

## Ribosomal proteins L34 and S29 of amphioxus *Branchiostoma belcheri tsingtauense*: cDNAs cloning and gene copy number\*

Lina Liu, Shicui Zhang<sup>✉</sup>, Zhenhui Liu, Hongyan Li, Mei Liu, Yongjun Wang and Lifang Ma

Department of Marine Biology, Ocean University of China, Qingdao 266003, PR China;

<sup>✉</sup>e-mail: sczhang@ouc.edu.cn

Received: 20 December, 2004; revised: 23 February, 2005; accepted: 19 April, 2005  
available on-line: 11 July, 2005

The complete cDNA and deduced amino-acid sequences of ribosomal proteins L34 (*AmphiL34*) and S29 (*AmphiS29*) from the amphioxus *Branchiostoma belcheri tsingtauense* were identified in this study. The *AmphiL34* cDNA is 435 nucleotides in length and encodes a 118 amino-acid protein with calculated molecular mass of 13.6 kDa. It shares 53.6–67.5% amino-acid sequence identity with its eukaryotic counterparts including human, mouse, rat, pig, frog, catfish, fruit fly, mosquito, armyworm, nematode and yeast. The *AmphiS29* cDNA comprises 453 nucleotides and codes for a 56 amino-acid protein with a calculated molecular mass of 6.6 kDa. It shows 66.1–78.6% amino-acid sequence identity to eukaryotic S29 proteins from human, mouse, rat, pig, zebrafish, seahorse, fruit fly, nematode, sea hare and yeast. *AmphiL34* contains a putative nucleolar localization signal, while *AmphiS29* has a zinc finger-like domain. A phylogenetic tree deduced from the conserved sequences of *AmphiL34* and *AmphiS29* and other known counterparts indicates that the positions of *AmphiL34*/*AmphiS29* are intermediate between the vertebrate and invertebrate L34/S29. Southern blot analysis demonstrates the presence of one copy of the L34 gene and 2–3 copies of the S29 gene in the genome of the amphioxus *B. belcheri tsingtauense*. This is in sharp contrast to the existence of 7–9 copies of the L34 gene and 14–17 copies of the S29 gene in the rat genome. These data suggest that housekeeping genes like *AmphiL34* and *AmphiS29* have undergone large-scale duplication in the chordate lineage.

**Keyword:** amphioxus, *Branchiostoma*, ribosomal protein, L34, S29, copy number

Ribosomes are complex RNA–protein organelles that mediate the sequential addition of amino acids to the carboxyl end of the growing polypeptide chain, according to the blueprints encoded by mRNA (Ramakrishnan, 2002). In eukaryotes, ribosomes contain about 80 structural proteins, besides rRNA. Ribosomal proteins are highly conserved, and encoded by housekeeping genes, since their activity is required for the growth and maintenance of all cell kinds (Wool, 1979). The study of ribosomal proteins in a variety of organisms may improve our understanding of ribosome structure and evolution, and elucidate the role of ribosomal proteins in the basic mechanisms of protein synthesis.

The genes encoding ribosomal proteins L34 and S29 have been identified in some eukaryotic

species including human (Chan *et al.*, 1993; Frigerio *et al.*, 1995), mouse (Strausberg & Feingold, 2002), rat (Chan *et al.*, 1993; Aoyama *et al.*, 1989), pig (Wintero *et al.*, 1996), frog (Vaccaro, 2003), fish (Karsi *et al.*, 2002), fruit fly (AY071577 and AY075533 in GenBank), mosquito (Niu & Fallon, 1999), armyworm (AY072290 in GenBank), sea hare (Moccia *et al.*, 2003), nematode (NM\_069929 and NM\_064862 in GenBank) and yeast (Wood *et al.*, 2002). Amphioxus or lancelet, a cephalochordate, is widely known as a living fossil, representing an organism transitional from invertebrates to vertebrates in phylogeny (Stokes & Holland, 1998; Zhang *et al.*, 2002). Liu *et al.* (2004a) and Liu *et al.* (2004b) have recently reported the cloning of ribosomal proteins S15a, L19, S20 and L10 cDNAs from amphioxus *Branchiostoma belcheri tsingtauense*. However, no information has

\*GenBank accession No. AY168761 (*AmphiL34*); GenBank accession No. AY264807 (*AmphiS29*).

**Abbreviations:** ORF, open reading frame; SDS, sodium dodecyl sulfate; UTR, untranslated region.

been available so far for L34 and S29 in this evolutionarily important animal.

Gene and genome duplication has been an interesting topic to biologists for decades (Ohno, 1970; Meyer & Malaga-Trillo, 1999; Sankoff, 2001). It is proposed that two rounds of large-scale gene duplication took place during the early chordate evolution: one occurred close to the origin of vertebrates, and the other close to the origin of jawed vertebrates (Holland *et al.*, 1994; Sharman & Holland, 1996; Sidow, 1996; Meyer & Scharl, 1999). This has been substantially evidenced by comparison of the number of the luxury protein genes such as *Hox* (Holland & Garcia-Fernandez, 1996), *Otx* (William & Holland, 1998), *Msx* (Sharman *et al.*, 1999) and *hedgehog* (Shimeld, 1999). Evolutionarily, whether the housekeeping genes like L34 and S29 also follow the two-round duplication rule remains unknown, and data concerning comparison of housekeeping gene copy numbers in different species are basically lacking. The aims of the present study were thus to characterize L34 and S29 cDNAs from the amphioxus *B. belcheri tsingtauense* and to determine their gene copy numbers in its genome.

## MATERIAL AND METHODS

The amphioxus gut cDNA library was constructed as described by Liu *et al.* (2002). cDNA clones were selected at random for sequencing. Both strands of all selected clones were sequenced with an ABI PRISM 377XL DNA Sequencer and all sequences were then analyzed for coding probability with the DNATools program (Rehm, 2001).

Initial comparison against the GenBank protein database was performed using the BLAST network server at the National Center for Biotechnology Information (Altschul *et al.*, 1997). Multiple protein sequences were aligned using the MegAlign program by the CLUSTAL method in the DNASTAR software package (Burland, 2000). Phylogenetic tree

was constructed by the neighbor-joining method within the Philip 3.5c software package (Felsenstein, 1993) using 1000 bootstrap replicates. The accession numbers of the ribosomal protein sequences in the GenBank database used for comparison are listed in Table 1 and Table 2.

Genomic DNA for Southern blotting analysis was isolated from adult amphioxus. A total of 30 amphioxus were ground in liquid nitrogen, and the ground powder was suspended in 15 ml of lysis buffer containing 10 mM Tris/HCl (pH 8.0), 100 mM EDTA and 0.5% SDS. After treatment with proteinase K (100 mg/ml, final concentration) at 55°C for 3 h, it was cooled to room temperature and mixed with equal volume of Tris/HCl saturated phenol (pH 8.0). The mixture was centrifuged at 5000 × g at 4°C for 20 min, and the supernatant was pooled and then mixed with equal volume of phenol/chloroform (1:1, v/v). The mixture was centrifuged as above and the supernatant was collected. DNA was precipitated by ethanol, and digested with various restriction enzymes (2 units per µg DNA) at 37°C for 20 h. The digested DNA was separated on a 1% agarose gel using 1×TBE (89 mM Tris/borate and 2 mM EDTA) and transferred onto Nylon membrane (Osmonics Inc.). The membranes were hybridized with digoxigenin (DIG)-labeled DNA probes produced with a DIG DNA labeling kit (Roche). Hybridized bands were visualized according to the instruction of the detection kit.

## RESULTS AND DISCUSSION

The first cDNA clone encoding amphioxus ribosomal protein L34, *AmphiL34*, was identified from the gut cDNA library as revealed by Blast search. Figure 1 shows the nucleotide and deduced amino-acid sequences of *AmphiL34* cDNA (GenBank accession number: AY168761). It consisted of 435 bp, including a 33 bp 5' untranslated region (UTR), an open reading frame (ORF) of 357 bp and a 45 bp

**Table 1. Representative members of the eukaryotic ribosomal protein L34 family**

Protein	Organism (Abbreviation)	Accession number	Amino acids	Source
L34Hs	Human <i>Homo sapiens</i> (Hs)	BC070208	117	GenBank
L34Mm	House mouse <i>Mus musculus</i> (Mm)	BC058118	117	GenBank
L34Rn	Norway rat <i>Rattus norvegicus</i> (Rn)	P11250	117	Swiss-Prot
L34Ss	Pig <i>Sus scrofa</i> (Ss)	Q29223	117	Swiss-Prot
L34Xl	African clawed frog <i>Xenopus laevis</i> (Xl)	AY079177	117	GenBank
L34Ip	Channel catfish <i>Ictalurus punctatus</i> (Ip)	AF401588	117	GenBank
L34Bb	Amphioxus <i>Branchiostoma belcheri</i> (Bb)	AY168761	118	GenBank
L34Dm	Fruit fly <i>Drosophila melanogaster</i> (Dm)	AY071577	162	GenBank
L34Aa	Asian tiger mosquito <i>Aedes albopictus</i> (Aa)	AF144549	130	GenBank
L34Sf	Fall armyworm <i>Spodoptera frugiperda</i> (Sf)	AY072290	119	GenBank
L34Ce	Nematode <i>Caenorhabditis elegans</i> (Ce)	NM_069929	110	GenBank
L34Sp	Fission yeast <i>Schizosaccharomyces pombe</i> (Sp)	O42846	112	Swiss-Prot

Table 2. Representative members of the eukaryotic ribosomal protein S29 family

Protein	Organism (Abbreviation)	Accession number	Amino acids	Source
S29Hs	Human <i>Homo sapiens</i> (Hs)	BC032813	56	GenBank
S29Mm	House mouse <i>Mus musculus</i> (Mm)	S71578	56	GenBank
S29Rn	Norway rat <i>Rattus norvegicus</i> (Rn)	BC058150	56	GenBank
S29Ss	Pig <i>Sus scrofa</i> (Ss)	NM_001001633	56	Swiss-Prot
S29Dr	Zebrafish <i>Danio rerio</i> (Dr)	AY561513	56	GenBank
S29Hc	Tiger tail seahorse <i>Hippocampus comes</i> (Hc)	AY357068	56	GenBank
S29Bb	Amphioxus <i>Branchiostoma belcheri</i> (Bb)	AY264807	56	GenBank
S29Dm	Fruit fly <i>Drosophila melanogaster</i> (Dm)	AY075533	56	GenBank
S29Ce	Nematode <i>Caenorhabditis elegans</i> (Ce)	NM_064862	56	GenBank
S29Ac	California sea hare <i>Aplysia californica</i> (Ac)	AF486840	56	GenBank
S29Sc	Baker's yeast <i>Saccharomyces cerevisiae</i> (Sc)	P41058	56	Swiss-Prot

3' UTR. The ORF encoded a 118 amino-acid protein with a calculated molecular mass of about 13.6 kDa, and an isoelectric point of 11.44. The 5' UTR had an oligopyrimidine tract CTTTCGCCATTTT upstream of the start codon ATG, which consists of a C residue at the cap site, followed by a sequence of polypyrimidines (Amaldi & Pierandrei-Amaldi, 1990; Perry & Meyuhas, 1990; Sugawara *et al.*, 1992). It should be noted that there are two purines interrupting the polypyrimidine tract and this was verified by sequencing both strands of the 5' UTR. The oligopyrimidine tract possibly plays a critical role in translation control (Levy *et al.*, 1991). The 3' UTR included a polyadenylation signal AATAAA 16 bases upstream of the poly(A) site, which is required for post-translational cleavage-polyadenylation of the 3' end of the pre-mRNA (Proudfoot & Brownlee, 1976).

The deduced protein sequence of *AmphiL34* was compared with that of other known L34 from

diverse organisms in the GenBank database (Fig. 2). It showed that at the amino-acid level, *AmphiL34* shared more than 57.3% identity with its homologues in vertebrates such as human, mouse, rat, pig, frog and catfish, and more than 53.6% identity with those in other eukaryotes including invertebrates like fruit fly, mosquito, armyworm and nematode and fungi like yeast. Resembling other L34 proteins (Niu & Fallon, 1999), the amino-acid sequence conservation in *AmphiL34* is particularly high among the 38 N-terminal residues, and among residues 92–110 near the C-terminus.

*AmphiL34* is a rather hydrophobic protein with 46 hydrophobic amino acids out of 118 residues, and has a high percentage of basic amino acids (18 lysines and 16 arginines) mostly located in the N-terminal half of the deduced protein sequence, and a low percentage of acidic amino acids (1 aspartic acid and 3 glutamic acid) mostly situated in the C-terminal half. The strong basic character of L34s including *AmphiL34*, may be instrumental for its binding to rRNA in the 60S subunit of eukaryotic ribosomes (Ulbrich *et al.*, 1979; Dudov & Perry, 1984; Wiedemann & Perry, 1984). Moreover, the tetrapeptide RXXX, which is critical for nucleolar localization (Quaye *et al.*, 1996), is found at residues 88–91 in L34 proteins including *AmphiL34* (Fig. 2).

The second identified cDNA clone encoded amphioxus ribosomal protein S29 or *AmphiS29*. Figure 3 shows the nucleotide and deduced amino-acid sequences of *AmphiS29* cDNA (GenBank accession number: AY264807). The cDNA comprised 453 bp, and contained a 5' UTR of 21 bp, an ORF of 171 bp and a 3' UTR of 261 bp. The ORF encoded a 56 amino-acid protein with a calculated molecular mass of about 6.6 kDa and an isoelectric point of 9.79. The 5' UTR of *AmphiS29* had an oligopyrimidine tract CTCTTTGCCGATC critical for translation control. Like in *AmphiL34*, the polypyrimidine tract was interrupted by three purines and this was verified by sequencing both strands of the 5' UTR. The 3' UTR possessed a polyadenylation signal AATAAA required for post-translational cleavage-polyadenylation of the 3' end of the pre-mRNA.

All ORFs in extract from sequence L34.TXT  
All frames 435 base pairs Position: 1 to 435

```

1  GCTTTCGCCATTTTCACCTGAAAGGTTGCCAACCATGGTGCAGCGGCTGACATACAGCGGGGACTGTGTACAAAC 75
1  M T V Q R L T Y R R R L S Y N

76  ACCAAGTGAACCCGAGCGGCTCTCTAAGACTCCAGGTGGAAGAACTGGTCTATCTGTACACTAAGAAGAAAGGCG 150
15  T K S N R R R I S K T P G G K L V Y L Y T K K K G

151  AGTGTTCACGATGTGGGGAGTGTAAAGGTCAAGGTGTTGGGGTGAAGGCACCAAGACCCAAAGAGCTGATGTCC 225
40  S V P V C G D C K V K L F G V K A P R P K K L M S

226  ATGTACAGGAGGCTGAAGCAGCTCACCAGAGCTTACGGGGATCCCGCTGCTGCAAGTGTGTCCGGAAAGGATT 300
65  M S R R L K H V T R A Y G G S R C S K C V R E R !

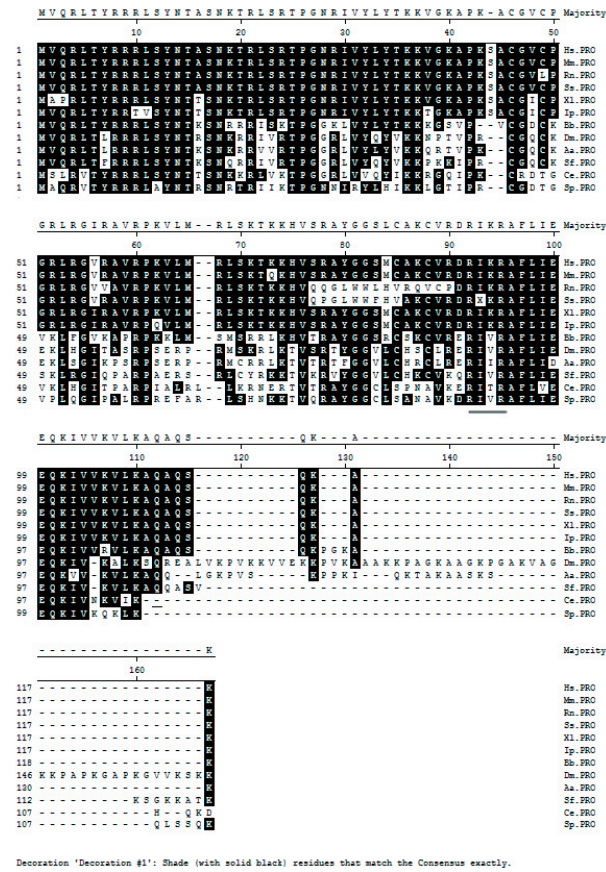
301  GTTGTGCTTTCTGATTGAGGAACAGAAGATTGTGGTGAAGTGTCTGAAGGCACAGGCTCAGTCACAGAAACCG 375
90  V R A F L I E E Q K I V V R V L K A Q A Q S Q K P

376  GGCAAGGCCAAGTAAACAGGACCAACAACAACACATGTAGACACAAATAAACACCAAGC 435
115  G K A K *

```

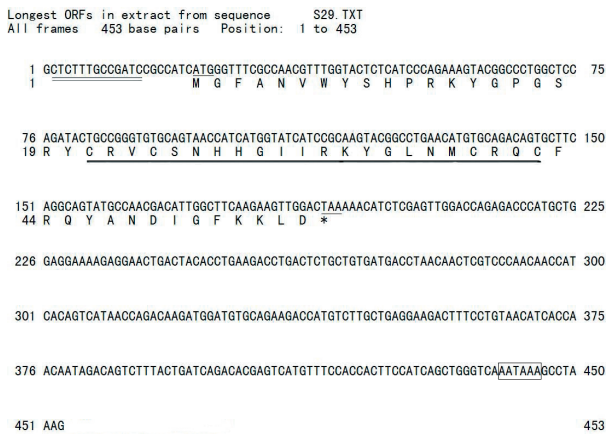
Figure 1. Nucleotide and deduced amino-acid sequences of amphioxus L34 gene (accession number in GenBank: AY168761).

The presumed translational start and termination sites are underlined, and the asterisk represents the stop codon. The potential polyadenylation signal upstream with respect to the poly(A) tail is boxed and the oligopyrimidine tract within the 5' UTR is double underlined. The putative nucleolar localization signal (RXXX) is indicated with dots under the letters.



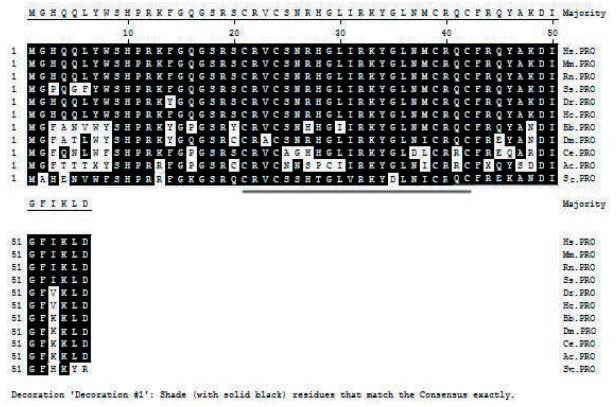
**Figure 2.** Amino-acid sequence alignment of representative L34s using the MegAlign program (DNASTAR) by CLUSTAL method.

Shaded residues are the amino acids that match the consensus. Gaps introduced into sequences to optimize the alignments are represented by (-). The putative nucleolar localization signals (RXXR) are underlined. See Table 1 for sequence reference.



**Figure 3.** Nucleotide and deduced amino-acid sequences of amphioxus S29 gene (accession number in GenBank: AY264807).

The presumed translational start and termination sites are underlined, and the asterisk represents the stop codon. The potential polyadenylation signal upstream with respect to the poly(A) tail is boxed and the oligopyrimidine tract within the 5' UTR is double underlined and the zinc finger-like domain is marked by heavy bar.



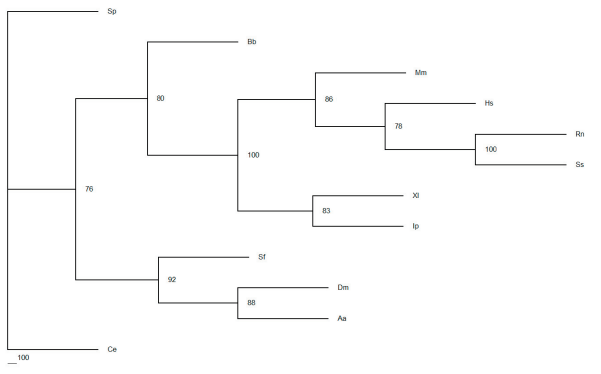
**Figure 4.** Amino-acid sequence alignment of representative S29s using the MegAlign program (DNASTAR) by CLUSTAL method.

Shaded residues are the amino acids that match the consensus. The zinc finger-like domains are underlined. See Table 2 for sequence reference.

Comparison of the deduced AmphiS29 amino-acid sequence with that of its counterparts in the GenBank database (Fig. 4) showed high identity to S29 from mammals (76.8%) including human, mouse, rat, pig and zebrafish (78.6%), seahorse (76.8%), fruit fly (69.8%), nematode (75%), sea hare (71.4%) and yeast (66.1%). Like other known S29 proteins, AmphiS29 has a zinc finger-like domain, -C-X<sub>2</sub>-C-X<sub>14</sub>-C-X<sub>2</sub>-C-, at position 21-42 (Fig. 4). Zinc fingers are small DNA-binding peptide motifs that have the potential to coordinate a zinc ion and bind to nucleic acids, mostly to DNA, and a few to RNA (Klug & Rhodes, 1987). AmphiS29 also has an excess of basic residues over acidic ones (10:2), which is close to the proportion of basic residues *versus* acidic ones in the amino-acid sequence of S29 from human, mouse, rat, pig, zebrafish and seahorse (11:2), fruitfly (10:3), nematode (12:4), sea hare (10:3) and yeast (11:4).

Two phylogenetic trees were constructed based on the conserved sequences of AmphiL34 and AmphiS29 and their counterparts from more than 10 representative species including invertebrates and vertebrates. Both trees demonstrated that the relationship among different species well reflected the established phylogeny of the chosen organisms and that the positions of AmphiL34/AmphiS29 were intermediate between the vertebrate and invertebrate L34/S29 (Figs. 5, 6). These result are in line with the notion that amphioxus is an organism transitional from invertebrates to vertebrates in phylogeny (Stokes & Holland, 1998; Zhang *et al.*, 2001).

The visible homology of AmphiL34 and AmphiS29 to their known counterparts extends the range of species in which these proteins are highly conserved, and the high conservatism of L34 and S29 amino-acid sequences from human to yeast sug-

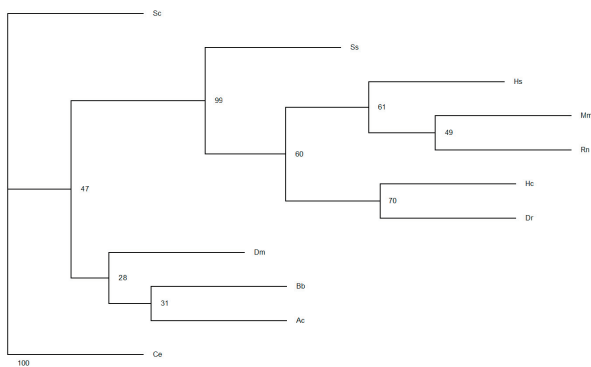


**Figure 5.** Phylogenetic tree constructed from different amino-acid sequences of L34 by the neighbor-joining method within the Philip 3.5c software package using 1000 bootstrap replicates.

The numbers at the branches are bootstrap percentages supporting the given branching pattern. L34 of yeast *Schizosaccharomyces pombe* was used as the outgroup. See Table 1 for sequence reference.

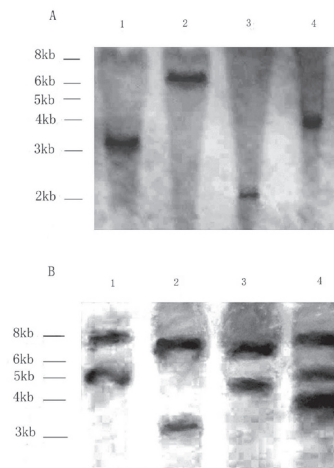
gests that they have been subjected to a strong selective pressure during evolution.

To analyze the copy number of the *AmphiL34* and *AmphiS29* genes, DIG-labeled cDNA probes of *AmphiL34* and *AmphiS29* were used to hybridize with digests of amphioxus genomic DNA with *Hind*III, *Bst*XI, *Eco*RV and *Bgl*II or *Eco*RI, *Pst*I, *Eco*RV and *Hind*III. The enzymes used do not digest *AmphiL34* or *AmphiS29* cDNA sequences, respectively. For *AmphiL34* there is a single hybridization band for each of the enzymes *Hind*III, *Bst*XI, *Eco*RV and *Bgl*II (Fig. 7A), while for *AmphiS29* there are two hybridization bands for each of the enzymes *Eco*RI, *Pst*I and *Eco*RV, and three hybridization bands for the enzyme *Hind*III (Fig. 7B). Considering the possibility that *AmphiS29* contains introns which can be digested by these restriction enzymes,



**Figure 6.** Phylogenetic tree constructed from different amino-acid sequences of S29 by the neighbor-joining method within the Philip 3.5c software package using 1000 bootstrap replicates.

The numbers at the branches are bootstrap percentages supporting the given branching pattern. S29 of yeast *Saccharomyces cerevisiae* was used as the outgroup. See Table 2 for sequence reference.



**Figure 7.**

**A.** Southern blotting analysis of *B. belcheri tsingtauense* genomic DNA for L34.

Lane 1 *Hind*III Lane 2 *Bst*XI Lane 3 *Eco*RV Lane 4 *Bgl*II.

**B.** Southern blotting analysis of *B. belcheri tsingtauense* genomic DNA for S29.

Lane 1, *Eco*RI; Lane 2, *Pst*I; Lane 3, *Eco*RV; Lane 4, *Hind*III.

one can infer the presence of one copy of the L34 gene and not more than 2–3 copies of the S29 gene in the genome of *B. belcheri tsingtauense*. It is of great interest to note that 7–9 copies of the L34 gene and 14–17 copies of the S29 gene are present in the rat genome (Aoyama *et al.*, 1989; Chan *et al.*, 1993). It is clear from the comparison of the number of the *AmphiL34* and *AmphiS29* genes with that of the rat L34 and S29 genes that duplication of the L34 and S29 genes occurred in vertebrates like rat. It therefore appears that housekeeping protein genes, such as L34 and S29, also underwent large-scale duplication during early chordate evolution, reinforcing the large-scale gene duplication proposal (Holland *et al.*, 1994; Sharman & Holland, 1996; Sidow, 1996).

#### Acknowledgements

This work was supported in part by a grant (30470203) from the Natural Science Foundation of China to S.C. Zhang.

#### REFERENCE

- Altschul SF, Madden TL, Schaffer AA, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**: 3389–3402.
- Amaldi F, Pierandrei-Amaldi P (1990) Translational regulation of the expression of ribosomal protein genes in *Xenopus laevis*. *Enzyme* **44**: 93–105.
- Aoyama Y, Chan YL, Wool IG (1989) The primary structure of rat ribosomal protein L34. *FEBS Lett* **249**: 119–122.
- Burland TG (2000) DNASTAR's Lasergene sequence analysis software. *Methods Mol Biol* **132**: 71–91.

- Chan Yuen-Ling, Suzuki K, Olvera J, Wool IG (1993) Zinc finger-like motifs in rat ribosomal proteins S27 and S29. *Nucleic Acids Res* **21**: 649–655.
- Dudov KP, Perry RP (1984) The gene family encoding the mouse ribosomal protein L32 contains a uniquely expressed intron-containing gene and an unmutated processed gene. *Cell* **37**: 457–468.
- Felsenstein J (1993) *PHYLIP (Phylogeny Inference Package)*. Department of Genetics, University of Washington, Seattle.
- Frigerio JM, Dagorn JC, Iovanna JL (1995) Cloning, sequencing and expression of the L5, L21, L27a, L28, S5, S9, S10 and S29 human ribosomal protein mRNAs. *Biochim Biophys Acta* **1262**: 64–68.
- Holland PW, Garcia-Fernandez J (1996) Hox genes and chordate evolution. *Dev Biol* **173**: 382–395.
- Holland PW, Garcia-Fernandez J, Williams NA, Sidow A (1994) Gene duplication and the origin of vertebrate development. *Dev Suppl* 125–133.
- Karsi A, Patterson A, Feng J, Liu Z (2002) Translational machinery of channel catfish: I. A transcriptomic approach to the analysis of 32 40S ribosomal protein genes and their expression. *Gene* **291**: 177–186.
- Klug A, Rhodes D (1987) Zinc fingers: a novel protein fold for nucleic acid recognition. *Cold Spring Harb Symp Quant Biol* **52**: 473–482.
- Levy S, Avni D, Hariharan N, Perry RP, Meyuhas O (1991) Oligopyrimidine tract at the 5' end of mammalian ribosomal protein mRNAs is required for their translational control. *Proc Natl Acad Sci USA* **88**: 3319–3323.
- Liu Z, Zhang S, Yuan J, Sawant MS, Wei J, Xu A (2002) Molecular cloning and phylogenetic analysis of AmphUbf80, a new member of ubiquitin family from the amphioxus *Branchiostoma belcheri tsingtauense*. *Curr Sci* **83**: 50–53.
- Liu M, Zhang S, Liu Z, Yuan J, Xu A (2004a) Identification of the ribosomal proteins S20 and L10 from the amphioxus *Branchiostoma belcheri tsingtauense*. *Indian J Marine Sci* **33**: 231–237.
- Liu Z, Zhang S, Liu M, Xu A (2004b) Identification of the ribosomal proteins S15a and L19 from the amphioxus *Branchiostoma belcheri tsingtauense*. *OPHELIA* **58**: 23–27.
- Meyer A, Malaga-Trillo E (1999) Vertebrate genomics: more fishy tales about Hox genes. *Curr Biol* **9**: R210–R213.
- Meyer A, Scharl M (1999) Gene and genome duplications in vertebrates: the one-to-four (-to-eight in fish) rule and the evolution of novel gene functions. *Curr Opin Cell Biol* **11**: 699–704.
- Moccia R, Chen D, Lyles V, Kapuya E, Yaping E, Kalachikov S, Spahn CM, Frank J, Kandel ER, Barad M, Martin KC (2003) An unbiased cDNA library prepared from isolated *Aplysia* sensory neuron processes is enriched for cytoskeletal and translational mRNAs. *J Neurosci* **23**: 9409–9417.
- Niu LL, Fallon AM (1999) The ribosomal protein L34 gene from the mosquito, *Aedes albopictus*: exon-intron organization, copy number, and potential regulatory elements. *Insect Biochem Mol Biol* **29**: 1105–1117.
- Ohno S (1970) *Evolution by Gene Duplication*. Springer-Verlag, New York, Heidelberg, Berlin.
- Perry RP, Meyuhas O (1990) Translational control of ribosomal protein production in mammalian cells. *Enzyme* **44**: 83–92.
- Proudfoot NJ, Brownlee GG (1976) 3' Non-coding region sequences in eukaryotic messenger RNA. *Nature* **263**: 211–214.
- Quaye IK, Toku S, Tanaka T (1996) Sequence requirement for nucleolar localization of the rat ribosomal protein L31. *Eur J Cell Biol* **69**: 151–155.
- Ramakrishnan V (2002) Ribosome structure and the mechanism of translation. *Cell* **108**: 557–572.
- Rehm BH (2001) Bioinformatic tools for DNA/protein sequence analysis, functional assignment of genes and protein classification. *Appl Microbiol Biotechnol* **57**: 579–592.
- Sankoff D (2001) Gene and genome duplication. *Curr Opin Genet Dev* **11**: 681–684.
- Sharman AC, Holland PWH (1996) Conservation, duplication and divergence of developmental genes during chordate evolution. *Netherlands J Zool* **46**: 47–67.
- Sharman AC, Shimeld SM, Holland PW (1999) An amphioxus *Msx* gene expressed predominantly in the dorsal neural tube. *Dev Genes Evol* **209**: 260–263.
- Shimeld SM (1999) The evolution of the hedgehog gene family in chordates: insights from amphioxus hedgehog. *Dev Genes Evol* **209**: 40–47.
- Sidow A (1996) Gen(om)e duplications in the evolution of early vertebrate. *Curr Opin Genet Dev* **6**: 715–722.
- Stokes MD, Holland ND (1998) The lancelet: also known as amphioxus, this curious creature has returned to the limelight as a player in the phylogenetic history of the vertebrates. *Am Sci* **86**: 552–560.
- Strausberg RL, Feingold EA, Grause LH, Derge JG, Klausner RD, Collins FS, Wagner L, Shennan CM *et al.* (2002) Generation and initial analysis of more than 15000 full-length human and mouse cDNA sequence. *Proc Natl Acad Sci USA* **99**: 16899–16903.
- Sugawara A, Shiga K, Takasawa S, Yonekura H, Yamamoto H, Okamoto H (1992) Sequence of the chicken rig gene encoding ribosomal protein S15. *Gene* **108**: 313–314.
- Ulbrich N, Lin A, Wool IG (1979) Identification by affinity chromatography of the eukaryotic ribosomal proteins that bind to 5.8S ribosomal ribonucleic acid. *J Biol Chem* **254**: 8641–8615.
- Vaccaro MC (2003) Sequencing and characterization of the *Xenopus laevis* ribosomal protein L34 cDNA. *Gene* **318**: 163–167.
- Wiedemann LM, Perry RP (1984) Characterization of the expressed gene and several processed pseudogenes for the mouse ribosomal protein L30 gene family. *Mol Cell Biol* **4**: 2518–2528.
- Williams NA, Holland PW (1998) Molecular evolution of the brain of chordates. *Brain Behav Evol* **52**: 177–185.
- Wintero AK, Fredholm M, Davies W (1996) Evaluation and characterization of a porcine small intestine cDNA library: analysis of 839 clones. *Mamm Genome* **7**: 509–517.
- Wood V, Gwilliam R, Stewart A, Sgouros J, Peat N, Hayles J, Baker S, Basham D (2002) The genome sequence of *Schizosaccharomyces pombe*. *Nature* **415**: 871–880.
- Wool IG (1979) The structure and function of eukaryotic ribosomes. *Annu Rev Biochem* **48**: 719–754.
- Zhang SC, Yuan JD, Li HY (2001) Amphioxus—model animal for insights into the origin and evolution of the vertebrates. *Chin Bull Life Sci* **13**: 214–218 (in Chinese with English abstract).
- Zhang SC, Yuan JD, Liu ZH (2002) Gene duplication and origin of the vertebrates. *Chin Bull Life Sci* **15**: 18–20 (in Chinese with English abstract).