Zoological Studies **36**(3): 159-169 (1997)

Zoological Studies

Genomic Structure and Promoter Region of c-*fos* Gene of Round-spotted Pufferfish, *Tetraodon nigroviridis* (Syn. *T. fluviatilis*)

Mau-Sun Chang[1], Jiann-Horng Leu[2], Chen-Wen Yao[3], Fore-Lien Huang[1] and Chang-Jen Huang[2,*]

[1]*Department of Zoology, National Taiwan University, Taipei, Taiwan 106, R.O.C.*

[2]*Institute of Biological Chemistry, Academia Sinica, Taipei, Taiwan 115, R.O.C. Tel:886-2-7855696 Fax:886-2-7889759*

[3]*Graduate Institute of Life Science and Institute of Preventive Medicine, National Defense Medical Center, Taipei, Taiwan 100, R.O.C.*

(Accepted February 17, 1997)

Mau-Sun Chang, Jiann-Horng Leu, Chen-Wen Yao, Fore-Lien Huang and Chang-Jen Huang (1997) Genomic structure and promoter region of c-*fos* gene of round-spotted pufferfish, *Tetraodon nigroviridis* (Syn. *T. fluviatilis*). *Zoological Studies* **36**(3): 159-169. The round-spotted pufferfish c-*fos* gene was isolated from a liver genomic library. Complete DNA sequences, including 0.75-kb of the promoter region were determined. Our data indicate that this gene is composed of 4 exons and 3 introns spanning about 3 kb. This gene encodes a protein of 374 amino acids with a molecular mass of 42 262 Da. The amino acid sequence identities between this puffer fish c-fos protein and fugu (another pufferfish species), human, mouse, and chicken c-fos protein are 90%, 58%, 58%, and 53%, respectively. Two potential transcription initiation sites were located by primer extension analysis. Examination of 0.75-kb of a 5'-flanking sequence revealed potential binding sites for a variety of transcription factors such as CREB, E2A, Ets-1, GATA-1, HNF-5, c-Myb, and Sp1. When the 0.9-kb DNA fragment (-746 to +153) was placed upstream of the chloramphenicol acetyltransferase (CAT) reporter gene and transfected into a carp CF cell line, it could drive the synthesis of CAT enzyme 2.7 times more efficiently than could the RSV promoter. Although lacking the known serum-response element, the CAT activity of this region still could be induced by serum.

Key words: Genomic DNA sequence, c-*fos*, Round-spotted pufferfish *Tetraodon nigroviridis*, Chloramphenicol acetyltransferase (CAT).

The proto-oncogene *fos* is the normal cellular homolog of the transforming gene, v-*fos*, of the FBJ murine osteosarcoma virus (Curran et al. 1983, Verma and Sassone-Corsi 1987). The c-Fos oncoproteins can associate with the c-Jun proteins to form the gene regulator AP-1 (activator protein-1) which regulates transcription of a specific set of genes associated with cellular proliferation and differentiation (Angel and Karin 1991). The transient expression of Fos is one of the earliest responses to numerous external stimuli, such as phorbol esters, serum, and calcium ionophores. However, overexpression of c-Fos in transgenic mice causes osteosarcomas and chondrosarcomas (Ruther et al. 1989). On the other hand, mice lacking Fos develop osteopetrosis and exhibit altered hematopoiesis. Thus, these results suggest that Fos also plays important roles in the regulation of osteoclast-macrophage lineage determination (Johnson et al. 1992, Grigoriadis et al. 1994).

The mammalian AP-1 family is now known to consist of multiple proteins, including those that are Fos related (c-Fos, FosB, Fra-1, and Fra-2) and those that are Jun related (c-Jun, JunB, and JunD) (Angel and Karin 1991). The consensus sequence of its recognition site has been identified as TGA(G/C)T(A/C)A, which can also act as a TPA- and serum-inducible enhancer element (Lee et al. 1987). All of the Jun proteins can form both homo- and heterodimers (such as JunB/c-Jun) which thus

*To whom correspondence and reprint requests should be addressed.

**The sequence data in this paper has been submitted to the EMBL/GenBank Data Libraries under the accession number U53520.

bind to the AP-1 binding site. On the other hand, Fos proteins are not able to form homodimeric complexes and therefore do not bind DNA by themselves. However, Fos proteins can associate with any of the Jun proteins to form stable heterodimers that have higher DNA-binding activity than the Jun dimers alone (Angel and Karin 1991).

In our previous report (Chang et al. 1996), we determined the genomic structure of the carp JAK1 kinase gene and found 2 AP-1 binding sites (TGACTCA and TGAGTAA) present in the promoter region of this gene. We are interested in the transcriptional regulation of the carp JAK1 kinase gene by AP-1. As an initial step to target this issue, we first cloned the c-*fos* gene of the round-spotted pufferfish, *Tetraodon nigroviridis*, by screening a genomic library and herein present the complete sequence as well as the promoter region of this gene. The activity of the putative promoter region of this gene has also been investigated by transient expression in carp epitheloid CF cells (Chen and Kou 1986).

MATERIALS AND METHODS

Isolation of genomic clones

By using the lambda FIXII as a cloning vector (Stratagene, La Jolla, CA, USA), a round-spotted pufferfish, *Tetraodon nigroviridis*, liver genomic library was constructed and contained approximately 5×10^5 independent clones. The amplified library was then used to isolate 15- to ~18 kilobase (kb) genomic DNA clones containing the gene that encodes round-spotted puffer c-Fos. Degenerate primers were designed to fit the amino acid sequences that are highly conserved in the Fos and Fos-related proteins from several different species, such as human (van Straaten et al. 1983), mouse (van Beveren et al. 1983), and chicken (Fujiwara et al. 1987). The amino acid sequences of the 2 opposing primers are ^{134}PEEEEK and ^{194}EFILAAH. With these primers and genomic DNA as template, a PCR (Mullis and Faloona 1987) product of 270 base pairs (bp) was obtained (data not shown). Amplified DNA fragments were then purified and ligated into pGEM-T (Promega, Madison, WI, USA). Among the 12 clones sequenced, 10 clones contained the same *fos*-related sequences. Thus, DNA clones containing the *fos*-related gene were then used as a probe to screen the above round-spotted pufferfish genomic library. The probe was labeled using a DIG DNA Labeling Kit (Boehringer Mannheim, Mannheim, Germany). Approximately 1×10^6 amplified clones were plated at a density of 5×10^4 plaque forming units/150-mm Petri dish. Hybridization and washing were carried out as previously described (Sambrook et al. 1989). In brief, nitrocellulose (Schleicher and Schuell, Dassel, Germany) lifts of the phage plates were hybridized at 42 °C overnight in 50% (v/v) formamide containing 5 × SSC (1 × SSC = 150 mM NaCl, 15 mM sodium citrate, pH 7.0), 0.1% (w/v) SDS, 0.5% (w/v) N-lauroylsarcosine, 2% Blocking Reagent (Boehringer Mannheim), and a DIG-labeled probe. Following hybridization, filters were washed in 2 × SSC, 0.1% (w/v) SDS at 25 °C for 30 min, then in 0.1% SSC, and 0.1% (w/v) SDS at 25 °C for 30 min. Signals were detected using the DIG Luminescent Detection Kit for Nucleic Acids (Boehringer Mannheim). Clones of interest were further plaque purified by 3 more screening cycles. Phage DNA was prepared from individual pure clones as described (Sambrook et al. 1989). Restriction fragments of each phage DNA were characterized by Southern analysis.

Subcloning and sequencing

Phage DNA was digested with *Not* I, *Sac* I, and *Sal* I and subcloned into either the pUC18 (Boehringer Mannheim) or pBluescript (Stratagene) vector. Each subclone was sequenced by the dideoxy chain termination method (Sanger et al. 1977) with Sequenase (US Biochemical, Cleveland, OH, USA) according to the manufacturer's instructions. The nucleotide sequences of the 1st and last 300 to 400 nucleotides of each DNA fragment were determined with M13/pUC direct and reverse sequence primers. Based on the resulting nucleotide sequences, 2 oligonucleotides of 20 nucleotides in length were synthesized and used as primers to determine the nucleotide sequence of the next 300 to 400 nucleotides in each direction of the double-stranded DNA. The exon-intron boundaries were determined by alignment of the resulting sequences with exons of the human (van Straaten et al. 1983), mouse (van Beveren et al. 1983), and chicken (Fujiwara et al. 1987) c-*fos* gene. Sequence assembly and alignment were performed using the Genetics Computer Group software program. The transcription factor recognition site data bases (releases 7.3 and 6.5) were used to identify transcription factor motifs within the 5' flanking region of the round-spotted pufferfish c-*fos* gene.

Primer extension analysis

An antisense oligonucleotide (5'-TCAGCGT-TAAAAGACGTAAACATC-AT-3') corresponding to 26 nucleotides of the 1st translation start site (Fig. 5) was labeled at the 5'-end with [γ-^{32}P]ATP by T4 polynucleotide kinase and purified on a Stratagene NucTrap probe purification column. Labeled primer was annealed to 5 μg of poly (A)$^+$ RNA prepared from round-spotted pufferfish brain total RNA and extended as described by Sambrook et al. (1989). The extended products were analyzed on a 5% polyacrylamide/7 M urea sequencing gel. The sizes of the extended products were inferred from a sequencing ladder of the c-*fos* gene obtained from the same primer used for primer extension.

Plasmid construction for functional analysis

A 0.9-kb DNA fragment encompassing the 5'-flanking region was subcloned into polylinker regions of the reporter vector pCAT-Basic (Promega) using PCR primers 5'-TTGGCATGCAAAAGTCG-CACTATAATTTT-3' (nucleotides -746 to -727) and 5'-GGGGTCGACTTAAAAG ACGTAAACATCAT-3' (nucleotides +134 to +153). Each primer encodes a unique restriction site (underlined) for the convenience of subcloning (*Sph* I, GCATGC; and *Sal* I, GTCGAC). The pRSV-CAT (Gorman et al. 1983) and pJP1-CAT (Chang et al. 1996) were utilized as positive controls.

Transfection, CAT, and β-galactosidase assay

Carp fin epitheloid cells, CF (Chen and Kou 1986), were maintained in Leibovitz's L-15 medium supplemented with 10% fetal calf serum at 27 °C. Approximately 5 × 10^5 cells were plated in 60-mm culture dishes before transfection. After 12 h plating, cells were washed twice with Leibovitz's L-15 medium and incubated with DNA-Lipofectamine complexes containing 4 μg of the different CAT constructs together with 1 μg of pSV-β-galactosidase vector (Promega) in duplicate. Transfection was carried out for 5 h, after which cells were washed with fresh Leibovitz's L-15 medium and fed with the same medium supplemented with 5% fetal calf serum. After the expected transfecting time, cells were harvested, washed in phosphate-buffered saline, and lysed with 25 mM Tris phosphate, pH 7.8, containing 2 mM dithiothreitol, 2 mM EDTA, 10% glycerol, and 1% Triton X-100 at room temperature for 30 min. The total lysates were scraped from the dish and transferred to microcentrifuge tubes. Cell debris were removed by centrifugation at 12 000 rpm for 10 min and the extracts were frozen at −70 °C. Protein concentration was measured by the Bio-Rad protein concentration quick-assay method (Bio-Rad, Richmond, CA, USA). CAT and β-galactosidase activities in these cell extracts were measured according to previously described procedures (Herbomel et al. 1984). Acetylated products of the CAT assay were separated by thin layer chromatography, developed with chloroform-methanol (95:5, v/v), visualized by autoradiography, and quantified by using the PhosphoImager (Bio-Imaging Analyzer BAS 2000, Fuji, Japan).

RESULTS

Genomic structure of the round-spotted pufferfish c-*fos* gene

Two positive clones, F1 and F2, were isolated, and their restriction enzyme maps were almost identical. Therefore, only the F1 clone was further characterized and sequenced. As shown in Fig. 1, the coding region (from translation start codon

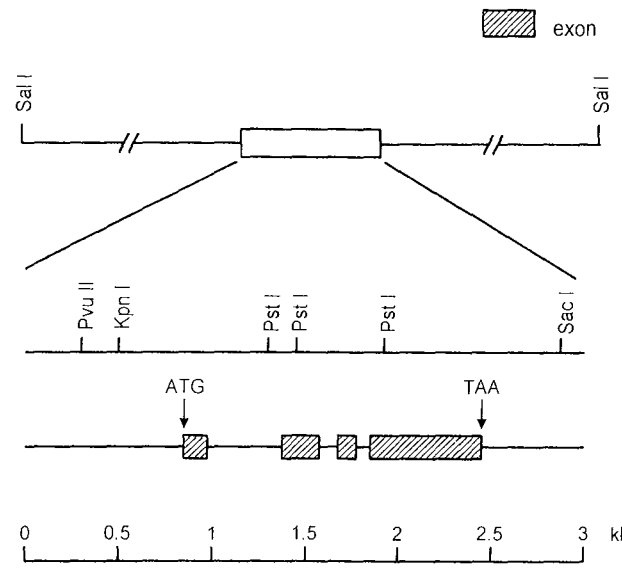

Fig. 1. Physical map of the round-spotted pufferfish c-*fos* gene. The structure of the gene is represented with regard to the organization of the exons and introns. Exons are indicated to scale by hatched boxes whereas introns and the 5'- and 3'-flanking regions are indicated by solid lines. The entire gene spans about 3 kb in length and contains 4 exons. The restriction map shows cleavage sites for endonucleases *Kpn* I, *Pst* I, *Pvu* II, *Sac* I, and *Sal* I.

```
ATGATGTTTACGTCTTTTAACGCTGAGTGCGACTCTTCTTCCCGCTGCTCTGCCTCCCCGTCCGACAATGTCTACTATCC   80
M  M  F  T  S  F  N  A  E  C  D  S  S  S  R  C  S  A  S  P  S  D  N  V  Y  Y  P    27
GTCCCCGGCAGGGTCCTACTCCAGCATGGGCTCCCCGCAGTCTCAGgtacggccggctcagtgttgcatattagacgaag   160
 S  P  A  G  S  Y  S  S  M  G  S  P  Q  S  Q                                       42
tttcaggcagtttgtgtattttatgtcccgtttggcgagccggaaaacgtgatacagctagtttaatgctttttaacata   240
acagcacaagcttttcaggtgatccaattcaatctgatctttcctaacaacacactgacctccacttgcaataaattcgt   320
tcatctggaaagtcctcagagagtgtgtaacgtcatagctgacatcagagggatatttttacattctctagcttttaata   400
gaagctctccaatcttcattcataaccatccttgcaaagctgcagcagcttctcggtctcgaaagtgcagggctgagctt   480
ccctttccctatttctccacagGATTTGACTGACCTGACAGCATCAAGTGCCTCCTTCGTCCCCACGGTCACAGCCATCT   560
                                D  L  T  D  L  T  A  S  S  A  S  F  V  P  T  V  T  A  I  S    62
CGACCAGCCCGGATCTGCAGTGGATGGTGCAGCCTTTGGTCTCCTCGGTGGCTCCTTCTCGCAGGGCTCACCCCTACAGT   640
  T  S  P  D  L  Q  W  M  V  Q  P  L  V  S  S  V  A  P  S  R  R  A  H  P  Y  S     88
CCCAGTCCCTCCTACAAGAGAACCGTCATGAGGTCTGGAGCCTCCAAGCCACACGCCAAGAGGGGTCGGGTGGAACAGgt   720
P  S  P  S  Y  K  R  T  V  M  R  S  G  A  S  K  P  H  A  K  R  G  R  V  E  Q      114
aattagaactccgccttgtgttcccggggttttgtaaccgatagcaacacctgtttggccgtcagccgtcgctgaagcct   800
aaacttttgttgtcgtcactcagACTACACCCGAGGAGGAGGAGAAGAAAAGAATCCGCAGGGAGAGAAATAAGCAGGC   880
                                T  T  P  E  E  E  E  K  K  R  I  R  R  E  R  N  K  Q  A    133
AGCAGCTAAATGTCGTAACAGGAGGCGAGAACTCACAGATTCTTTGCAAGCGgtaagactcgctcacatctgctgtcttc   960
 A  A  K  C  R  N  R  R  R  E  L  T  D  S  L  Q  A                                  150
cagtgaaacaagttctcacatttgttcgcagtcttgtgcaacagGAAACCGATCAGTTAGAGGCTGAGAAATCCAGCCTG  1040
                                                E  T  D  Q  L  E  A  E  K  S  S  L    162
CAGAACGATATTGCCAATCTTCTGAAGGAGAAGGAGCGTCTGGAGTTCATTCTGGCTGCCCACCAGCCCATCTGCAAGAT  1120
Q  N  D  I  A  N  L  L  K  E  K  E  R  L  E  F  I  L  A  A  H  Q  P  I  C  K  I    189
CCCCTCCCAGATGGACTCAGACTTCCCTGTGGTCTCCATGTCTCCGGTGCACGCCTACCTCTCTCCACCGCTGCCTCCACGC  1200
 P  S  Q  M  D  S  D  F  P  V  V  S  M  S  P  V  H  A  Y  L  S  T  A  A  S  T  Q    216
AGCCACAGACCTCCGTCCCAGAGGCCACCACCGTCACCTCAAGCCACTCGACATTCACCTCAACCTCCAACTCCATTTTC  1280
  P  Q  T  S  V  P  E  A  T  T  V  T  S  S  H  S  T  F  T  S  T  S  N  S  I  F     242
GGCAGCAACAGCGACTCCCTCCTCTCCACCGCCACCGTGTCCGACAGCGTGGTGAAGATGACTGACCTGGAATCCTCCGT  1360
G  S  N  S  D  S  L  L  S  T  A  T  V  S  D  S  V  V  K  M  T  D  L  E  S  S  V    269
CCTGGAAGAGTCGTTGGACCTGCTGGCAAAGACGGAGGTGGAGACGGTCGAGGTTCCCGACGTCAACCTGTCCAGCTCCC  1440
 L  E  E  S  L  D  L  L  A  K  T  E  V  E  T  V  E  V  P  D  V  N  L  S  S  S  L    296
TCTACACAGCCCAGGACTGGGAGCCTCTCCACGCCACCATCGGCAGCAGTGACTTTGAGCCCCTGTGCACGCCCCGTGGTG  1520
  Y  T  A  Q  D  W  E  P  L  H  A  T  I  G  S  S  D  F  E  P  L  C  T  P  V  V     322
ACGTGCACGCCCGGCCTGCACCACCATCACATCCTCTTTTGTCTTCACCTTTCCAGAGGCCGAGACCTTCCCCACCTGCTG  1600
T  C  T  P  A  C  T  T  I  T  S  S  F  V  F  T  F  P  E  A  E  T  F  P  T  C  C    349
CGTCGCCCACAGGAGAGGAAGCAACAGCAACGACCAGTCCTCTGACTCCCTCAGCTCCCCGACCCTGCTGGCCCTCTAAA  1680
 V  A  H  R  R  G  S  N  S  N  D  Q  S  S  D  S  L  S  S  P  T  L  L  A  L  *     374
GACTCTTCCAAAAAGAAAAGGCTGCCTCTCTCAGTCCACCTTTCTCCATCTATGAGACGTGCGCGGATCACAGGGCTATG  1760
GAATGGACTTGCACCAGGAAGCAAAGTGTTTATAATATCATTTGCCTTTATCTACTTGCTTATAAGACCCATGTAGCAAG  1840
TTAAAAAGCATGTTTTAACTAAAACCAGTTAGATGTTTAATGAGTTACGGTATCATTTAACGGTGCTATAGAGACAACGG  1920
AACCTGTATTAGTGCTGATATATTTGGAGCAATAGTTCTATTGGTCTTCTGTTTTCTGGTTGATGGATGTGAGGGTGCTT  2000
TGTGTTTGTGGAGCTCAAGAACTGACGTGCTTTGTCCCAACCATGGTCTGAGTGGTCTATCTGTACGAAGTTTTCTGTTA  2080
AACAAACGTGGCTTTAATTT                                                               2100
```

Fig. 2. Complete nucleotide sequence of the round-spotted pufferfish c-*fos* gene. The nucleotide sequence of a 3-kb DNA fragment containing the promoter region of the pufferfish c-*fos* gene was determined. Only 2100 nucleotides encompassing the coding region as well as the 3'-untranslated region are shown in this figure. The other promoter sequences are shown in Fig. 5. All nucleotide sequences in this figure and Fig. 5 have been deposited in GenBank with the accession number U53520.

ATG to stop codon TAA) of the round-spotted pufferfish c-*fos* gene spans about 1.7 kb and is composed of 4 exons and 3 introns. The complete DNA sequence of this region as well as the 3'-untranslated region were determined and are shown in Fig. 2. All exon/intron boundaries identified conformed to the GT/AG splice donor/acceptor rule (Breathnach et al. 1978). The sizes of the 3 introns were 376 bp (intron 1), 106 bp (intron 2), and 72 bp (intron 3), respectively.

The open reading frame of these 4 exons encodes a protein of 374 amino acids with a molecular mass of 42 262 Da. As shown in Fig. 3, the deduced amino acid sequence contains 2 methionines, as also observed in avian and mammalian Fos proteins (van Beveren et al. 1983, van Straaten et al. 1983, Fujiwara et al. 1987). The overall amino acid sequence identities between the round-spotted pufferfish c-Fos protein and those of fugu (Trower et al. 1996), human (van Straaten et al. 1983), mouse (van Beveren et al. 1983), and chicken (Fujiwara et al. 1987) are 90%, 58%, 58%, and 53%, respectively. In addition, the round-spotted pufferfish c-Fos protein has several well-conserved regions when compared with other mammalian Fos proteins, especially in the putative DNA binding region, the leucine zipper domain and the carboxyl-terminal transactivation region (Landschulz et al. 1988, Sassone-Corsi et al. 1988).

Determination of the transcription initiation site

The transcription initiation site of the round-spotted pufferfish c-*fos* gene was determined by primer extension of the 5'-end of the round-spotted puffer liver poly $(A)^+$ RNA. Two major extended products were revealed (Fig. 4). The exact position of the extended product was determined by aligning the sequencing ladder obtained with the same primer. Therefore, the extended products corresponded to sites at 135 and 133 nucleotides upstream of the 1st initiator methionine codon (Fig. 5).

Characteristics of the 5'-flanking region of the pufferfish c-*fos* gene

In order to identify sequence elements that might be involved in the transcriptional regulation of the round-spotted pufferfish c-*fos* gene, the nucleotide sequence of 880 bp of the 5'-upstream region of the c-*fos* gene relative to the translation start site was analyzed by computer and is shown in Fig. 5. One TATA box, which is generally located at a position about 30 nucleotides upstream of the RNA start site (Breathnach and Chambon 1981), was observed 39 nucleotides upstream of the RNA start site. Two putative binding sites for HNF-5 (Grange et al. 1991), a ubiquitous gene activator induced in response to various extracellular signals, were found at the position of nucleotides −392 to −386 and −106 to −100. Another potential binding site for CREB (Gonzalez and Montminy 1989) was observed at the position of nucleotides −64 to −57. Three proto-oncogene products c-Myb (Faisst and Meyer 1992), E2A (Murre et al. 1989), and Ets-1 (Wasylyk et al. 1993) binding sites were also found at the positions of nucleotides −430 to −425, −402 to −396, and −314 to −307, respectively. A site for Sp1 (Jackson et al. 1990) was found at the positions of nucleotides −327 to −320. In addition, 2 consensus GATA-1 (Orkin 1992) binding sequences were also observed at the position of nucleotides −546 to −541 and −164 to −159.

Functional analysis of the promoter region

To verify whether the 5'-flanking region of the round-spotted pufferfish c-*fos* gene exhibits functional promoter activity, the genomic DNA fragment containing the 5'-upstream region (−746 to +153) was fused to the CAT reporter gene in pCAT-Basic (Promega) to create pPF-fos-CAT. Following transfection into carp CF cells, this chimeric gene produced about 518 times the promoter activity of the pCAT-Basic, indicating that this fragment contains a functional eukaryotic promoter. In addition, pPF-fos-CAT displayed stronger CAT activity, approximately 11.5 and 2.7 times the promoter activity of pJP1-CAT (Chang et al. 1996) and pRSV-CAT (Gorman et al. 1983), respectively (Fig. 6). This finding suggests that the promoter of the round-spotted pufferfish c-*fos* gene may be useful to drive the expression of foreign genes at a higher level than that of viral promoter in fish cell lines.

To understand whether the CAT activity of pPF-fos-CAT could be induced by serum stimulation, carp CF cells were transfected with pPF-fos-CAT and starved for 24 h in medium without serum. Then, 5% serum was added and CF cells were harvested at various times. As shown in Fig. 7, serum induced CAT activity of the round-spotted pufferfish c-*fos* promoter region 4-fold at 8 h and 16 h, and approximately 1.3-fold at 24 h. Furthermore, although the CAT activity of pPF-fos-CAT was dramatically reduced at 32 h, the effect of serum stimulation was still obvious.

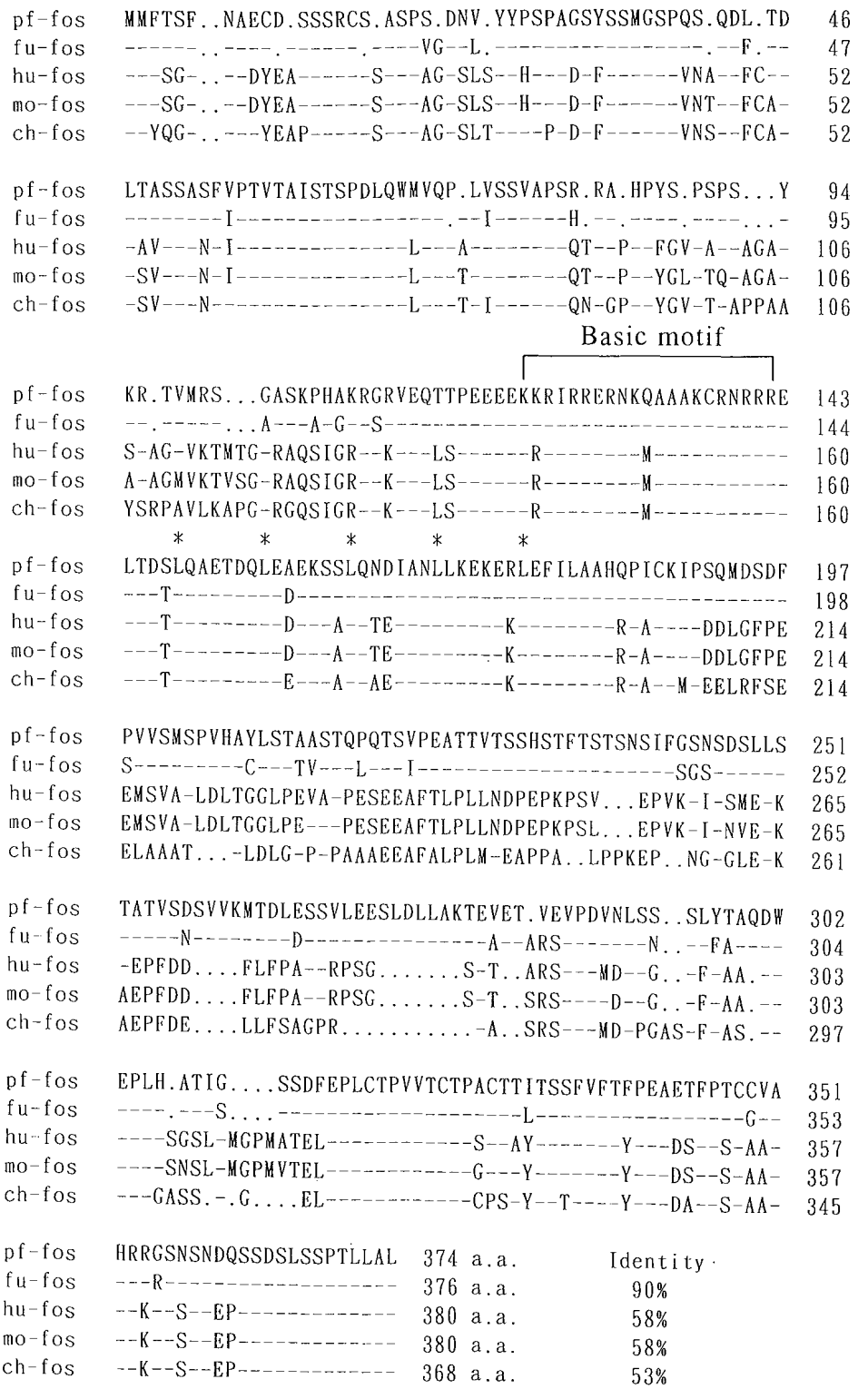

Fig. 3. Amino acid comparison of the round-spotted pufferfish, *Fugu*, human, mouse, and chicken c-Fos proteins. Dashes indicate amino acids identical to the round-spotted pufferfish c-Fos whereas dots indicate gaps introduced to generate an optimal alignment. The sequences of c-Fos proteins from *Fugu* (Trower et al. 1996), human (van Straaten et al. 1983), mouse (van Beveren et al. 1983), and chicken (Fujiwara et al. 1987) were retrieved from the databases. The identity of each Fos protein to the round-spotted pufferfish c-Fos is indicated at the end of each sequence. Asterisks represent leucine amino acids of the leucine zipper, which occur five times in the c-Fos protein. The putative DNA binding region is also indicated by the basic motif just before the leucine zipper domain.

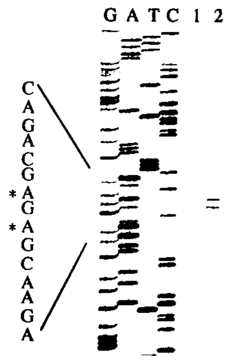

Fig. 4. Determination of the transcription initiation site of the round-spotted pufferfish *c-fos* gene. [γ-^{32}P]-labeled primer (see Fig. 5) was annealed to 10 μg yeast tRNA (lane 1, as a negative control) or 5 μg poly $(A)^+$ RNA from round-spotted puffer liver (lane 2) and extended with reverse transcriptase. The sequencing ladder of the round-spotted puffer *c-fos* gene labeled G, A, T, and C was obtained by using the same primer and electrophoresed on the same gel. The extended products and their corresponding sites are indicated by asterisks.

DISCUSSION

The pufferfish *Fugu rubripe* (*Fugu*) has been shown to have a compact genome and thus is used as a model for vertebrate genome analysis (Brenner et al. 1993). *Fugu* belongs to the order Tetraodontiformes and is widely consumed in Japan. Its genome is estimated to be approximately 404 Mb, 7.5 times smaller than that of humans. Recently, many *Fugu* genes have been cloned and characterized, such as the glucose-6-phosphate dehydrogenase gene (Mason et al. 1995), the p55 gene (Elgar et al. 1995), Huntington's disease (HD) gene (Baxendale et al. 1995), dopamine receptor genes (Macrae and Brenner 1995), and the chaperonin-containing t-complex polypeptide 1 gene (Yoda et al. 1995). All these genes are compact and contain small introns, except the *Fugu* dopamine receptor genes. It is noteworthy that the *Fugu* HD gene consists of 64 exons and

```
       .         .         .         .         .         .         .
-746 AAAAGTCGCACTATAATTTTATATTGCGTTGGAAATTATATATATATATAAAAAAAAAAAAAATCTTTGTT

       .         .         .         .         .         .         .
-676 CTGGTATCATCGTCAATGCCCTAATATAATTCTTAGTAAAAATACAGGTATAAAAATGCAACGTATTTTA
       .         .         .         .         .         .       GATA-1
-606 ATACAGGCATTTTAATACAGGTATACAAATGGCACGTATTGTTTGAAACCTTTAAACATGTGATAAGTGT

       .         .         .         .         .         .         .
-536 TAACACGTACACTAGCCGTGTGTTTTTACTGGAACTGCTCAAACAATACGTGGTGTTTAAGTTTACTTTT
       .         .         .         .    c-Myb        .         .    E2A
-466 GATGTACTTTTGTAGTGTTGGATGTACTTTTGCTGACCGTTATCCCAGTTCTTTAGTTTTCACAACAGCT
       .    HNF-5.         .         .         .         .         .
-396 GTTGTGTTTGTTGTCTTACGTAACGTGGTGCAGATGTCGCCAATGCCGGAGAGCCTGTTTCCCAGCAGGC
       . Spl     .    Ets-1.         .         .         .         .
-326 CCGCCCCTTTTACAGGATGCACCATATTTGGAAAATGACGTACTGGGGATCCCGTTTCTGGAATATTCCA

       .         .         .         .         .         .         .
-256 TTTATTGGAGGTACCATTAAACGACAACTTTTACGTTAGAATTTCCGCTTTAAACGGTAAAAATTCTCAA
       .         .         . GATA-1  .         .         .         .
-186 ATATGTTAAGTTTTTGCGTGTTAGATAGAAGAACGTGAGTGGTTTTGTGTTACTGCCGCGTTGTTAGGTC
       .         .HNF-5    .         .         .         .   CREB  .
-116 GGTGTTTTGGTGTTTGCTTCGCGGCGTTTGAAAGGGAACGCTCCCCTTTTGCTGACGTCATCAGCTCAGG
       .TATA box .         .         .         .   * *     .         .
 -46 CTATATAAGCGCGCGTGCGGCGACTCCGCCGAAAGTTTCAGACGAGAGCAAGAGAAGCCAGAGCACGTCC
       .         .         .         .         .  -2 +1    .         .
 +25 GCGGGCTCACACACCGACCTGACCGAAGATTCCCTCCTTTCCAGGGTTGAAATCCCGACCGCAGCCTTTG

       .         .         .         .         .         .         .
 +95 TTTGGATTTACCAGCTTGCTTTTTTTCGCCCGGACAGCGATGATGTTTACGTCTTTTAACGCTGAGTGCG
                                             <-------------------------
```

Fig. 5. Nucleotide sequence of the 5'-flanking region of the round-spotted pufferfish *c-fos* gene. The candidate transcription start sites by primer extension (see Fig. 4) are indicated with nucleotide numbers (−2), (+1), and asterisks, which are located at 135 and 133 bp upstream of the 1st translation start site. A 26-mer antisense oligonucleotide used for primer extension analysis is underlined. Potential binding sites for a variety of transcription factors are also marked and underlined. Nucleotide +134 in this figure is the same as nucleotide 1 of Fig. 2.

spans only 23 kb, compared to 170 kb of the human HD gene. Moreover, the intron/exon boundaries of these genes between human and *Fugu* are highly conserved. This strongly suggests the feasibility of using the pufferfish genome as a model system to analyze human genes.

In this study, we cloned and characterized the c-*fos* gene from a different species of the round-spotted pufferfish, *Tetraodon nigroviridis*, which is easily obtained at cheaper prices from a local aquarium. The round-spotted pufferfish has a genome size of 380 Mb which is believed to be the smallest of any vertebrate and is 8 times smaller than that of humans (Hinegardner and Rosen 1972, Brenner et al. 1993). Recently, the *Fugu* c-*fos* gene was identified to be linked to the familial Alzheimer disease (AD3) locus (Trower et al. 1996). The relative gene order in the AD3 locus is the same in *Fugu* and human genomes. However, the genome size in the AD3 locus is 12.4 kb in *Fugu*, compared to > 600 kb in humans. Once again, these are consistent with the observation that intergenic distances are much smaller in the *Fugu* genome compared to the human genome (Brenner et al. 1993). The overall amino acid and nucleotide sequence identities of the c-*fos* gene between round-spotted pufferfish and *Fugu* is 90.4% and 86.8% , respectively. Sizes of the 3 introns of the round-spotted pufferfish c-*fos* gene is 376 bp, 106 bp, and 72 bp, compared to 397 bp, 104 bp, and 102 bp in the *Fugu* c-*fos* gene. Moreover, the last 200 nucleotides of the 1st intron and 400 nucleotides of the 3-untranslated region of both c-*fos* genes are very similar, e.g., 90% and 84% identical, respectively (data not shown). These results indicate that the intron size of the

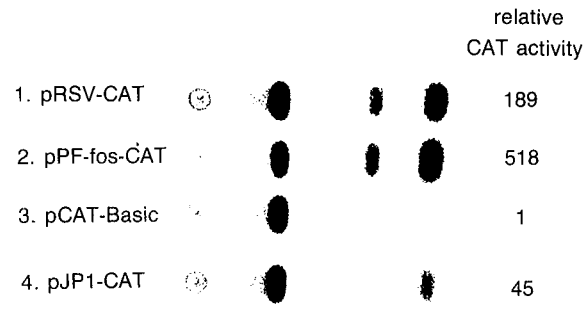

Fig. 6. Analysis of the promoter activity of the 5'-flanking region of the round-spotted pufferfish c-*fos* gene fused to the CAT reporter gene. Each chimeric gene was cotransfected with pSV-β-galactosidase DNA into CF cells and assayed for CAT and β-galactosidase activities as described under Materials and Methods. CAT activity in the individual experiments was corrected for variations in transfection efficiency by normalizing the value to the β-galactosidase activity in the same extract. The normalized activity of each promoter construct was then expressed relative to that of pCAT-Basic, with pCAT-Basic assigned a relative activity of 1.0. The data represent the mean of triplicate transfection experiments for each plasmid.

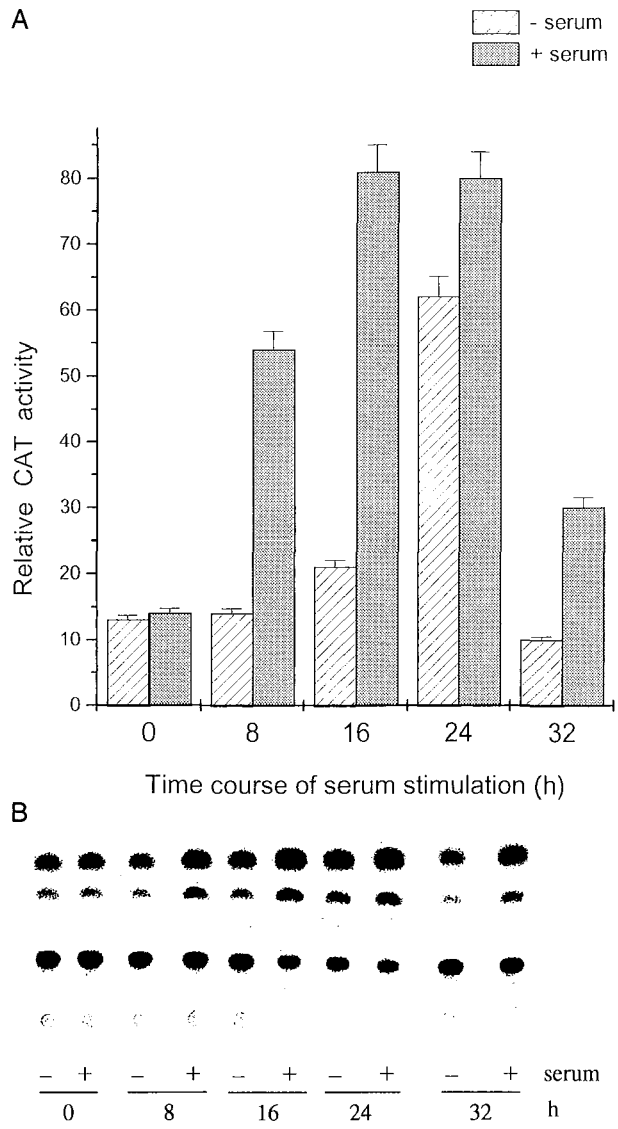

Fig. 7. Time course of the CAT activity of pPF-fos-CAT in carp CF cells with and without the supplement of 5% fetal calf serum. Initially, carp CF cells were transfected with pPF-fos-CAT DNA and pSV-β-galactosidase for 5 h, then were grown in medium without serum for another 19 h. Next, culture medium with 5% fetal calf serum was added and the CF cells were harvested at various times. Cell extracts were assayed for CAT and β-galactosidase activities. The relative CAT activity is represented by the mean result of triplicate transfection and normalized to the β-galactosidase activity. Acetylated chloramphenicol derivatives were separated from unacetylated chloramphenicol by thin-layer chromatography as shown in panel B. Relative activities at different times were quantified by using a PhosphoImager and are presented in panel A.

round-spotted pufferfish is smaller than that of *Fugu*. Therefore, the round-spotted pufferfish is an alternative species for genomic analysis.

Although the compact genome of *Fugu* has been proposed to as advantageous for detecting conserved regulatory elements present in the non-coding region (Aparicio et al. 1995), our studies indicate that the promoter region of the round-spotted pufferfish c-*fos* gene is different from that of mammalian c-*fos* genes. In mammals, there are several potential regulatory elements, such as the sis-inducible element (SIE), serum-response element (SRE), AP1 site, and calcium/cAMP-response element (Ca^{2+}/CRE) present in the c-*fos* 5'-flanking region (Wang and Howells 1994). On the contrary, the putative promoter region of the round-spotted pufferfish c-*fos* gene has other potential binding sites for transcription factors, such as CREB, E2A, Ets-1, GATA-1, HNF-5, c-Myb, and Sp1 (Fig. 5). Moreover, mammalian c-*fos* transcription increases in response to various stimuli, such as growth factors, cytokines, phorbol esters, neurotransmitters, heat shock, UV irradiation, and cAMP (Angel and Karin 1991). The SRE within the human (van Straaten et al. 1983), mouse (van Beveren et al. 1983), and rat (Wang and Howells 1994) c-*fos* promoter has been reported to be required for induction of c-*fos* by serum and growth factors. It is interesting to note that this element has not been found in the promoter region of the round-spotted pufferfish c-*fos* gene. However, the CAT activity of this region increased significantly at 16 h and 24 h after the addition of 5% calf fetal serum (Fig. 7). It seems that there may be another SRE-like element within the round-spotted pufferfish c-*fos* gene and this needs further investigation.

All of the round-spotted pufferfish and mammalian c-*fos* genes have CRE sites within the 5'-flanking region, but their sequences are slightly different. The CRE sequence of the round-spotted pufferfish c-*fos* gene is TGACGTCA, whereas that of the rat c-*fos* gene is TGACGTAG. It is interesting to note that the spacing between the CRE and the TATA box, found at nt −45 to −39, differs for mammalian species, being 11 nt in the round-spotted pufferfish, 19 nt in the rat, 23 nt in the human, and 25 nt in the mouse. Whether these differences affect the transcriptional activity of the CRE remains to be investigated.

Acknowledgements: This research was supported by grants from the National Science Council (NSC85-2311-B-001-046) and from the Department of Health (DOH85-HR-507), Taiwan, Republic of China.

REFERENCES

Angel A, M Karin. 1991. The role of Jun, Fos and the AP-1 complex in cell-proliferation and transformation. Biochim. Biophys. Acta **1072:** 129-157.

Aparicio S, A Morrison, A Gould, J Gilthorpe, C Chaudhuri, P Rigby, R Krumlauf, S Brenner. 1995. Detecting conserved regulatory elements with the model genome of the Japanese puffer fish, *Fugu rubripes*. Proc. Natl. Acad. Sci. USA **92:** 1684-1688.

Baxendale S, S Abdulla, G Elgar, D Buck, G Micklem, R Durbin, G Bates, S Brenner, S Beck, H Lehrach. 1995. Comparative sequence analysis of the human and pufferfish Huntingtons disease genes. Nature Genet. **10:** 67-75.

Breathnach R, C Benoist, K Ohare, F Gannon, P Chambon. 1978. Ovalbumin gene: evidence for a leader sequence in mRNA and DNA sequences at the exon-intron boundaries. Proc. Natl. Acad. Sci. USA **75:** 4853-4857.

Breathnach R, P Chambon. 1981. Organization and expression of eukaryotic split genes coding for proteins. Annu. Rev. Biochem. **50:** 349-383.

Brenner S, G Elgar, R Sandford, A Macrae, B Venkatesh, S Aparicio. 1993. Characterization of the pufferfish (*Fugu*) genome as a compact model vertebrate genome. Nature **366:** 567-571.

Chang MS, GD Chang, JH Leu, FL Huang, CK Chou, CJ Huang, TB Lo. 1996. Expression, characterization and genomic structure of carp JAK1 kinase gene. DNA and Cell Biol. **15:** 827-844.

Chen SN, GH Kou. 1986. Establishment, characterization and application of 14 cell lines from warm-water fish. *In* E Kurstak, Y Kuroda, eds. Invertebrate and fish tissue culture. Tokyo: Japan Societies Press, pp. 218-227.

Curran T, WP MacConnell, F van Straaten, IM Verma. 1983. Structure of the FBJ murine osteosarcoma virus genome molecular cloning of its associated helper virus and the cellular homolog of the v-*fos* gene from mouse and human cells. Mol. Cell. Biol. **3:** 914-921.

Elgar G, F Rattray, J Greystrong, S Brenner. 1995. Genomic structure and nucleotide sequence of the p55 gene of the puffer fish *Fugu rubripes*. Genomics **27:** 442-446.

Faisst S, S Meyer. 1992. Compilation of vertebrate-encoded transcription factors. Nucleic Acids Res. **20:** 3-26.

Fujiwara KT, K Ashida, H Nishina, H Iba, N Miyajima, M Nishizawa, S Kawai. 1987. The chicken c-*fos* gene: cloning and nucleotide sequence analysis. J. Virol. **61:** 4012-4018.

Gonzalez GA, MR Montminy. 1989. Cyclic AMP stimulates somatostatin gene transcription by phosphorylation of CREB at serine 133. Cell **59:** 675-680.

Gorman C, R Padmanabhan, BH Howard. 1983. High efficiency DNA-mediated transformation of primate cells. Science **221:** 551-553.

Grange T, J Roux, G Rigaud, R Pictet. 1991. Cell-type specific activity of two glucocorticoid responsive units of rat tyrosine aminotransferase gene is associated with multiple binding sites for C/EBP and a novel liver-specific nuclear factor. Nucleic Acids Res. **19:** 131-139.

Grigoriadis AE, ZQ Wang, MG Cecchini, W Hofstetter, R Felix, HA Fleisch, EF Wagner. 1994. c-Fos: a key regulator of

osteoclast-macrophage lineage determination and bone remodeling. Science **266:** 443-448.

Herbomel P, B Bourachot, M Yaniv. 1984. Two distinct enhancers with different cell specificities coexist in the regulatory region of Polyoma. Cell **39:** 653-662.

Hinegardner R, DE Rosen. 1972. Cellular DNA content and the evolution of teleostean fishes. Am. Naturalist **106:** 621-644.

Jackson SP, JJ MacDonald, S Lees-Miller, R Tjian. 1990. GC box binding induces phosphorylation of Sp1 by a DNA-dependent protein kinase. Cell **63:** 155-165.

Johnson RS, BM Spiegelman, V Papaioannou. 1992. Pleiotropic effects of a null mutation in the c-*fos* proto-oncogene. Cell **71:** 577-586.

Lee W, P Mitchell, R Tjian. 1987. Purified transcription factor AP-1 interacts with TPA-inducible enhancer elements. Cell **49:** 741-752.

Macrae AD, S Brenner. 1995. Analysis of the dopamine receptor family in the compact genome of the puffer fish *Fugu rubripes*. Genomics **25:** 436-446.

Mason PJ, DJ Stevens, L Luzzatto, S Brenner, S Aparicio. 1995. Genomic structure and sequence of the *Fugu rubripes* glucose-6-phosphate dehydrogenase gene (G6PD). Genomics **26:** 587-591.

Mullis KB, FA Faloona. 1987. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. Methods Enzymol. **155:** 335-350.

Murre C, PS McCaw, D Baltimore. 1989. A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. Cell **56:** 777-783.

Orkin SH. 1992. GATA-binding transcription factors in hematopoietic cells. Blood **80:** 575-581.

Ruther U, D Komitowski, FR Schubert, EF Wagner. 1989. c-Fos expression induces bone tumors in transgenic mice. Oncogene **4:** 861-865.

Sambrook J, EF Fritsch, T Maniatis. 1989. Molecular cloning: a laboratory manual. New York: Cold Spring Harbor Lab. Press.

Sanger F, S Nicklen, AR Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:** 5463-5467.

Sassone-Corsi P, JC Sisson, IM Verma. 1988. Transcriptional autoregulation of the proto-oncogene *fos*. Nature **334:** 314-319.

Trower MK, SM Orton, IJ Purvis, P Sanseau, L Riley, C Cristodoulou, D Burt, CG See, G Elgar, R Sherrington, EI Rogaev, PS George-Hyslop, S Brenner, CW Dykes. 1996. Conservation of synteny between the genome of the pufferfish (*Fugu rubripes*) and the region on human chromosome 14 (14q24.3) associated with familial Alzheimer disease (AD3 locus). Proc. Natl. Acad. Sci. USA **93:** 1366-1369.

van Beveren C, F van Straaten, T Curran, R Muller, IM Verma. 1983. Analysis of FBJ-MuSV provirus and c-*fos* (mouse) gene reveals that viral and cellular *fos* gene products have different carboxy termini. Cell **32:** 1241-1255.

van Straaten F, R Muller, T Curran, C van Beveren, IM Verma. 1983. Complete nucleotide sequence of a human c-*fos* gene: deduced amino acid sequence of the human c-Fos protein. Proc. Natl. Acad. Sci. USA **80:** 3183-3187.

Verma IM, P Sassone-Corsi. 1987. Proto-oncogene *fos*: complex but versatile regulation. Cell **51:** 513-514.

Wang WW, RD Howells. 1994. Sequence of the 5-flanking region of the rat c-*fos* proto-oncogene. Gene **143:** 261-264.

Wasylyk B, SL Hahn, A Giovane. 1993. The Ets family of transcription factors. Eur. J. Biochem. **211:** 7-18.

Yoda T, T Morita, K Kawatsu, K Sueki, T Shibata, Y Hamano. 1995. Cloning and sequencing of the chaperonin-encoding *Cctd* gene from *Fugu rubripes*. Gene **166:** 249-253.

圓斑河豚 (*Tetraodon nigroviridis* [Syn. *T. fluviatilis*]) c-*fos* 基因結構與起動子活性分析

張茂山[1] 呂健宏[2] 姚振文[3] 黃火鍊[1] 黃銓珍[2]

圓斑河豚屬於豚形目四齒豚科，廣泛分佈於東南亞及東印度的河口區域，在臺灣容易自水族館購得，其基因體大小約爲 380 Mb，只有人類基因體大小的八分之一，適於從事基因體結構分析與比較。

在鯉魚蛋白激酶 *JAK1* 基因的啟動子區域有兩處轉錄因子 AP1 結合的核酸序列，爲了研究 AP1 如何調控 *JAK1* 基因的轉錄，我們便著手魚類 c-*fos* 基因的分子選殖。在本報告中，我們自圓斑河豚染色基因體基因庫篩選得到 c-*fos* 基因並定其基因結構。實驗的結果顯示圓斑河豚 c-*fos* 基因共包含 4 個外顯子和 3 個內隱子，可轉譯一個含 374 個胺基酸的蛋白質，其分子量約爲 42 kd ，其胺基酸序列與虎河豚 (*Fugu rubripes*)、人類、小白鼠與雞的相同度分別爲 90%、58%、58% 與 53%。我們利用引子延展分析決定了兩個可能的轉錄起始點，並且在圓斑河豚 c-*fos* 基因 0.75 kb 的啟動子區域中也發現了 CREB、E2A、Ets-1、GATA-1、HNF-5、c-Myb 及 Sp1 等轉錄因子的核酸結合序列。另外在氯黴素乙醯轉換酵素活性分析的實驗中，我們發現在圓斑河豚 c-*fos* 基因 −746 到 +153 的啟動子區域具有較對照組 pRSV-CAT 高出 2.7 倍的氯黴素乙醯轉換酵素活性，而此啟動子區域在 5% 胎牛血清的刺激下，其酵素活性亦明顯增加，證明圓斑河豚 c-*fos* 基因 −746 到 +153 的區域是一有功能的啟動子區域。

關鍵詞：基因結構, 起動子, 圓斑河豚 (*Tetraodon nigroviridis*)。

[1] 國立臺灣大學動物系

[2] 中央研究院生物化學研究所

[3] 國防醫學院生命科學研究所及預防醫學研究所