Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biochemical and Biophysical Research Communications 314 (2004) 249–258
BBRC
www.elsevier.com/locate/ybbrc

Identification of 5′-upstream region of pufferfish ribosomal protein L29 gene as a strong constitutive promoter to drive GFP expression in zebrafish

Ming-Huang Chang,[a,b] Chih-Ming Chou,[c] Yueh-Chun Hsieh,[d] I-Ching Lu,[d] M. Kothai Nachiar Devi,[d] Juei-Pei Chang,[d] Tzong-Fu Kuo,[a,*] and Chang-Jen Huang[d,*]

[a] *Graduate Institute of Veterinary Medicine, National Taiwan University, Taipei, Taiwan, ROC*
[b] *Department of Veterinary Medicine, National Chia-Yi University, Chia-Yi, Taiwan, ROC*
[c] *Department of Biochemistry, Taipei Medical University, Taipei, Taiwan, ROC*
[d] *Institute of Biological Chemistry, Academia Sinica, Taipei, Taiwan, ROC*

Received 25 November 2003

Abstract

The genomic structure of *Tetraodon fluviatilis L*29 gene was determined and its promoter activity was analyzed in COS-1 cells and zebrafish embryos. The *TfL*29 gene comprises four exons and three introns, spanning approximately 1.7 kb. The 5′-upstream 2.2-kb of the first exon contains 10 E-boxes and many putative binding motifs for transcription factors GATA-1, AML-1a, c-Myb, Oct-1, CdxA, and NRF-2. Promoter activity assay showed that the distal 2.2-kb fragment not only had high luciferase activity in COS-1 cells, but also strong and ubiquitous GFP expression in a variety of tissues in zebrafish embryos. On the other hand, there are no TATA or CAAT boxes within a 300-bp region upstream from the transcription initiation site. Although this region has high luciferase activity in COS-1 cells, it is not sufficient to drive GFP expression in zebrafish embryos. In this proximal 300-bp region, there are two E-boxes, two CdxA sites, and one NRF-2 site that is immediately downstream of the transcription start site.

The ribosome is responsible for protein synthesis in all organisms. The mammalian and catfish ribosomes consist of four RNA molecules and 79 protein subunits. The large 60S ribosome consists of three RNA molecules and 47 protein subunits while the 40S ribosome one 18S rRNA and 32 protein subunits [1–3]. In mammal, the ribosomal protein L29 (RPL29) is identical to a heparin interacting protein (HIP), which is a small and highly basic protein that can bind to heparin/heparan sulfate [4,5]. Interestingly, the catfish RPL29 is homologous to mammalian RPL29/HIP, but it contains only 64 amino acids as compared to 159 amino acids in human [3]. The catfish RPL29 lacks the C-terminal 95 amino acid residues of human RPL29/HIP. In addition to ribosomal L29, other ribosomal proteins such as RPS3, RPS14, RPL4, RPL7 [6], and RPL6 [7] also have extraribosomal functions.

Most ribosomal protein genes are housekeeping genes with a strong constitutive expression. Analysis on genomic structures of ribosomal protein genes has been reported in higher eukaryotes, such as human [8] and *Xenopus* [9]. Studies on promoter activity of ribosomal protein genes have been reported in various cell types [7,10], but only a few in both cells and zebrafish embryos [11]. Here we reported the molecular cloning of a pufferfish (*Tetraodon fluviatilis*) ribosomal protein *L*29 gene (*TfL*29) and performed the promoter assay in COS-1 cells and zebrafish embryos. The round-spotted pufferfish, *T. fluviatilis*, has a genome size of 380 Mb which is eight times smaller than that of human [12,13]. Another pufferfish *Takifugu rubripes* (fugu) also has a compact genome and is used as a model for comparative genome analysis of vertebrates [13,14]. The complete genomic sequences of fugu have been determined in 2002 [15].

* Corresponding authors. Fax: +886-2-2788-9759 (C-J. Huang).
E-mail address: cjibc@gate.sinica.edu.tw (C.-J. Huang).

0006-291X/$ - see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2003.12.080

Table 1
Intron-exon organization of the *TfL29* gene

Exon No.	Exon size (bp)	3′ end of the exon	5′ end of the intron	Intron size (bp)	3′ end of the intron	5′ end of the next exon
1	32	GTG TGG CAG	GTagtcaaat	445	cctcctgcAG	AA ATG GCC AA
2	39	CAC AAC CAG T	GTaggtgtaa	198	tcctaaacAG	CT CGT AAG GC
3	65	CTG AAA GGG	GTatgtaaag	702	tttattgcAG	GTT GAC TCC A
4	182	tatgctttgtttgacaaaaaaataaa				

Comparison of fugu and human genomic sequences identified several conserved regulatory sequences in the promoter and intronic regions [16]. These regulatory elements have been reported to be functional in mammalian cell lines and transgenic mice [17].

We have previously determined the promoter activity of all four *T. fluviatilis* JAK kinase genes in zebrafish embryos [18]. In this study, we analyzed the expression of *TfL29* gene and performed the promoter assay in COS-1 cells and zebrafish embryos. We found that a distal 2.2-kb region had high promoter activity in both cells and embryos and a proximal 300-bp region had high reporter activity only in COS-1 cells, but not in zebrafish embryos (see Table 1).

Materials and methods

Materials. All the restriction enzymes were purchased from the Promega Biosciences (Madison, WI, USA) or New England Biolabs (Beverly, MA, USA). Chemicals were purchased from Merck (Darmstadt, Germany) or Sigma (MO, USA).

Cell cultures. Monkey kidney fibroblast COS-1 cells were cultured in Dulbecco's modified Eagle's medium, supplemented with 10% fetal bovine serum (FBS, HyClone, Utah), penicillin G (50 U/ml), streptomycin (50 μg/ml), and L-glutamine (2 mM) in a humidified atmosphere of 5% CO_2 at 37 °C. Various culture reagents used were purchased from HyClone (Logan, Utah).

Fish. Zebrafish (*Danio rerio*) were maintained at 28 °C on a 14 h-light/10 h-dark cycle. Embryos were incubated at 28 °C and different developmental stages were determined according to the description in Zebrafish Book [19].

Isolation of the round-spotted pufferfish genomic DNA. The genomic DNA of the round-spotted pufferfish (*T. fluviatilis*) was prepared as previously reported [18,20]. In brief, genomic DNA was prepared from the liver with DNAzol reagent (Life Technologies, Gaithersburg, MD), precipitated by addition of 100% ethanol, and then removed by spooling onto a pipette tip. DNA was dissolved in TE buffer (10 mM Tris–HCl, 1 mM EDTA, pH 8.0).

Total RNA isolation and first-stranded cDNA synthesis. Total RNA was isolated from various tissues (brain, gill, heart, intestine, kidney, muscle, and ovary) of *T. fluviatilis* using the RNAzol reagent (Tel-Test) according to the instructions of the manufacturer. After treatment with RQ1 RNase-Free DNaseI (Promega Biosciences, WI), 50–100 μg of total RNA from each tissue was used for the first strand cDNA synthesis in a 25 μl reaction mixture containing 10 pmol of oligo(dT) primer and 100 ng random primer (Promega), 30 U RNasin (Promega), 1 mM dNTP, 10 mM dithiothreitol, and 300 U Superscript II RT (Invitrogen Life technologies, CA). The reaction mixture was incubated at 42 °C for 1 h. Two microliters of the cDNA products was used for subsequent PCR amplification.

Isolation of the full-length L29 cDNA from pufferfish (T. fluviatilis). In order to isolate the cDNA covering the complete open reading frame (ORF) of *T. fluviatilis* L29, according to the sequences of one fugu EST (Accession No. CA846627), PCR amplification was performed in a 50 μl reaction mixture containing 2 μl first strand cDNA, 0.5 μg forward primer [L29F1, 5′-ATG GCN AA(A/G) TCN AA(A/G) AA(T/C) CA(T/C) AC-3′] and reverse primer [L29R1, 5′-TT(T/C) TT(A/G) TT(A/G) TG(T/C) TT(T/C) TTN GC(A/G) AA-3′], 1.5 mM $MgCl_2$, 0.2 mM dNTP, and 3 U ExTaq (Takara Shuzo, Shiga, Japan). The samples were incubated in a thermal cycler (Hybaid MultiBlock System, Hybaid Limited, MA) at 96 °C for 3 min; 30–45 cycles of 96 °C for 30 s, 50 °C for 30 s, and 72 °C for 30 s; and the final extension at 72 °C for 15 min. All PCR products were ligated into the pGEM-T easy vector (Promega Biosciences, Medison, WI) and individual clone was subjected to sequencing analysis.

The 5′- and the 3′-ends of *T. fluviatilis L29* mRNA were obtained by the RACE PCR technique using the Marathon cDNA amplification kit (Clontech Lab, CA). The adaptor-ligated double-strand cDNA as template was used for PCR amplification using the program of 95 °C for 5 min; 40 cycles of 95 °C for 30 s, 65 °C for 30 s, and 72 °C for 45 s; and the final extension at 72 °C for 5 min. The 5′-RACE was performed using an adaptor-specific sense primer, AP1: 5′-CTA ATA CGA CTC ACT ATA GGG C-3′ and a L29 antisense primer, L29R5: 5′-CTA TTT CGC CTG AAC AGC CTC CCG TG -3′. The PCR product was reamplified with AP2 primer: 5′-ACT CAC TAT AGG GCT CGA GCG GC-3′ and a nested L29 antisense primer, L29R4: 5′-TAG CGA TGA GAC CTG GGC CTC TTG-3′. For 3′-RACE, cDNA was amplified with a L29 sense primer, L29F2: 5′-ATG GCC AAG TCC AAG AAC CAC AC AAC-3′ and an oligo(dT) primer. All RACE products were cloned into pGEM-T easy vector (Promega) and subjected to sequence analysis.

Cloning of the promoter region of pufferfish L29 gene. The DNA sequences of the 5′-upstream 2.2-kb DNA fragment before the first 5′-UTR of the TfL29 mRNA were obtained by searching the *Tetraodon nigroviridis* database using BLAST program. Two specific primers were designed (TfL29F1: 5′-GGG ATT GGA TCC TCG GGT CAT CAT TAC TCA GC-3′; TfL29R1: 5′-CGT TTC ATT GAC ACT CCT GTG TTC AAT ACC GTC-3′) and used to perform PCR amplification by using *T. fluviatilis* genomic DNA as template. The PCR product was 2.6 kb containing the first intron. The PCR product was cloned into pGEM-T easy vector (Promega, Madison, WT) and sequenced.

RT-PCR analysis of T. fluviatilis L29 mRNA. PCR amplification was performed in a 50 μl reaction mixture containing 200 ng L29 primers (L29F2, 5′-ATG GCC AAG TCC AAG AAC CAC AC-3′ and L29R2, 5′-CTA TTT CGC CTG AAC AGC CTC CCG TG-3′) or β-actin primers (ActF, 5′-CCT CCG GTC GTA CCA CTG GTA T-3′ and ActR, 5′-CAA CGG AAG GTC TCA TTG CCG ATC GTG-3′), 1.5 mM $MgCl_2$, 0.2 mM dNTP, and 2.5 U ExTaq (Takara Shuzo, Shiga, Japan). The samples were incubated in a thermal cycler (Hybaid MultiBlock System, Hybaid Limited, MA) at 96 °C for 3 min; 30 to 45 cycles of 96 °C for 30 s, 50 °C for 30 s, and 72 °C for 30 s; and the final extension at 72 °C for 5 min. A negative control was performed in the absence of first strand cDNA. All PCR products were separated on a 1% agarose gel.

DNA sequence analysis. DNA sequence analysis was performed by using PRISM Ready Reaction Big-Dye Termination Cycle sequencing Kit (Applied Biosystems, CA) on an Applied Biosystems 310 automated DNA sequencer. Sequence analysis was performed by using the Clustal X, PSORT II (http://psort.ims.u-tokyo.ac.jp/), TMAP (http://www.mbb.ki.se/tmap/), and GenDoc programs.

Construction of reporter plasmids. In order to assess the promoter activity in COS-1 cells, different sizes of the promoter region of the *TfL*29 gene without the first intron in the 5′-UTR were cloned into pGL3-basic vector. The 5′-upstream 2.2-kb promoter region of the *TfL*29 gene was cloned into polylinker regions of the reporter vector pGL3-Basic (Promega) using PCR primers 5′-ATC GAG CTC GGG ATT GGA TCC TCG GGT CAT CAT TAC-3′ (nucleotides −2196 to −2170) and (5′-CCG CTC GAG CTG CCA CAC GAA CAG ACG ACA AC-3′; nucleotides +13 to +35). The oligonucleotide primers have additional sequences of *Sac*I and *Xho*I site at the 5′-end, respectively. The PCR product was purified from the gel, digested with *Sac*I and *Xho*I, and then cloned into the same sites of pGL3-Basic. Thus, the pTfL29-2.2k contained the flanking region from nucleotides −2196 to +35. Similarly, pTfL29-0.7k and pTfL29-0.3k were constructed by PCR amplification using different forward primers (5′-ATC GAG CTC GGG GTC TCT GTG GGC TCT TCT GTG G-3′; nucleotides −639 to −615) and (5′-ATC GAG CTC ACA ACA TTA TTG ATA TTT ATT AAT C-3′; nucleotides −254 to −230). Through PCR and subcloning, the reporter plasmids were obtained and designated pTfL29-2.2k, pTfL29-0.7k, and pTfL29-0.3k.

For the construction of GFP reporter plasmids, similar strategy was used as described above for luciferase reporter constructs except using primers carrying any another restriction sites (*Hin*dIII and *Kpn*I restriction sites). PCR product was digested with *Hin*dIII and *Kpn*I, and then cloned into the same sites of pEGFP-1 (Clontech).

Transactivation assay. For transactivation assay, transfections were performed in six-well plates. A 1 μg of reporter plasmid was transfected using Lipofectamin/Plus kit (Life Technologies) into COS-1 cells. Transfected cells were harvested at 48 h after infection and then assayed for luciferase activity using a dual luciferase assay kit FireLite purchasing from Packard (Groninige, BK, Netherland) according to manufacturer's instructions. Final luciferase activity was obtained after normalization with *Renilla* luciferase and blank pGL3-basic vector.

Microinjection of reporter constructs into zebrafish embryos. The reporter plasmids were linearized by digestion with restriction enzyme *Sca*I purified with Gene-Spin 1-4-3 DNA extraction kit (Protech technology enterprise, Taiwan). DNA concentration was adjusted to 100 μg/ml in 0.1 M KCl solution containing 0.5% phenol red and 100–200 pl was microinjected into the zebrafish embryo at one-cell stage by using Narishige IM 300 microinjector system (Narishigi Scientitific Instrument Lab., Tokyo, Japan). Embryos at 48 and 72 h postfertilization were observed under an Olympus IX70-FLA inverted fluorescence microscope. Images were taken by using the SPOT system (Diagnostic Instruments, Sterling Heights, Michigan) and assembled by PhotoShop program (Adobe System, California).

Results

Isolation of T. fluviatilis L29 cDNA

To isolate *TfL29* cDNA, we used the coding region of the fugu L29 (Accession No. CA846627) to search the GenBank database for related expression sequence tag (EST) sequences by using the program tBLAST. No corresponding EST clone was found in *T. nigroviridis*, which is identical to *T. fluviatilis*. Therefore, we used the degenerated primers to amplify the DNA fragment encoding the highly conserved region (amino acid residues 1–52) and then used 5′- and 3′-RACE to obtain the 5′- and 3′-untranslated regions (UTR). The assembled full-length *TfL29* cDNA consists of 303 bp containing an open reading frame of 207 bp encoding a protein of 68 amino acid residues. The *TfL29* cDNA sequence was deposited in GenBank with an accession number of AY466493. The amino acid sequence comparison indicates that Tetraodon and Fugu L29 proteins are highly homologous with 87% identity whereas TfL29 is also highly homologous to other fish L29 proteins such as catfish L29 (76% identity). Like catfish L29 [3], the TfL29 is also homologous to human L29 protein in the N-terminal region and lacks the C-terminal 91 amino acid residues (Fig. 1).

Genomic structure of T. fluviatilis L29 gene

We have used the 303 bp of *TfL*29 cDNA (Accession No. AY466493) as query to search the GenBank nonredundant database using BLAST program. The *TfL*29 cDNA matched to 4 non-contiguous regions in a *T. nigroviridis* BAC clone CH211-278F21 (GenBank accession no. BX088688). Subsequently, alignment of the sequence of BAC CH211-278F21 to that of *TfL*29 cDNA by BLAST 2 program indicated that *TfL29* cDNA is contained within four putative exons (Fig. 2). Using these putative exons as a model, a sequence alignment was produced such that each intron concurred with the conscence GT/AG intron donor/acceptor site rule [21]. Exon 1 contains 32 bp of 5′-UTR and exon 2 the next 2 bp of the 5′-UTR and the first 37 bp of the coding sequences of the *TfL*29 cDNA. Exon 3 contains the next 65 bp of the coding sequences while exon 4 contains the last 105 bp of the coding sequences and 77 bp of 3′-UTR. The size of introns varied considerably, ranging from 198 bp (intron 2) to 702 bp (intron 3) with an average size of 448 bp. The *TfL*29 gene spans approximately 1.7 kb (Fig. 2).

Expression profiles of T. fluviatilis L29 mRNA

RT-PCR was performed to determine the expression pattern of *T. fluviatilis L29* transcripts in various tissues. Our results showed that *Tf*L29 is ubiquitously expressed in all tissues examined, which is consistent with that most ribosomal protein genes are housekeeping genes.

Characteristics of the 5′-flanking regions of TfL29gene

In order to identify sequence elements involved in the control of the pufferfish *Tf*L29 gene expression, a DNA fragment corresponding to the 5′-upstream region of the *Tf*L29 gene was sequenced. Analysis of these sequences revealed numerous putative binding sites for transcription factors. As shown in Fig. 3, no TATA box, which is

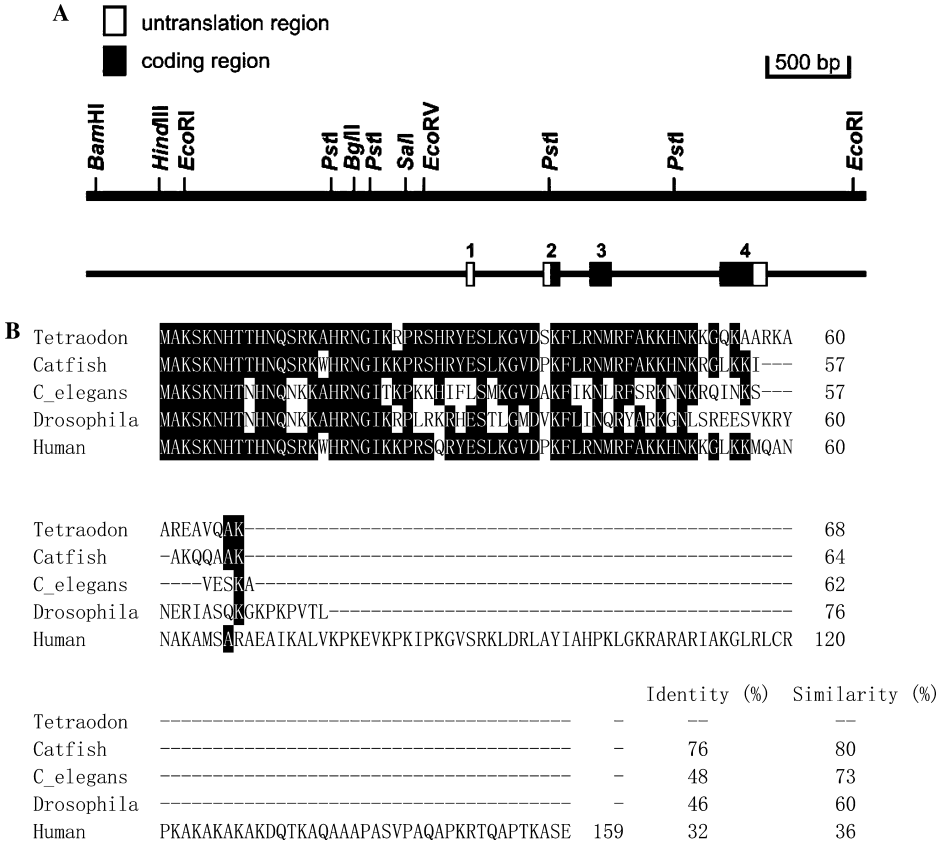

Fig. 1. Genomic organization of *T. fluviatilis L*29 gene and alignment of amino acid sequences of TfL29 with other ribosomal L29 proteins. (A) Genomic organization of *T. fluviatilis L*29 gene. Exons are indicated by boxes numbered from 1 to 4. The coding regions are shown as filled boxes whereas the 5′- and 3′-untranslated regions are shown as open boxes. Introns and the 5′- and 3′-flanking regions are indicated by solid lines. The restriction map was shown above the genomic structure. (B) Alignment of amino acid sequences of TfL29 with other ribosomal L29 proteins. Amino acid sequences of various ribosomal L29 proteins from *Tetraodon*, catfish, *C. elegans*, *Drosophila*, and human were aligned using CLUSTAL X program. Gaps are introduced to optimize alignment and shown as dashes. Sequences used are: catfish L29 (AAK95156), *C. elegans* L29 (NM_070270), *Drosophila* L29 (Q24154), and human L29 (P47914). Identical amino acids are shown in black boxes. Percentage identity and similarity of amino acid sequence between *Tetraodon* L29 and other L29 proteins are shown.

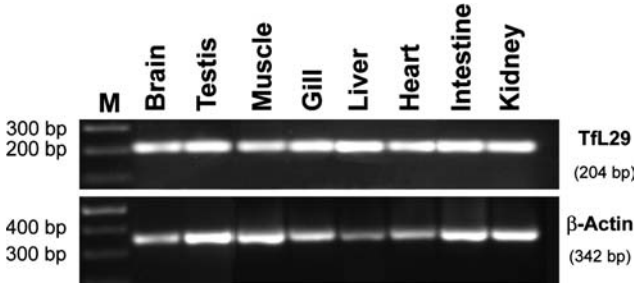

Fig. 2. Tissue distribution of *T. fluviatilis L*29 mRNA by RT-PCR. Total RNA (5–10 μg) from tissues of *T. fluviatilis* was subjected to RT-PCR analysis. The resulting PCR products were electrophoresed on 1.2% agarose gel containing ethidium bromide. A negative control was run simultaneously. A DNA fragment of 204 bp was amplified from different tissues using *TfL*29 specific primers. The intensity of 342 bp DNA fragment using β-actin specific primers amplified from *T. fluviatilis* tissues was used to evaluate the relative amount of cDNA used in each PCR.

generally located at a position about 30 bp upstream of the transcription start site [22], was identified. Seven putative binding sites for a ubiquitously expressed transcription factor, CdxA/cdx-1 [23], were found at −183, −237, −299, −348, −1392, −1501, and −1967. Ten E-box elements, known to be important for muscle-specific expression [24], were also found in this promoter region. In addition to the above motifs, the 5′-flanking region also contained potential DNA motifs for Oct-1 [25] at −697 and −922; GATA [26] at −1761; c-Myb [27] at −1212; and AML-1a [28] at −1444. Finally, a binding motif for a ubiquitously expressed transcription factor, NRF-2 [29], was found at +7 immediately downstream of the transcription start site (Fig. 3). Two TC-rich stretches are present near its transcription start site (underlined).

Functional analysis of the 5′-upstream region of the TfL29 gene in COS-1 cells

To determine whether the 5′-flanking region of the *TfL*29 gene exhibits promoter activity, different genomic DNA fragments containing the 5′-upstream region were fused to the luciferase reporter gene in pGL3-Basic

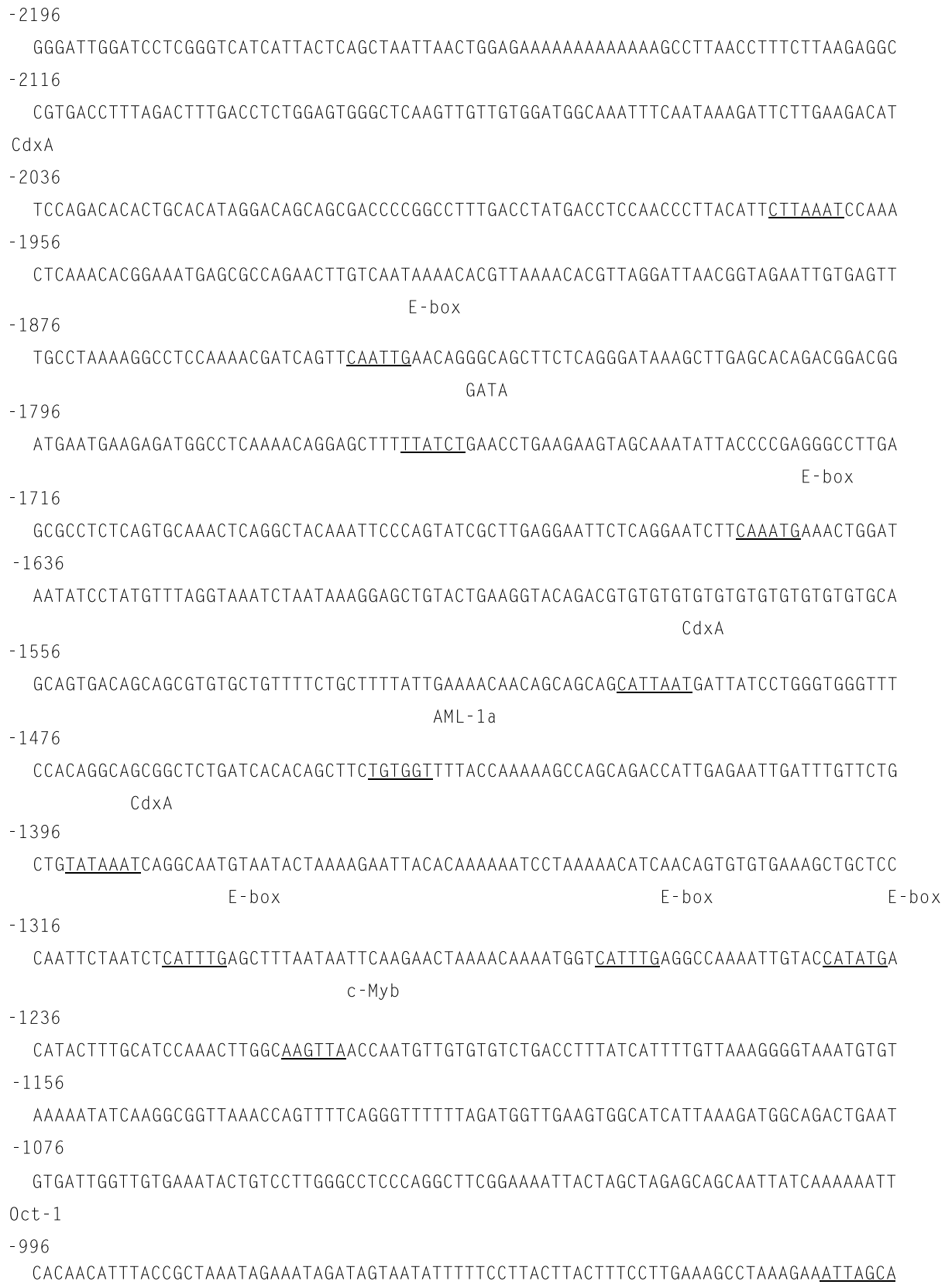

Fig. 3. Nucleotide sequence of the 5′-flanking region of the pufferfish *L29* gene. The putative transcription start site determined by 5′-RACE is indicated with nucleotide number (+1). Potential binding sites for a variety of transcription factors are also marked and underlined. Nucleotide sequences of the first intron that disrupts the 5′-untranslated region are also shown in small letters. The translation initiation codon ATG is shown in bold type. All sequences in this figure appear in GenBank with an Accession No. AY466376.

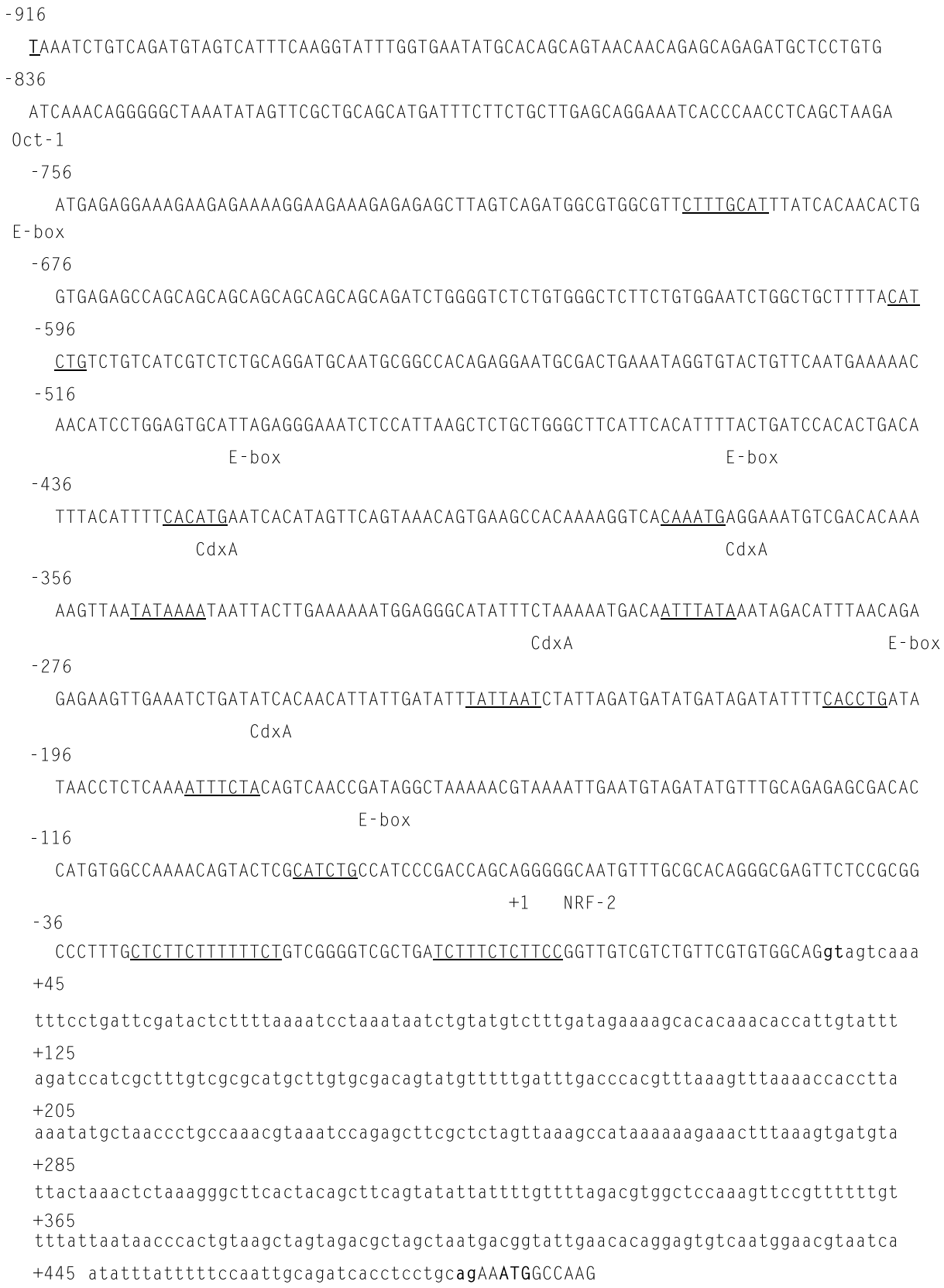

Fig. 3. (*continued*)

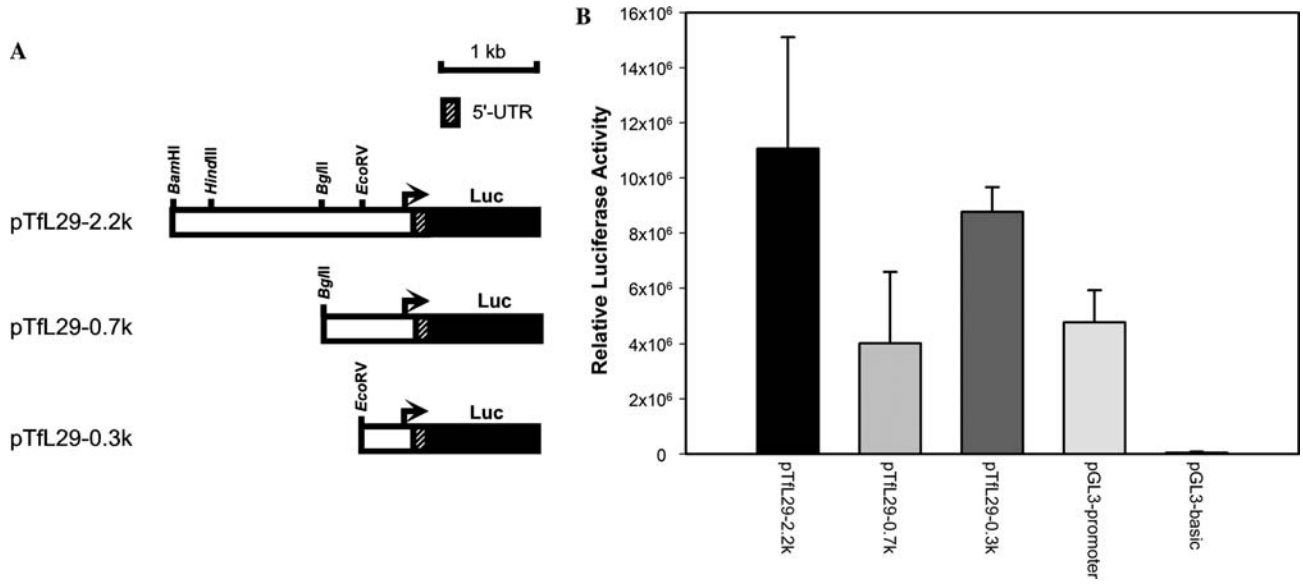

Fig. 4. Analysis of the promoter activity of various fragments of the 5′-flanking region of *TfL29* gene in COS-1 cells. COS-1 cells were cotransfected with one microgram of each reporter construct (A) and pSV-β-galactosidase. Cell lysates were prepared at 48 h after transfection. Luciferase activity (B) in an individual experiment was corrected for variation in transfection efficiency by normalizing the value to the β-galactosidase activity in the same extract. The data represent means of triplicate transfection experiments for each plasmid. pGL3-basic was used as negative control while pGL3-promoter as positive control.

vector as described in the Materials and methods. The luciferase activity was examined by transient transfection into COS-1 cells [30]. As shown in Fig. 4, the luciferase activities of pTfL29-2.2k-Luc, pTfL29-0.7k-Luc, and pTfL29-0.3k-Luc were about 2.6-, 0.8-, and 2-fold compared to that of SV40 basic promoter in pGL3-promoter. These data indicated that the proximal 0.3-kb region of the *TfL29* promoter contained strong promoter activity in COS-1 cells. In addition, negative regulatory sequences (−0.3 to −0.7 kb) and positive regulatory sequence (−0.7 to −2.2 kb) are also found.

Functional analysis of the 5′-upstream region of the TfL29 gene in zebrafish embryos

To investigate the promoter activity in vivo, a promoter assay system using zebrafish embryos, in a mosaic fashion, was employed to study transcriptional regulation of gene expression involved in development [31,32]. In this study, we used the GFP gene as the reporter gene. Three expression constructs such as pTfL29-2.2k-GFP, pTfL29-0.7k-GFP, and pTfL29-0.3k-GFP were constructed as described in the Materials and Methods and microinjected into zebrafish embryos at one-cell stage, respectively. Two representative expression patterns for each construct are shown in Fig. 5. It is interesting to note that pTfL29-2.2k-GFP displayed stronger GFP expression and wider tissue distribution (panels a and b) while pTfL29-0.7k showed moderate expression pattern (panels c and d). In contrast, pTfL29-0.3k-GFP displayed few GFP signals in a few regions (panels e and f). In pTfL29-2.2k-GFP-injected embryos, GFP fluorescence could be found in a variety of tissues, such as brain, muscle and epidermis, notochord, eye, heart, and yolk sac. A few circulating GFP blood cells were also found (data not shown). These data indicate that the distal 2.2-kb promoter of *TfL29* gene is a strong ubiquitous promoter in zebrafish embryo.

Discussion

In this study, we determined the genomic structure of *T. fluviatilis L29* gene and analyzed the promoter activity in COS-1 cells and zebrafish embryos. As shown in Figs. 4 and 5, we found that the proximal 300-bp promoter region is functional in COS-1 cells, but it is not sufficient for the expression of GFP reporter gene in zebrafish. On the other hand, the 2.2-kb 5′-flanking region contains many regulatory elements that are sufficient to express GFP gene with ubiquitous expression in zebrafish embryos. The proximal 300-bp region is devoid of a TATA box, but contains three CdxA/cdx-1 binding motifs and two E-box elements (Fig. 3). In addition to those motifs, another putative binding site for the transcription factor NRF-2, also known as GA-binding protein, was found immediately downstream of the transcription start site. This factor is ubiquitously expressed in all tissues [29]. The possible interaction of NRF-2 with CdxA/cdx-1 and myogenic transcription factor that binds to the E-box may affect the assembly of the transcriptional complex. In a 355-bp

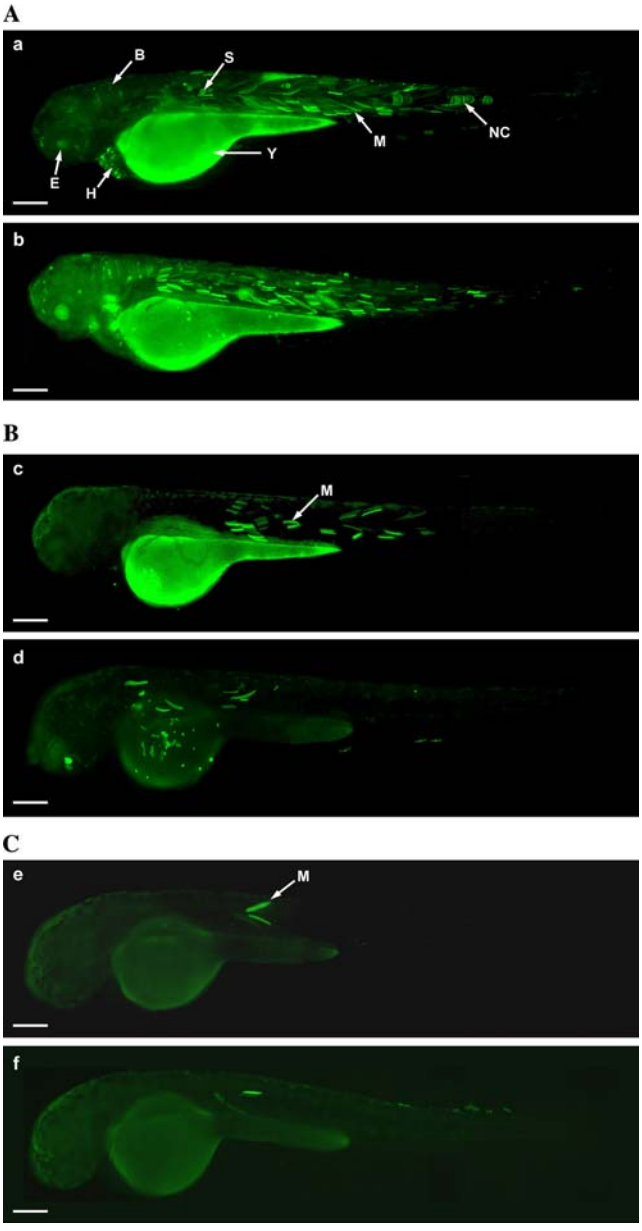

Fig. 5. Transient GFP expression of pTfL29-GFP constructs in zebrafish embryos. Various reporter constructs, such as pTfL29-2.2k-GFP (A), pTfL29-0.7k-GFP (B), and pTfL29-0.3k-GFP (C), were microinjected into zebrafish embryos at 1-cell stage. Zebrafish embryos at 48 h postfertilization with GFP fluorescence were selected for image analysis. Embryos are shown as lateral view with anterior to the left. Experiments were repeated at least three times and at least 100 fertilized eggs were used for injection of each construct at one time. Two representative patterns for each construct were shown. E, eye; Y, yolk ball and yolk extension; H, heart; B, brain; S, skin; M, muscle; and NC, notochord. Scale bars represent 100 μm.

promoter region of murine L29/HIP gene, there are many binding motifs for transcription factor including Sp1, HNF-3β, NF-κB, and NRF-2 [10]. Interestingly, it is common that the NRF-2 binding site is present in a minimal promoter region and required for the in vitro expression of several ribosomal protein genes, such as RPL32 [33], RPS16 [34], RPL27a [35], and the murine Hip/RPL29 [10].

The distal 2.2-kb promoter region of the pufferfish L29 gene had high luciferase activity in COS-1 cells and high GFP expression in zebrafish embryos. Analysis of this region reveals that many binding motifs for transcription factors such as GATA-1, c-Myb, AML-1a, Oct-1, CdxA, and E-box were found (Fig. 3). The former three transcription factors are known to be involved in hematopoiesis. GATA regulatory motifs were first identified in the promoters of globulin and other erythroid-specific genes [36]. The presence of GATA-1 binding site in the pufferfish *L29* gene promoter may suggest that GATA-1 transcription factor possibly regulates *L29* gene expression in hematopoietic cells. In addition to the GATA-1, the c-Myb is a transcriptional activator that plays an important regulatory role in cell proliferation and differentiation of hematopoietic cells [27]. Transcription factor AML-1 is known to be a key regulator of hematopoiesis and plays an important role in development of all hematopoietic lineages [28]. Oct-1 is one of the transcription factors with POU homeo domain and is expressed in most tissues including the brain [25]. Moreover, Oct-1 plays very important roles in the transcriptional regulation of gonadotropin-releasing hormone and aldolase C gene expression in the brain [37,38]. CdxA/cdx-1 is one of the vertebrate caudal proteins that play important roles in establishment of the body plan during early development [23]. CdxA and cdx-1 are homologous proteins and originate from chicken and mouse, respectively. As shown in Fig. 5, the visible GFP expression regions of pTfL29-2.2k-GFP-injected embryos are yolk sac, muscle, notochord, eye, heart, and brain. A few fluorescent blood cells are also observed (data not shown). Mutation or deletion analysis will be needed to confirm the relevance of GFP expression sites to the presence of potential regulatory elements in the 5′-flanking region of the pufferfish *L29* gene.

The consensus sequence of the E-box is CANNTG. This binding motif is recognized by myogenic regulatory factors (MRFs) [24], such as the MyoD, Myf5, myogenin, and MRF4 [39,40]. Multiple E-boxes have been found in the promoters of muscle-specific genes including creatine kinase, myosin light chain, and myogenin [41,42]. In this study, we found that there are 10 E-boxes present in the 5′-upstream 2.2-kb of the pufferfish *L29* gene. Among them, two sites are located in the proximal 300-bp region while five sites in the distal 960-bp region. As shown in Fig. 5, the 2.2-kb promoter region containing 10 E-boxes had stronger and wider GFP expression in the muscle whereas the 0.7-kb promoter region (five E-boxes) moderate GFP expression in the muscle. In contrast, the minimal 300-bp promoter region containing two E-boxes had low GFP expression in the muscle fiber. The relevance of promoter activity to

the presence of potential regulatory E-box elements in the 5′-flanking region of the pufferfish *L29* gene needs further mutation analysis.

Like the catfish L29, *T. fluviatilis* L29 is highly homologous to human ribosomal protein L29 at the N-terminal 53 amino acid residues, but lacks the C-terminal 91 amino acid residues of human L29/HIP (Fig. 1). Two potential heparin/heparin sulfate-binding motifs have been identified in human L29/HIP protein [43]. One is located at the N-terminus (aa 43-58: MRFAKKHNKKGLKKMQ) while the other at the C-terminus (aa 119-134: RPKAKAKAKAKDQTK). A similar sequence of MRFAKKHNKKGQKAAR was also found at the C-terminus of *T. fluviatilis* L29 protein. Further investigation is needed to determine whether the pufferfish L29 could bind to heparin/heparan sulfate.

In the early development of generation of transgenic GFP zebrafish, a non-fish promoter from *Xenopus elongation factor* 1α gene was used to drive GFP expression ubiquitously in zebrafish [44]. Recently, a fish promoter from medaka *elongation factor* 1α gene was also used to generate transgenic GFP zebrafish, in which GFP fluorescence was observed in a variety of tissues, such as kidney, liver, heart, gill, ovary, and testis, except for the skeletal muscle [45]. Moreover, other ubiquitous promoters from fish β-actin genes, such as zebrafish [46], medaka [47], and carp [48], were also used in the generation of transgenic zebrafish, in which GFP was expressed in many tissues such as the epidermis, blood vessels, muscle, notochord, fin ray, gut, eyes, and yolk sac [47]. In this study, we identified the 2.2-kb 5′-flanking region of *T. fluviatilis* ribosomal protein *L29* gene as a strong constitutive promoter that can drive GFP expression in a variety of tissues in zebrafish, such as epidermis, blood cells, heart, muscle, notochord, brain, eyes, and yolk sac. Thus, this promoter is useful for the generation of transgenic zebrafish with ubiquitous GFP expression.

Acknowledgments

We thank Dr. P. W. Hsiao for critically reading the manuscript and helpful discussion. This research was supported by grants from the National Science Council (NSC-91-3112-P-001-018-Y-3), Taiwan, Republic of China.

References

[1] I.G. Wool, Y.L. Chan, A. Gluck, Structure and evolution of mammalian ribosomal proteins, Biochem. Cell Biol. 73 (1995) 933–947.

[2] A. Karsi, A. Patterson, J. Feng, Z. Liu, Translational machinery of channel catfish: I. A transcriptomic approach to the analysis of 32 40S ribosomal protein genes and their expression, Gene 291 (2002) 177–186.

[3] A. Patterson, A. Karsi, J. Feng, Z. Liu, Translational machinery of channel catfish: II. Complementary DNA and expression of the complete set of 47 60S ribosomal proteins, Gene 305 (2003) 151–160.

[4] S. Liu, S.E. Smith, J. Julian, L.H. Rohde, N.J. Karin, D.D. Carson, cDNA cloning and expression of HIP, a novel cell surface heparin sulfate/heparin-binding protein of human uterine epithelial cells and cell lines, J. Biol. Chem. 271 (1996) 11817–11823.

[5] L.H. Rohde, J. Julian, A. Babaknia, D.D. Carson, Cell surface expression of HIP, a novel heparin/heparan sulfate binding protein, of human uterine epithelial cells and cell lines, J. Biol. Chem. 27 (1996) 11824–11830.

[6] I.G. Wool, Extraribosomal functions of ribosomal proteins, Trends Biochem. Sci. 21 (1996) 164–165.

[7] J. Wang, X. Yang, P. Zhou, H. Han, Cloning of mouse genomic ribosomal protein L6 gene and analysis of its promoter, Biochim. Biophys. Acta 1576 (2002) 219–224.

[8] M. Yoshihama, T. Uechi, S. Asakawa, K. Kawasaki, S. Kato, S. Higa, N. Maeda, S. Minoshima, T. Tanaka, N. Shimizu, N. Kenmochi, The human ribosomal protein genes: sequencing and comparative analysis of 73 genes, Genome Res. 12 (2002) 379–390.

[9] F. Amaldi, O. Camacho-Vanegas, B. Cardinall, F. Cecconi, C. Crosio, F. Loreni, P. Mariottini, L. Pellizzoni, P. Pierandrei-Amaldi, Structure and expression of ribosomal protein genes in *Xenopus laevis*, Biochem. Cell Biol. 73 (1995) 969–977.

[10] C.B. Kirn-Safran, S. Dayal, P.A. Martin-DeLeon, D.D. Carson, Cloning, expression, and chromosome mapping of the murine *Hip*/*Rp1*29 gene, Genomics 68 (2000) 210–219.

[11] A. Molina, A. Iyengar, L.F. Marins, F. Biemar, S. Hanley, N. Maclean, T.J. Smith, J.A. Martial, M. Muller, Gene structure and promoter function of a teleost ribosomal protein: a tilapia (*Oreochromis mossambicus*) L18 gene, Biochim. Biophys. Acta 1520 (2001) 195–202.

[12] R. Hinegardner, D.E. Rosen, Cellular DNA content and the evolution of teleostean fishes, Am. Nat. 106 (1972) 621–644.

[13] S. Brenner, G. Elgar, R. Sandford, A. Macrae, B. Venkatesh, S. Aparicio, Characterization of the pufferfish (*Fugu*) genome as a compact model vertebrate genome, Nature 366 (1993) 265–268.

[14] G. Elgar, R. Sandford, S. Aparicio, A. Macrae, B. Venkatesh, S. Brenner, Small is beautiful: comparative genomics with the pufferfish (*Fugu rubripes*), Trends Genet. 12 (1996) 145–150.

[15] S. Aparicio, J. Chapman, E. Stupka, N. Putnam, J.M. Chia, P. Dehal, A. Christoffels, S. Rash, S. Hoon, A. Smit, et al., Whole-genome shotgun assembly and analysis of the genome of *Fugu rubripes*, Science 297 (2002) 1301–1310.

[16] P. Gilligan, S. Brenner, B. Venkatesh, Fugu and human sequence comparison identifies novel human genes and conserved non-coding sequences, Gene 294 (2002) 35–44.

[17] S. Brenner, B. Venkatesh, W.H. Yap, C.F. Chou, A. Tay, S. Ponniah, Y. Wang, Y.H. Tan, Conserved regulation of the lymphocyte-specific expression of lck in the Fugu and mammals, Proc. Natl. Acad. Sci. USA 99 (2002) 2936–2941.

[18] J.H. Leu, S.J. Yan, T.F. Lee, C.M. Chou, S.T. Chen, P.P. Hwang, C.K. Chou, C.J. Huang, Complete genomic organization and promoter analysis of the round-spotted pufferfish *JAK*1, *JAK*2, *JAK*3, and *TYK*2 genes, DNA Cell Biol. 19 (2000) 431–446.

[19] M. Westerfield, The Zebrafish Book, third ed., University of Oregon Press, Eugene, OR, USA, 1995.

[20] J.H. Leu, M.S. Chang, C.W. Yao, C.K. Chou, S.T. Chen, C.J. Huang, Genomic organization and characterization of the promoter region of the round-spotted pufferfish (*Tetraodon fluviatilis*) *JAK*1 kinase gene, Biochim. Biophys. Acta 1395 (1998) 50–56.

[21] R. Breathnach, C. Benoist, K. O'hare, F. Gannon, P. Chambon, Ovalbumin gene: evidence for a leader sequence in mRNA and DNA sequences at the exon–intron boundaries, Proc. Natl. Acad. Sci. USA 75 (1978) 4853–4857.

[22] R. Breathnach, P. Chambon, Organization and expression of eucaryotic split genes coding for proteins, Annu. Rev. Biochem. 50 (1981) 349–383.
[23] Y. Margalit, S. Yarus, E. Shapira, Y. Gruenbaum, A. Fainsod, Isolation and characterization of target sequences of the chicken CdxA homeobox gene, Nucleic Acids Res. 21 (1993) 4915–4922.
[24] S. Apone, S.D. Hauschka, Muscle gene E-box control elements. Evidence for quantitatively different transcriptional activities and the binding of distinct regulatory factors, J. Biol. Chem. 270 (1995) 21420–21427.
[25] M.A. Groenen, R.J. Dijkhof, J.J. van der Poel, R. van Diggelen, E. Verstege, Multiple octamer binding sites in the promoter region of the bovine alpha s2-casein gene, Nucleic Acids Res. 20 (1992) 4311–4318.
[26] M. Merika, S.H. Orkin, DNA-binding specificity of GATA family transcription factors, Mol. Cell. Biol. 13 (1993) 3999–4010.
[27] H. Nakagoshi, T. Nagase, C. Kanei-Ishii, Y. Ueno, S. Ishii, Binding of the c-myb proto-oncogene product to the simian virus 40 enhancer stimulates transcription, J. Biol. Chem. 265 (1990) 3479–3483.
[28] E. Ogawa, M. Maruyama, H. Kagoshima, M. Inuzuka, J. Lu, M. Satake, K. Shigesada, Y. Ito, PEBP2/PEA2 represents a family of transcription factors homologous to the products of the *Drosophila* runt gene and the human AML1 gene, Proc. Natl. Acad. Sci. USA 90 (1993) 6859–6863.
[29] F.C. de la Brousse, E.H. Birkenmeier, D.S. King, L.B. Rowe, S.L. McKnight, Molecular and genetic characterization of GABP beta, Genes Dev. 8 (1994) 1853–1865.
[30] S.C. Sung, T.J. Fan, C.M. Chou, J.H. Leu, Y.L. Hsu, S.T. Chen, Y.C. Hsieh, C.J. Huang, Genomic structure, expression and characterization of a STAT5 homologue from pufferfish (*Tetraodon fluviatilis*), Eur. J. Biochem. 270 (2003) 239–252.
[31] S. Higashijima, H. Okamoto, N. Ueno, Y. Hotta, G. Eguchi, High-frequency generation of transgenic zebrafish which reliably express GFP in whole muscles or the whole body by using promoters of zebrafish origin, Dev. Biol. 192 (1997) 289–299.
[32] Q. Long, A. Meng, H. Wang, J.R. Jessen, M.J. Farrell, S. Lin, GATA-1 expression pattern can be recapitulated in living transgenic zebrafish using GFP reporter gene, Development 124 (1997) 4105–4111.
[33] T. Yoganathan, N.K. Bhat, B.H. Sells, A positive regulator of the ribosomal protein gene, beta factor, belongs to the ETS oncoprotein family, Biochem. J. 287 (1992) 349–353.
[34] R.R. Genuario, R.P. Perry, The GA-binding protein can serve as both an activator and repressor of ribosomal protein gene transcription, J. Biol. Chem. 271 (1996) 4388–4395.
[35] J. Kusuda, M. Hirai, R. Tanuma, M. Hirata, K. Hashimoto, Genomic structure and chromosome location of RPL27A/Rpl27a, the genes encoding human and mouse ribosomal protein L27A, Cytogenet. Cell Genet. 85 (1999) 248–251.
[36] T. Evans, M. Reitman, G. Felsenfeld, An erythrocyte-specific DNA-binding factor recognizes a regulatory sequence common to all chicken globin genes, Proc. Natl. Acad. Sci. USA 85 (1988) 5976–5980.
[37] S.A. Eraly, S.B. Nelson, K.M. Huang, P.L. Mellon, Oct-1 binds promoter elements required for transcription of the GnRH gene, Mol. Endocrinol. 12 (1998) 469–481.
[38] H. Skala, A. Porteu, M. Thomas, M.F. Szajnert, H. Okazawa, A. Kahn, F. Phan-Dinh-Tuy, Upstream elements involved in vivo in activation of the brain-specific rat aldolase C gene. Role of binding sites for POU and winged helix proteins, J. Biol. Chem. 273 (1998) 31806–31814.
[39] T.J. Hinterberger, D.A. Sassoon, S.J. Rhodes, S.F. Konieczny, Expression of the muscle regulatory factor MRF4 during somite and skeletal myofiber development, Dev. Biol. 147 (1991) 144–156.
[40] M.A. Rudnicki, P.N. Schnegelsberg, R.H. Stead, T. Braun, H.H. Arnold, R. Jaenisch, MyoD or Myf-5 is required for the formation of skeletal muscle, Cell 75 (1993) 1351–1359.
[41] S.L. Amacher, J.N. Buskin, S.D. Hauschka, Multiple regulatory elements contribute differentially to muscle creatine kinase enhancer activity in skeletal and cardiac muscle, Mol. Cell. Biol. 13 (1993) 2753–2764.
[42] M.V. Rao, M.J. Donoghue, J.P. Merlie, J.R. Sanes, Distinct regulatory elements control muscle-specific, fiber-type-selective, and axially graded expression of a myosin light-chain gene in transgenic mice, Mol. Cell. Biol. 16 (1996) 3909–3922.
[43] D.E. Hoke, S.R. LaBrenz, M. Hook, D.D. Carson, Multiple domains contribute to heparin/heparan sulfate binding by human HIP/L29, Biochemistry 39 (2000) 15686–15694.
[44] A. Amsterdam, S. Lin, N. Hopkins, The *Aequorea victoria* green fluorescent protein can be used as a reporter in live zebrafish embryos, Dev. Biol. 171 (1995) 123–129.
[45] M. Kinoshita, S. Kani, K. Ozato, Y. Wakamatsu, Activity of the medaka translation *elongation factor 1alpha*-A promoter examined using the GFP gene as a reporter, Dev. Growth Differ. 42 (2000) 469–478.
[46] S. Higashijima, H. Okamoto, N. Ueno, Y. Hotta, G. Eguchi, High-frequency generation of transgenic zebrafish which reliably express GFP in whole muscles or the whole body by using promoters of zebrafish origin, Dev. Biol. 192 (1997) 289–299.
[47] K. Hamada, K. Tamaki, T. Sasado, Y. Watai, S. Kani, Y. Wakamatsu, K. Ozato, M. Kinoshita, R. Kohno, S. Takagi, M. Kimura, Usefulness of the medaka beta-actin promoter investigated using a mutant GFP reporter gene in transgenic medaka (*Oryzias latipes*), Mol. Mar. Biol. Biotechnol. 7 (1998) 173–180.
[48] P.D. Gibbs, M.C. Schmale, GFP as a genetic marker scorable throughout the life cycle of transgenic zebrafish, Mar. Biotechnol. 2 (2000) 107–125.