 2.7 mM KCl, 10 mM $NaHPO_4$, 2 mM KH_2PO_4, 60 mM EDTA pH 7.4). After three times washing of homogenate, the cell density was adjusted to 5×10^8/ml, little bit higher than that in conventional method because amphioxus genome is about 17% of that in vertebrates, and mixed with 1% InCert agarose (Bio-Whittaker Molecular Applications) to generate plugs. The plugs are treated with proteinase K and then subjected to a brief pre-electrophoresis to remove the degraded DNA. Genomic DNA was partially digested by *Eco*R I/*Eco*R I methylase competition and selected in the size ranging from 100–180 kb by two successive pulsed field gel electrophoresis (PFGE). Following the electroelution and desaltification, DNA segments were ligated to vector pBACe3.6 and used to transform DH10B cells electrically. Total of 45,312 colonies are manually picked and arrayed into one hundred and eighteen 384-well microtiter plates that contain LB medium supplemented with 10% glycerol and 12.5 μg/ml chloramphenicol. After overnight incubation at 37°C, the plates were duplicated and stored at –80°C.

To evaluate the quality of the amphioxus BAC library, we extracted plasmid DNA from 318 randomly selected single colonies completely digested the DNA with *Not* I and separated them by PFGE. Figure 1a is a representative electrophoretic profile. Typically, the DNA fragments generated by *Not* I contained an 8.7 kb vector band and one or more larger bands. It was estimated that the recombination ratio of the library is about 90.3% and average inserted fragment size is 120 kb with the largest size at 330 kb. Figure 1b shows the size distribution of inserts in the BAC library, and meanwhile, it also exhibits most recombined DNA fragments range of 100–150 kb. Based on the estimation of 400 Mb amphioxus genome (http://bacpac.chori.org), we could deduce that the BAC library represents at least 12 genome-equivalents of amphioxus and allows a 99.9995% probability of recovering any specific sequence of the animal according to the formulas $R=N\cdot I/GS$ and $P=1-\mathrm{Exp}[N\cdot \ln(1-I/GS)]$, where R is genome representation of a library, N the number of recombinant clones, I average insert fragment size and GS the genome size of organism.

For further evaluation and interested clones selection, we performed PCR-based library screening using primers derived from *Amphi-Pax* gene family (Table 1). Previous studies show that *Pax* family encode a family of transcription factors which play key roles in animal embryonic development and there are four *Pax* genes in amphioxus, namely *Amphi-Pax 1* (Holland, 1996) *Amphi-Pax 2/5/8* (Kreslova et al., 2002), *Amphi-Pax* 3/7 (Holland et al., 1999) and *Amphi-Pax 6* (Glardon et al., 1998), each represents one of the four well-defined *Pax* gene subfamilies (Balczarek et al., 1997), suggesting that the members of all four subfamilies were present in the genome of the most recent common ancestor of vertebrate and amphioxus. The exploration of *Amphi-Pax* genes is critical for understanding the tempo and mode of evolutionary processes as well as the functional differentiation of the genes. Present library screening identified a total of 26 positive clones carrying *Amphi-Pax* genes (Table 1). Sequencing verifications of those amplicons show that they are 98%, 97%, 98% and 94% similar to *Amphi-Pax 1*, *Amphi-Pax 2/5/8*, *Amphi-Pax 3/7* and *Amphi-Pax 6* genes of *B. floridae* respectively confirming those clones carry the *Amphi-Pax* genes of *B. belcheri*. Interestingly however, although the library is estimated to represent 12-fold coverage of the amphioxus genome, the average number of positive clones carrying the *Amphi-Pax* genes is only 6.5. The disagreement between estimated clone coverage and PCR-screening probably results from a bias in clone coverage of the genome due to the use of single

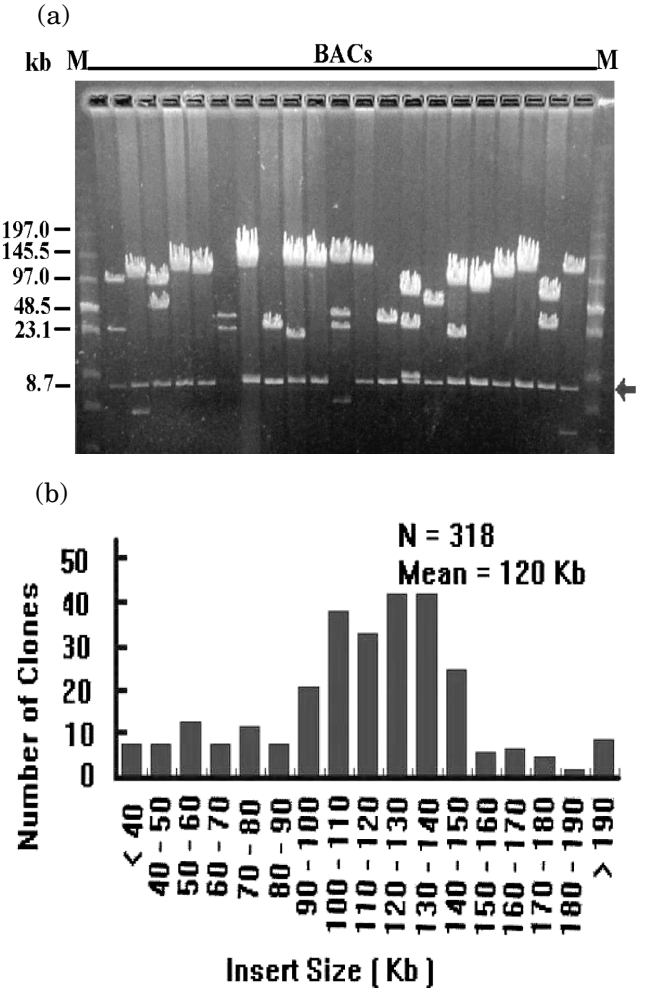

Fig. 1. Estimation (a) and distribution (b) of insert fragment size in BAC clones. Total of 318 randomly chosen BAC clones were digested with *Not* I. Then the digested DNAs were analyzed by pulsed-field gel electrophoresis on 1% agarose gels in 0.5 × TBE at 6 V/cm, with linear pulse ramping from 0.1 to 40 s for 14 hr at 14°C. The sizes of DNA markers (M) are indicated in kb, and the arrow denotes the vectors.

Table 1. PCR primers used for screening the BAC library in present study and number of positive clones revealed by each primer pair

Gene	Primer	5'-3' primer sequence	Annealing temperature	Ref. Seq. No.	Positive clones	Product Size
Amphi-Pax1	P1F	CATTTGGGGAGGTGAACCAGC	60°C	U20167	13	391 bp
	P1R	GTGTTGCCGATCTTGTTGCG				
Amphi-Pax2/5/8	P2F	TATGAGACAGGCTCAATCAAGC	57°C	AF053762	5	173 bp
	P2R	GGAACTGTGTCGTTGTCACA				
Amphi-Pax3/7	P3F	GCTCGGGGGAGTCTTCATCAA	57°C	AF165886	4	238 bp
	P3R	CCGGGGTTGTCTCTCTTGTA				
Amphi-Pax6	P6F	GCACAGTTCAAGCGGGAGTG	56°C	AJ223444	4	91 bp
	P6R	GGGATGTTCTCATTGGTACA				

restriction enzyme (*Eco*R I) to generate large DNA fragments, which could contribute to under- or over- representation of certain regions of the genome (Hong et al., 2003). Another explanation account for the phenomena might be incorrect genome size estimation of *B. belcheri*. If the fragments were randomly inserted to vector absolutely, we can calculate that the genome size of the amphioxus should be 700 Mb instead of 400 Mb.

In summary, we describe here the construction of a *B. belcheri* genomic DNA library with comprehensive coverage of the genome. The library quality is good enough for applications like chromosome walking, specific gene or region isolation, fluorescence in situ hybridization (FISH) and genome sequencing (Ding et al., 1999; Ripoll et al., 2000). As noted previously this is the second BAC library from *Branchiostoma*. The availability of this resource enables amphioxus interspecies comparison which would provides a much more robust determination of the latest common ancestor of extant chordates. Comparative genomic studies with vertebrate genomes would lead to better understanding of the molecular mechanism responsible for vertebrate novelties and provide insight into the origins of modern species. The knowledge would benefit both the biomedical and biological communities.

The authors thank Ms. Fan Xiaona for her technical aids. This work was supported by NSFC (No. 30470938), Natural Science Foundation of Fujian Province of China (No. D0510002) and grants from the Science and Technology Bureau of Xiamen (No. 3502Z20042015) and Key Laboratory of Cellular and Molecular Evolution, Kunming Institute of Zoology.

REFERENCES

Balczarek, K. A., Lai, Z. C., and Kumar, S. (1997) Evolution of functional diversification of the paired box (Pax) DNA-binding domains. Mol. Biol. Evol. **14**, 829–842.

Ding, Y., Johnson, M. D., Colayco, R., Chen, Y. J., Melnyk, J., Schmitt, H., and Shizuya, H. (1999) Contig assembly of bacterial artificial chromosome clones through multiplexed fluorescence-labeled fingerprinting. Genomics **56**, 237–246.

Gee, H. (1994). Vertebrate morphology: Return of the amphioxus. Nature **370**, 504–505.

Gibson-Brown, J. J., Osoegawa, K., McPherson, J. D., Waterston, R. H., De Jong, P. J., Rokhsar, D. S., and Holland, L. Z. (2003) A proposal to sequence the amphioxus genome submitted to the Joint Genome Institute of the US Department of Energy. J. Exp. Zoolog. Part B Mol. Dev. Evo. **300**, 5–22.

Glardon, S., Holland, L. Z., Gehring, W. J., and Holland, N. D. (1998) Isolation and developmental expression of the amphioxus *Pax-6* gene (*Amphi-Pax 6*): Insights into eye and photoreceptor evolution. Development **125**, 2701–2710.

Holland, L. Z. (1996) Muscle development in amphioxus: Mophology, biochemistry, and molecular biology. Israel. J. Zool. **42** (suppl), 235–246.

Holland, L Z., Schubert, M., Kozmik, Z., and Holland, N. D. (1999) *AmphiPax3/7*, an amphioxus paired box gene: insights into chordate myogenesis, neurogenesis, and the possible evolutionary precursor of definitive vertebrate neural crest. Evol. Dev. **1**, 153–165.

Hong, Y. S., Hogan, J. R., Wang, X., Sarkar, A., Sim, C., Loftus, B. J., Ren, C., Huff, E. R., Carlile, J. L., Black, K., Zhang, H. B., Gardner, M. J., and Collins, F. H. (2003) Construction of a BAC library and generation of BAC end sequence-tagged connectors for genome sequencing of the African malaria mosquito *Anopheles gambiae*. Mol. Genet. Genomics **268**, 720–728.

Krelova, J., Holland, L. Z., Schubert, M., Burgtorf, C., Benes, V., and Kozmik, Z. (2002) Functional equivalency of amphioxus and vertebrate *Pax2/5/8* transcription factors suggests that the activation of mid-hindbrain specific genes in vertebrates occurs via the recruitment of Pax regulatory elements. Gene **282**, 143–150.

Minguillon, C., Ferrier, D. E., Cebrian, C., and Garcia-Fernandez, J. (2002) Gene duplications in the prototypical cephalochordate amphioxus. Gene **287**, 121–128.

Mou, C. Y., Zhang, S. C., Lin, J. H. , Yang, W. L., Wu, W. Y., Wei, J. W., Wu, X. K., Du, J. C., Fu, Z. Y., Ye, L. T., Lu, Y., Xie, X. J., Wang, Y. L., and Xu, A. L. (2002) EST analysis of mRNAs expressed in neurula of Chinese amphioxus. Biochem. Biophys. Res. Commun. **299**, 74–84.

Nohara, M., Nishida, M., Manthacitra, V., and Nishikawa, T. (2004) Ancient phylogenetic separation between Pacific and Atlantic cephalochordates as revealed by mitochondrial genome analysis. Zoolog. Sci. **21** (2), 203–210.

Oda, H., Akiyama-Oda, Y., and Zhang, S. (2004) Two classic cadherin-related molecules with no cadherin extracellular repeats in the cephalochordate amphioxus: distinct adhesive specificities and possible involvement in the development of

multicell-layered structures. J. Cell. Sci. **117**, 2757–2767.

Panopoulou, G., Hennig, S., Groth, D., Krause, A., Poustka, A. J., Herwig, R., Vingron, M., and Lehrach, H. D. (2003) New evidence for genome-wide duplications at the origin of vertebrates using an amphioxus gene set and completed animal genomes. Genome Res. **13**, 1056–1066.

Ripoll, P. J., O'Sullivan, D. M., Edwards, K. J., and Rodgers, M. (2000) Technique for cloning and sequencing the ends of bacterial artificial chromosome inserts. Biotechniques **29**, 271–276.

Suzuki, M. M., and Satoh, N. (2000) Genes expressed in the amphioxus notochord revealed by EST analysis. Dev. Biol. **224**, 168–177.

Takahashi, T., and Holland, P. W. (2004) Amphioxus and ascidian *Dmbx* homeobox genes give clues to the vertebrate origins of midbrain development. Development **131**, 3285–3294.

Wang, Y. Q., Xu, Q. S., Peng, X. X., Zhou H. T. (2004) Taxonomic status of amphioxus *Branchiostoma blecheri* in Xiamen Beach estimated by homologous sequence of *Cyt b* gene. Acta Zoologica Sinica **50**, 202–208.

Xu, Q. S., Ma, F., and Wang, Y. Q. (2005) Morphological and 12S rRNA gene comparison of two Branchiostoma species in Xiamen waters. J. Exp. Zoolog. B. Mol .Dev. Evol. **304**B, 259–257.

Zhang, X. H., Yang, H., Huang, X. W., Wang, L. F., Chen, Z. K., Zhang, Y. J., Zhang, P. J., and Zhang, H. W. (2004) Expression pattern of *profilin* homologue in embryogenesis of amphioxus. Shi Yan Sheng Wu Xue Bao **37**, 283–288.