Characterization of the pufferfish *Otx2* cis-regulators reveals evolutionarily conserved genetic mechanisms for vertebrate head specification

Chiharu Kimura-Yoshida[1], Kuniko Kitajima[1], Izumi Oda-Ishii[1], E Tian[2], Misao Suzuki[3], Masayuki Yamamoto[4], Tohru Suzuki[5], Makoto Kobayashi[4], Shinichi Aizawa[6] and Isao Matsuo[1,*]

[1]Head Organizer Project, Vertebrate Body Plan Group, RIKEN Center for Developmental Biology, 2-2-3 Minatojima Minami Cho, Chuou-Ku, Kobe, Hyougo 650-0047, Japan
[2]Department of Morphogenesis, Institute of Molecular Embryology and Genetics, Kumamoto University, 2-2-1 Honjo, Kumamoto 860-0811, Japan
[3]Division of Transgenic Technology, Center for Animal Resources and Development, Kumamoto University, 2-2-1 Honjo, Kumamoto 860-0811, Japan
[4]The Center for Tsukuba Advanced Research Alliance and Institute of Basic Medical Sciences, University of Tsukuba, Tsukuba 305-8577, Japan
[5]Nutrition Division, National Research Institute of Aquaculture, Fisheries Research Agency, Nansei-cho, Watarai-gun, Mie 516-0193, Japan
[6]Animal Resources and Genetic Engineering Team and Vertebrate Body Plan Group, RIKEN Center for Developmental Biology, 2-2-3 Minatojima Minami Cho, Chuou-Ku, Kobe, Hyougo 650-0047, Japan
*Author for correspondence (e-mail: isao@cdb.riken.go.jp)

Accepted 23 September 2003

Development 131, 57-71
Published by The Company of Biologists 2004
doi:10.1242/dev.00877

Summary

The *Otx2* gene, containing a highly conserved *paired*-type homeobox, plays a pivotal role in the development of the rostral head throughout vertebrates. Precise regulation of the temporal and spatial expression of *Otx2* is likely to be crucial for proper head specification. However, regulatory mechanisms of *Otx2* expression remain largely unknown. In this study, the *Otx2* genome of the puffer fish *Fugu rubripes*, which has been proposed as a model vertebrate owing to its highly compact genome, was cloned. Consistently, *Fugu Otx2* possesses introns threefold smaller in size than those of the mouse *Otx2* gene. *Otx2* mRNA was transcribed after MBT, and expressed in the rostral head region throughout the segmentation and pharyngula periods of wild-type *Fugu* embryos. To elucidate regulatory mechanisms of *Otx2* expression, the expression of *Otx2-lacZ* reporter genes nearly covering the *Fugu Otx2* locus, from –30.5 to +38.5 kb, was analyzed, by generating transgenic mice. Subsequently, seven independent cis-regulators were identified over an expanse of 60 kb; these regulators are involved in the mediation of spatiotemporally distinct subdomains of *Otx2* expression. Additionally, these expression domains appear to coincide with local signaling centers and developing sense organs. Interestingly, most domains do not overlap with one another, which implies that cis-regulators for redundant expression may be abolished exclusively in the pufferfish so as to reduce its genome size. Moreover, these cis-regions were also able to direct expression in zebrafish embryos equivalent to that observed in transgenic mice. Further comparative sequence analysis of mouse and pufferfish intergenic regions revealed eight highly conserved elements within these cis-regulators. Therefore, we propose that, in vertebrate evolution, the *Otx2* promoter acquires multiple, spatiotemporally specific cis-regulators in order to precisely control highly coordinated processes in head development.

Key words: Homeobox, *Otx2*, Pufferfish, Gene regulation, cis-region, Brain development, Evolution, Local signaling center

Introduction

The establishment of the body plan of the vertebrate embryo requires the expression of regulatory genes in precisely controlled spatial domains, such that their products can function appropriately in the patterning process. Elucidation of the underlying mechanisms involved in this regulatory cascade initially depends on the identification of cis-acting regions that mediate spatial and temporal regulation. Homeobox genes are thought to participate in such developmental processes, thereby providing an ideal system with which to examine the genetic regulatory mechanism. Notably, Hox genes play a crucial role in the body plan of the trunk region, caudal to the hindbrain (Krumlauf, 1994). An intriguing co-linearity exists between the order of vertebrate Hox genes on the chromosome, and their temporal and spatial expression (Krumlauf, 1994). Much is known regarding the genetic regulatory mechanism of Hox expression (Nonchev et al., 1997; Duboule, 1998). However, in the head region, the regulatory mechanism of homeobox gene expression remains uncertain (Logan et al., 1993; Kammandel et al., 1999; Zerucha et al., 2000).

The vertebrate head is a complex structure, which is primarily composed of the skull, specific sense organs, muscles, glands, skin and the brain, which can be further subdivided into distinct and complex domains. Therefore, vertebrate head development requires highly coordinated processes. An understanding of the genetic mechanisms that control regional specification and morphogenesis of the embryonic rostral brain has only recently come to light; homeobox genes, Fgf, Wnt and other signaling molecules play an important role in these processes (reviewed by Rubenstein et al., 1998). Among these genes, *Otx2*, a *paired*-type homeobox gene, has been isolated as a mouse cognate of *orthodenticle* in *Drosophila* (Finkelstein and Perrimon, 1991; Simeone et al., 1992). The protein sequence of the homeodomains, and their patterns of expression, have all been highly conserved during evolution of the animal kingdom. Notably, mouse *Otx2* expression occurs earliest in the inner cell mass at the blastocyst stage, and in the visceral endoderm and epiblast at the prestreak stage (Acampora et al., 1995; Ang et al., 1996; Kimura et al., 2000; Kimura et al., 2001). During midstreak stages, the expression domain is restricted to the anterior region in all three germ layers corresponding to the presumptive rostral head (Simeone et al., 1993; Ang et al., 1994). At the pharyngula stage, *Otx2* expression is present in developing sense organs, such as the eyes, olfactory epithelium and the inner ear, as well as in the prospective forebrain and midbrain regions (Simeone et al., 1993; Mallamaci et al., 1996).

Proper regulation and function of this gene is crucial, as the mutation of *Otx2* leads to defects in the rostral head in several developmental processes (Acampora et al., 1995; Acampora et al., 1997; Matsuo et al., 1995; Ang et al., 1996; Suda et al., 1996; Suda et al., 1997; Suda et al., 2001; Tian et al., 2002; Hide et al., 2002). However, very little is known about how *Otx2* gene expression is dynamically regulated. The expression of homeobox genes is regulated primarily at the level of transcription. Accordingly, cis-elements and trans-factors that govern spatiotemporal *Otx2* expression may provide valuable insight into the genetic pathways that control anteroposterior axis formation, patterning of the head, and specification of the forebrain and midbrain. Previously, we identified cis-regulatory regions in the mouse *Otx2* 5′-flanking region, which are mainly responsible for the expression patterns in the head mesenchyme, anterior visceral endoderm and mesendoderm (Kimura et al., 1997; Kimura et al., 2000). However, we were not able to detect any cis-elements governing embryonic expression in the forebrain, midbrain and sense organs within the 80 kb mouse *Otx2* locus (–65 kb to +15 kb) (C.K.-Y. and I.M., unpublished). This finding suggests that cis-elements of mouse *Otx2* expression must lie outside of this surveyed 80 kb region.

To unravel the difficulty of identifying *Otx2* cis-regulatory regions, we took advantage of the compact genome of the pufferfish *Fugu rubripes* (*Fugu*), which displays a genome of 365 Mb, which is approximately eight times smaller than the human and mouse genomes (Brenner et al., 1993; Aparicio et al., 2002). The pufferfish genome has a similar number of genes and this fish shares a common body plan for specialized functions in higher vertebrates; consequently, it provides an ideal system with which to investigate conserved mechanisms of *Otx2* gene regulation. Coincidentally, it has been shown that several regulatory cis-elements have been functionally conserved between mammals and pufferfish (Marshall et al., 1994; Pöpperl et al., 1995; Aparicio et al., 1995; Kimura et al., 1997; Camacho-Hubner et al., 2000; Brenner et al., 2002).

In the present study, the *Fugu Otx2 (Fotx2)* gene was cloned. In addition, endogenous *Fotx2* expression was analyzed during *Fugu* embryogenesis. Subsequently, cis-acting regulators throughout the entire *Fotx2* genomic locus were surveyed by employing transgenic mice. Many distinct cis-regulators were identified, which directed temporally and spatially specific expression, mimicking endogenous *Otx2* expression. Additionally, these cis-regions possessed the ability to drive transgene expression in zebrafish embryos and contained evolutionarily conserved sequence elements. These results indicate that developmental domain-specific multiple cis-regulators control highly coordinated *Otx2* expression during vertebrate rostral head development.

Materials and methods

Cloning and sequencing of the *Fugu Otx2* (*Fotx2*) gene

Degenerate PCR primers were designed based on the amino acid sequences of the zebrafish and mouse *Otx2* genes. The conserved amino acid sequences GQNKV and WKFQV were utilized to synthesize the oligonucleotide primers PF1 (5′-GGAATTCGG-NCA(A/G)AA(T/C)AA(A/G)GT-3′) and PF2 (5′-TTGGATCCA-C(T/C)TG(A/G)AA(T/C)TTCCA-3′), respectively. A 590 bp PCR product was cloned into pBluescript (KS–) and sequenced. The *Fugu* genomic phage library (Kimura et al., 1997) was probed with the 590 bp *Fugu Otx2* fragment described above. DNA sequence determination of the 3.5 kb *Fugu Otx2* genomic region was performed by sequencing of both strands of double-stranded plasmid DNA.

Expression analysis of *Fotx2* by RT-PCR and in situ hybridization

Wild pufferfish embryos were maintained at 18°C, and sacrificed at subsequent developmental stages for RNA preparation and in situ hybridization as described (Suzuki et al., 2002). Primers for *Fotx2* expression, 5′-TTACGCGCGCCCAGTTAGACGTTTTGGAGG-3′ and 5′-GACGCCGGGGACTGGTTCAGATGGCTTGTG-3′, afforded a 586 bp product. Primers for β-actin expression, 5′-GGCACC-GCTGCCTCCTC-3′ and 5′-GCCAGGATGGAGCCTCC-3′, yielded a 359 bp product. Whole-mount in situ hybridization was conducted according to the method of Jowett and Lettice (Jowett and Lettice, 1994).

Transgene construction of the *Fotx2* gene

All constructs used in this study were generated using standard molecular cloning techniques (Sambrook and Russell, 2001). Generation of construct *F0placZ* was described previously, as construct #11 (Kimura et al., 1997). Construct *F1placZ* was generated by insertion of a 5′ 4.0 kb *Eco*RV-*Nsp*V genomic fragment, blunted by T4 DNA polymerase, into a *Sma*I site of construct *F0placZ*. Construct *F2placZ* was produced by insertion of a 5′ 2.5 kb *Eco*RV-*Bam*HI genomic fragment, blunted by T4 DNA polymerase, into a *Sma*I site of construct *F0placZ*. Construct *F3placZ* was generated by cloning of a 5′ *Sac*II-*Bam*HI genomic fragment into the *Sac*II and *Bgl*II sites of construct *F0placZ*. Construct *F4placZ* was generated by insertion of a 5′ 4.0 kb *Not*I-*Bam*HI genomic fragment into *Not*I and *Bgl*II sites of construct *F0placZ*. Construct *F5placZ* was generated by insertion of a 5′ 11.0 kb *Bam*HI-*Bam*HI genomic fragment into a *Bgl*II site of construct *F0placZ*. Construct *F6placZ* was generated by insertion of a 5′ 1.5 kb *Not*I-*Bam*HI genomic fragment into *Not*I and *Bgl*II sites of construct *F0placZ*. Construct *F7placZ* was produced by insertion of a 3′ 2.5 kb *Bst*EII-*Bam*HI genomic fragment blunted by T4 DNA polymerase into a *Sma*I site of construct *F0placZ*. Construct

F8placZ was generated by insertion of a 3′ 8.0 kb *Xba*I-*Bam*HI genomic fragment into *Xba*I and *Bgl*II sites of construct *F0placZ*. *F9placZ* was generated by insertion of a 3′ 5.0 kb *Xba*I-*Bam*HI genomic fragment into the *Xba*I and *Bgl*II sites of construct *F0placZ*. Construct *F10placZ* was produced by insertion of a 3′ 9.0 kb *Bam*HI-*Not*I genomic fragment into *Not*I and *Bgl*II sites of construct *F0placZ*. Construct *F11placZ* was generated by insertion of a 3′ 5.5 kb *Not*I-*Bam*HI fragment into *Not*I and *Bam*HI sites of construct *F0placZ*. Construct *F12placZ* was generated by insertion of a 3′ 6.3 kb *Bam*HI-*Bam*HI fragment into a *Bgl*II site of construct *F0placZ*. Construct *F13placZ* was produced by insertion of a 4.8 kb *Not*I-*Bam*HI fragment into *Not*I and *Bam*HI sites of construct *F0placZ*.

In order to construct the *F3hsplacZ* vector, a 2.4 kb *Sma*I-*Eco*RV fragment of *hsp-lacZ* (Kothary et al., 1989) was first inserted into *Sma*I and *Eco*RV sites of construct *pBSlacZIII* (Kimura et al., 1997), yielding the *pBShsplacZ* vector. Construct *F3hsplacZ* was then generated by cloning of a 5′ *Sac*II-*Bam*HI genomic fragment into the *Sac*II and *Bgl*II sites of construct *pBShsplacZ*. In order to construct *F3pGFP* and *F8pGFP* transgenes, the sequence of the translational start site of the *lacZ* gene in *F3placZ* and *F8placZ* was converted into a *Nco*I linker sequence by PCR-based mutagenesis, yielding *F3placZ(Nco)* and *F8placZ(Nco)*, respectively. Subsequently, a 740 bp *Nco*I-*Eco*RI fragment of the plasmid *pEGFP* (Clontech) was inserted between the *Nco*I and *Eco*RI sites of the *F3placZ(Nco)* and *F8placZ(Nco)* vectors, respectively. Construct *E6/7dplacZ* was produced by insertion of three genomic fragments, the 660 bp *Not*I-*Cla*I (blunted), the 960 bp *Bsg*I (blunted)-*Sph*I (blunted) and the 916 bp *Sap*I (blunted)-*Bam*HI, into *Not*I and *Bam*HI sites of construct *F0placZ*.

Production and genotyping of transgenic mice

Transgenic mice were generated by microinjection of fertilized eggs from CD-1, as described by Hogan et al. (Hogan et al., 1994). Transgenic mouse embryos were identified by PCR analysis as previously described (Kimura et al., 1997).

β-Gal staining and histological analysis of embryos

β-Gal staining was conducted as described previously (Kimura et al., 1997). For whole-mount views of 10.5 dpc to 13.5 dpc embryos, specimens were washed with PBS and dehydrated in graded ethanol. Specimens were then cleared in 1:2 benzyl alcohol/benzoate as described (Eng et al., 2001). For histological analysis, embryos were embedded in paraplast. Serial sections (10 μm) were produced and stained with Eosin.

GFP reporter analysis using zebrafish embryos

Zebrafish embryos were obtained by natural mating (Westerfield, 1995) and were staged accordingly (Kimmel et al., 1995). *F3pGFP* and *F8pGFP* constructs were linealized by digestion with *Sac*II and *Sal*I, respectively. Digested DNA was resuspended in water and injected into the blastomere of early one-cell stage embryos as described (Kobayashi et al., 2001). Embryos were fixed in PBS containing 4% paraformaldehyde, and examined under GFP-plus filters on a MZFLIII microscope (Leica) equipped with a C5810 chilled CCD camera (Hamamatsu Photonics).

Results

Isolation and characterization of the *Fugu Otx2* gene (*Fotx2*)

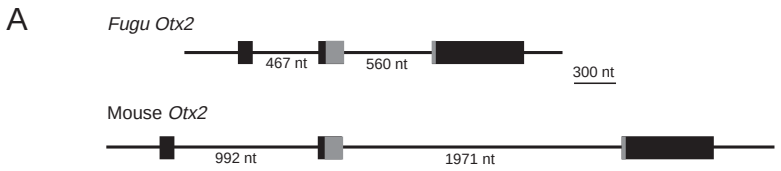
In order to identify the *Fugu Otx2* (*Fotx2*) gene, PCR of *Fugu* genomic DNA was used to isolate a 590 bp fragment of the *Fotx2* gene (see Materials and methods). Consequently, two genomic DNA clones, which contained the coding region and overlapped one another, were obtained from a *Fugu* genomic phage library (Kimura et al., 1997) (Fig. 3). A 3.5 kb *Pst*I-*Bam*HI fragment hybridized with the 590 bp probe was used as the source of the complete DNA sequence of the *Fotx2* gene, which is available at GenBank (Accession Number AY 303542). The overall size of the *Fotx2* gene, from initiation to termination codons, is 1901 bp, which is approximately twofold smaller than the mouse gene (Fig. 1A) (Simeone et al., 1992) (http://www.ensembl.org/Mus_musculus/). The coding regions of the *Fotx2* gene, consisting of three exons, were nearly identical in size, 94, 152 and 624 nucleotides, respectively, to those in the mouse *Otx2* gene (Fig. 1A). By contrast, the first and second introns (467 and 560 nucleotides, respectively) of *Fotx2* are approximately threefold smaller than those of the mouse (Fig. 1A).

The general structure of the amino acid sequence of *Fotx2* is quite similar to the sequences of previously described members of the *Otx* family; each of these species shares a short N-terminal domain, a *paired*-type homeodomain and a long C terminus, including two repeats of the Otx tail motif (Fig. 1B) (Furukawa et al., 1997). Notably, *Fotx2* exhibits 96% homology with zebrafish *Otx2*, 93% homology with mouse *Otx2* and 90% homology with *Xenopus Otx2*, throughout the entire peptide sequence (Fig. 1B). Remarkably, the homeodomain sequence is fully conserved but for a single amino acid change at position 52 in *Xenopus Otx2* (Fig. 1B). Further phylogenetic analyses have indicated that *Fotx2* is closely related to other vetebrate *Otx2* genes; however, these data reveal that *Fotx2* is not closely related to either *Otx1* or *Crx* genes (Fig. 1C).

Expression of the *Otx2* gene during pufferfish development

In order to investigate *Fotx2* expression during development, RT-PCR analysis was performed in wild-type *Fugu* embryos. The developmental stage of *Fugu* is substantially similar to that of other teleosts, such as zebrafish (data not shown). *Fotx2* mRNA expression was not detected at 22 hours postfertilization (hpf), which is equivalent to the dome stage in zebrafish. However, it was observed initially as early as 30 hpf, equivalent to the shield stage in zebrafish (Fig. 2A). These data indicate that *Fotx2* mRNA is not present in the oocyte and, furthermore, that it is zygotically transcribed following midblastula transition. Subsequently, *Fotx2* transcripts were significantly upregulated around 54 hpf, an early somites stage (Fig. 2A). These transcripts were detected throughout the segmentation and pharyngula periods; additionally, transcripts were detected at 172 hpf when embryos had been hatched (Fig. 2A).

To examine *Fotx2* expression during embryogenesis more precisely, whole-mount in situ hybridization of *Fugu* embryos was performed. We found *Fotx2* expression in the developing rostral head (Fig. 2C-E). *Fotx2* expression was apparent in the anterior portion of the neural plate at 44 hpf, which is equivalent to the bud stage of zebrafish (Fig. 2C). In order to compare anteriorly localized *Fotx2* expression, pufferfish *no tail* expression was also examined (Fig. 2B). As expected, *no tail* expression was observed consistently at the opposite side to *Fotx2* expression; it was detected in the posterior portion of the body, including in the notochord and tail bud at 44 hpf (Fig. 2B). By 54 hpf stage, the 5-6 somites stage, *Fotx2* expression was elevated significantly in the prospective forebrain and midbrain regions (Fig. 2D). At the subsequent 78 hpf stage (18-



B

```
FrOtx2  MMSYLKQPPYTVNGLSLTTSGMDLLHPSVGYQATPRKQRRERTTFTRAQLDVLEALFAKT  60
DrOtx2  ------------------------------P-----------------------------  60
XlOtx2  ----------A-------A-------Q----P----------------I--------  60
MmOtx2  ----------A--------------------P-----------------------------  60

FrOtx2  RYPDIFMREEVALKINLPESRVQVWFKNRRAKCRQQQQQQQQNGSQNKVRPAKKKSSPTRE  120
DrOtx2  ------------------------------------------G---------------  120
XlOtx2  ---------------------------------------G------S---T--V--  120
MmOtx2  ---------------------------------------G-------------A--  120

FrOtx2  VSSESGASGQFTPPTSTTTVPAISTSTAPVSIWSPASISPLSDPLSTSSSCMQRSYPMTY  180
DrOtx2  A------------*-S-S------T-------------------------------  179
XlOtx2  ------T----S--*-S-S--V--S------------V----------------------  179
MmOtx2  ------T-------*-S-S--T-AS-S-------------------------------  179

FrOtx2  TQASSYSQGYAGSTSYFGGMDCGSYLAPMHHQLSGPGSTLSPMSTNAVTSHLNQSPASLS  240
DrOtx2  ----G---------------------T------T-----------------------P  239
XlOtx2  ----G--------------------S----------A-----G-----------VA--  239
MmOtx2  ----G---------------------T------P---A-----G---------------  239

FrOtx2  TQGYGTSGLGFNSTADCLDYKDQAASWKLNFNADCLDYKDQTSSWKFQVL           290
DrOtx2  -----A------------------S-----------------------           289
XlOtx2  S-A--A-S------*--------T-------------------------           288
MmOtx2  -----A-S------T--------T-------------------------           289
```

C

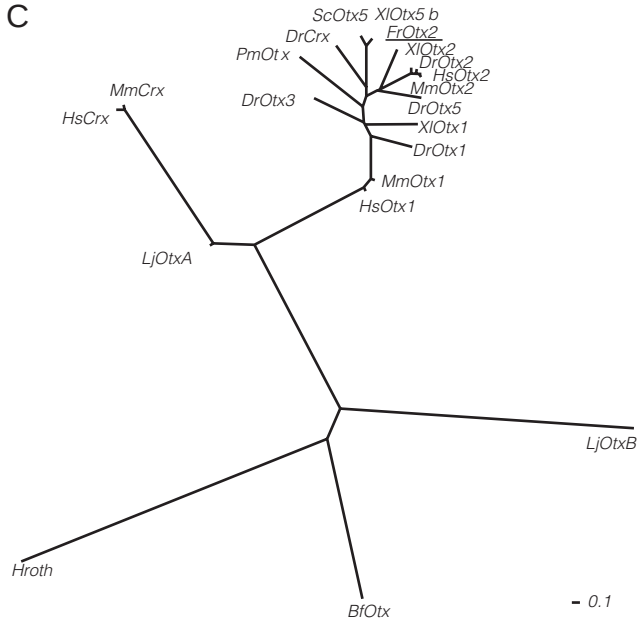


Fig. 1. Identity of the *Fugu rubripes Otx2* gene. (A) Comparison of *Fugu* and mouse *Otx2* genes. Filled boxes represent coding exons. Homeobox regions are shown in gray. (B) Comparison of the deduced amino acid sequences of the *Fugu Otx2* (*FrOtx2*), zebrafish *Otx2* (*DrOtx2*), *Xenopus Otx2* (*XlOtx2*) and mouse *Otx2* (*MmOtx2*). Homeodomains highly conserved between all *Otx2* genes are shaded. Otx protein sequences were retrieved from the GenBank database. Degenerate PCR primers are indicated by arrows and Otx tail motifs are underlined. (C) Phylogenetic tree of Otx/Crx family proteins. The tree was constructed using the NJ method. Protein sequences were retrieved from the GenBank database. Bf, amphioxus *Branchiostoma floridae*; Dr, zebrafish *Danio rerio*; Fr, *Fugu rubripes*; Hs, *Homo sapiens*; Lj, *Lampetra japonica*; Mm, *Mus musculus*; Pm, *Petromyzon marius*; Sc, dogfish *Scyliorhinus canicula*; Xl, *Xenopus laevis*.

20 somites stage), *Fotx2* expression was evident throughout the entire rostral head (Fig. 2E). These results indicate that pufferfish *Otx2* is expressed in the rostral head during *Fugu* embryogenesis.

Multiple and long-distance cis-regions combinatorially regulate the *Fotx2* locus in transgenic mice

In order to understand the cis-regulatory mechanism of *Fotx2* gene expression, the activity of reporter gene constructs harboring its genomic DNA were surveyed. First, five lambda phage contigs of genomic DNA from –30.5 kb to +38.5 kb (Fig. 3) were cloned and mapped. These contigs largely cover the entire *Fotx2* genomic locus (see Discussion). Next, 13 genomic fragments were tested for cis-activity via generation of transgenic mouse embryos carrying the reporter *lacZ* gene (Figs 3, 4). Transgene expression analysis was transiently conducted at 10.5 dpc; seven genomic fragments directing *lacZ* expression were identified (Figs 3, 4). All thirteen constructs contained a 2.4 kb *Fotx2* promoter, which displays no cis-activity during embryogenesis (Fig. 3; *F0placZ*) (Kimura et al., 1997). For seven of the 13 DNA fragments tested, *lacZ* expression occurred in spatial and temporal patterns that mimicked a subset of the normal *Otx2* expression pattern, with the exception of the F4 cis-region. This fine correlation suggests that these cis-regions can function independently of one another, and that their spacing relative to the promoter is not crucial for expression.

Most of the *Fotx2* cis-acting regions that we defined activated expression in multiple domains and at several developmental stages (Figs 3-11), suggesting that the *Fotx2* loci are composites of smaller more specific cis-regions. For example, the distant F3 region, spanning from –17 to –8.9 kb, directed transgene expression in the roof of the diencephalon and the medio-caudal telencephalon (Fig. 4A). The F4 region, which spans from –18 to –14 kb, directed *lacZ* expression in the premandibular, mandibular and hyoid arches (Fig. 4B; *F4placZ*). The further distant F5 region, which spans from –29 to –18 kb, governed *lacZ* expression in the lateral portion of the mesencephalon and the ventral diencephalons, including the zona limitans intrathalamica (ZLI) (Fig. 4C). The F8 region spanning +2.5 to +10.5 kb directed *lacZ* expression in the dorsal diencephalon and mediocaudal telencephalon (Fig. 4D; *F8placZ*). The F9

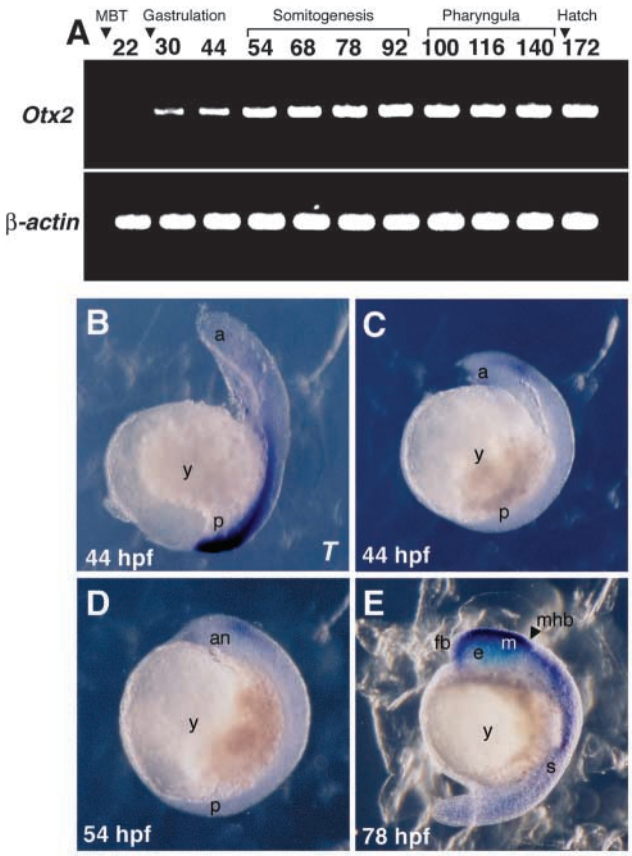


Fig. 2. Expression analysis of the *Fotx2* gene in wild-type pufferfish embryos. (A) Expression patterns of the *Fotx2* gene analyzed by RT-PCR. The *β-actin* fragment was amplified as a control for the quality of cDNA. (B-E) Expression patterns of the *Fugu* genes analyzed by whole-mount in situ hybridization. Pufferfish *no tail* expression at 44 hpf (B). *Fotx2* expression at 44 (C), 54 (D) and 78 hpf (E). MBT, midblastula transition; a, anterior; an, anterior neural tube; e, eye; fb, forebrain; m, mesencephalon; mhb, mid-hindbrain boundary; s, somites; p, posterior; y, yolk.

region (+10.5 to +15.5) governed *lacZ* expression in the cephalic mesenchyme (Fig. 4E; *F4placZ*) (Kimura et al., 1997). The further distant F11 region, spanning +22.8 to +26.5 kb, could drive *lacZ* expression in the trigeminal ganglions, cranial nerves, dorsal diencephalon, mesencephalon, rhombencephalon and spinal cord (Fig. 4F). The F12 region, harboring +26.5 to +32.8 kb, governed *lacZ* expression in the nasal pits, retina and first branchial groove (Fig. 4G).

Fotx2 cis-regions direct spatially and temporally distinct expression in the developing brain and sense organs

In order to characterize *Fotx2* cis-activity during embryogenesis more precisely, transgene expression was analyzed during embryogenesis. Each cis-acting region was found to drive transgene expression in a distinct subdomain of *Otx2* expression (Figs 5-9). Transgenic mice were established; each strain carried one transgene from *F3placZ*, *F5placZ*, *F8placZ*, *F11placZ* and *F12placZ*. Subsequently, their F1 or F2 hemizygous offspring were examined with respect to *lacZ* expression.

lacZ expression governed by *F3placZ* carrying the –17 to –8.9 kb region was observed initially in the restricted domain of the anterior neural plate at 8.5 dpc (Fig. 5A,B). At 9.5 dpc, *lacZ* expression was upregulated in the prospective dorsocaudal telencephalon and dorsal diencephalon (Fig. 5C). By 10.5 dpc, *lacZ* expression was further restricted in the diencephalic roof and dorsocaudal telencephalon (Fig. 5D,H). Histological analysis revealed that *lacZ* expression was present in the most caudal and medial aspects of the dorsal telencephalon, surrounding the lamina terminalis (Fig. 5E,F). At the level of the diencephalon, *lacZ* expression, which was evident in the most dorsal neural tube, occurred as two stripes excluding the dorsal-most roof (Fig. 5G). At 12.5 dpc, cis-activity remained in structures that developed into the choroid plexus; however, activity was not apparent in the cortical hem (Fig. 5I-L). In actuality, expression governed by the F3 cis-region appears to be closely related to Bmp and noggin expression (Fig. 5O) (Furuta et al., 1997; Hebert et al., 2002). This activity remained at 13.5 dpc (Fig. 5M,N).

In order to assess whether the F3 fragment can drive similar expression patterns for a heterogeneous promoter, the mouse *hsp68* promoter, which is widely used in transgenic mice, was employed in place of the pufferfish *Otx2* promoter (Kothary et al., 1989). Subsequently, we found that the F3 fragment was also able to direct expression under the control of the *hsp68* promoter in the diencephalic roof and the dorsocaudal telencephalon at 10.5 dpc (*n*=4/5; *F3hsp68lacZ*; Fig. 5P). This finding indicates that *lacZ* expression patterns are induced primarily by the F3 cis-region, and are not due to the *Fotx2* promoter activity itself.

lacZ activity driven by *F8placZ*, carrying the +2.5 to +10.5 kb region, was detected initially in the dorsal diencephalon and in restricted areas of the dorsocaudal telencephalon at 9.5 dpc (Fig. 6A,B). Subsequently, at 10.5 dpc, *lacZ* expression was upregulated in the dorsal diencephalon, including in the prospective epithalamus and the mediocaudal telencephalon (Fig. 6C,D). Further histological examination demonstrated that *lacZ* activity was present in the dorsomedial portion of the caudal telencephalon, corresponding to the edge of the embryonic cerebral cortex, which is referred to as the cortical hem (Fig. 6E-G) (Grove et al., 1998). From 9.5 to 10.5 dpc, the expression domains governed by the F8 cis-region appear to display partial overlap with those governed by the F3 cis-region. Later, at 12.5 and 13.5 dpc, *lacZ* activity was detected in both the diencephalon and the telencephalon, including the cortical hem; however, activity was absent in the choroid plexus, which is directed by *F3placZ* as described (Fig. 6H-K, Fig. 5). The expression pattern generated by the F8 cis-acting region appears to be similar to that of *Wnt3a* (Fig. 6L) (Grove et al., 1998; Lee et al., 2000).

lacZ expression governed by the F5 (–29.0 to –18.0 kb) cis-region was observed initially in the ventral portion of the diencephalon at 9.5 dpc (Fig. 7A). By 10.5 dpc, *lacZ* expression was detected in the ZLI, the ventral diencephalon and the lateral portion of the mesencephalon, with a longitudinal direction as described (see Fig. 7B,C). Further histological examination indicated activity, at 11.5 dpc, in the lateral mesencephalon, ZLI, and ventral portion of P2 and P3 diencephalon, including in the retromammillary region (Fig. 7D-G). Even at 13.5 dpc, *lacZ* expression remained in these structures (Fig. 7H,I).

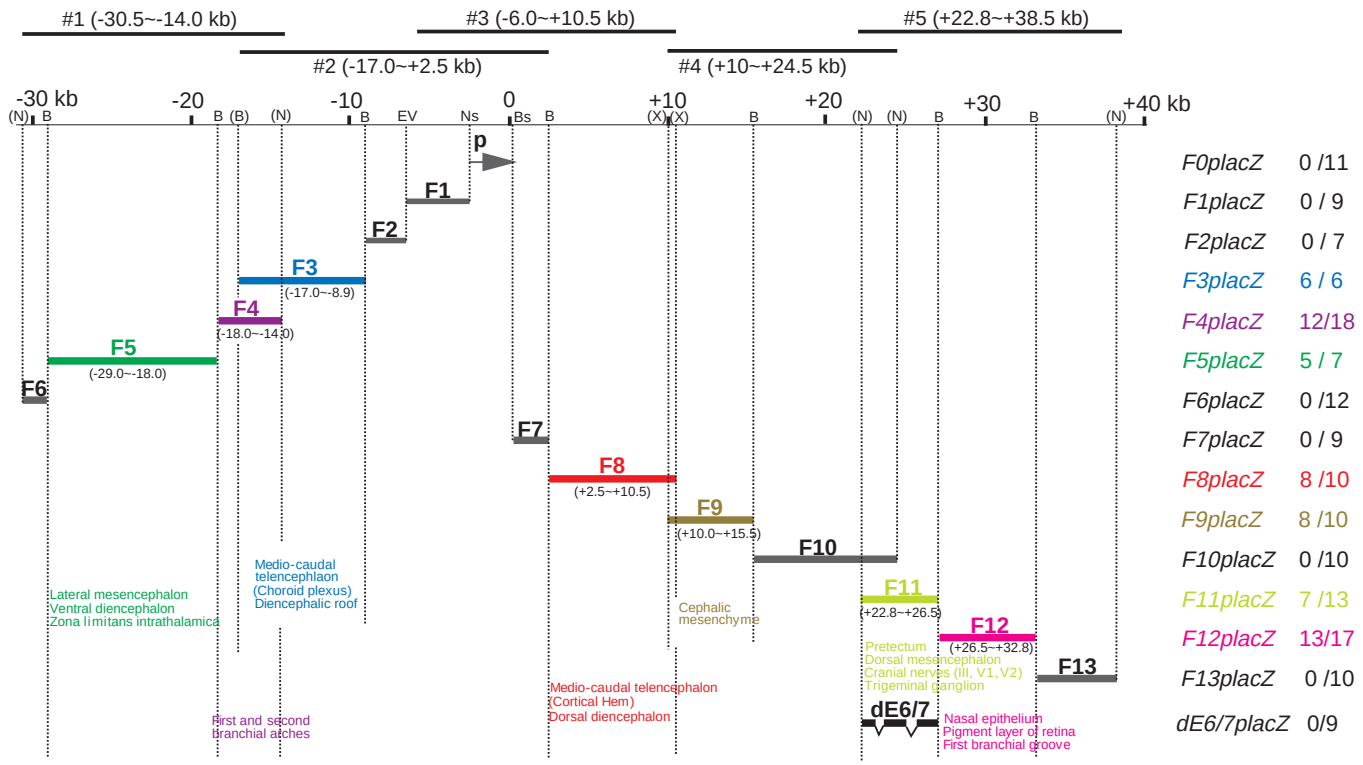


Fig. 3. Schematic diagram of the *lacZ* transgene constructs used to identify cis-acting regions throughout the *Fotx2* genomic locus. The translational start site is indicated (0). Each fragment examined for cis-activity is denoted by the bars marked F1-F13. Restriction enzyme sites used are indicated above. The colored genomic fragments consistently display specific *lacZ* activity at 10.5 dpc. By contrast, the fragments in black possess no cis-acting ability at 10.5 dpc. Tissues in which expression is driven by these fragments are indicated in colored letters below. The number of *lacZ*-positive embryos in the specific regions among transgenic embryos is indicated on the right with transgene names. In all these embryos, the patterns of expression were reliably identical although their levels were variable. B, *Bam*HI; Bs, *Bst*EII; EV, *Eco*RV; N, *Not*I; Ns, *Nsp*V; X, *Xba*I.

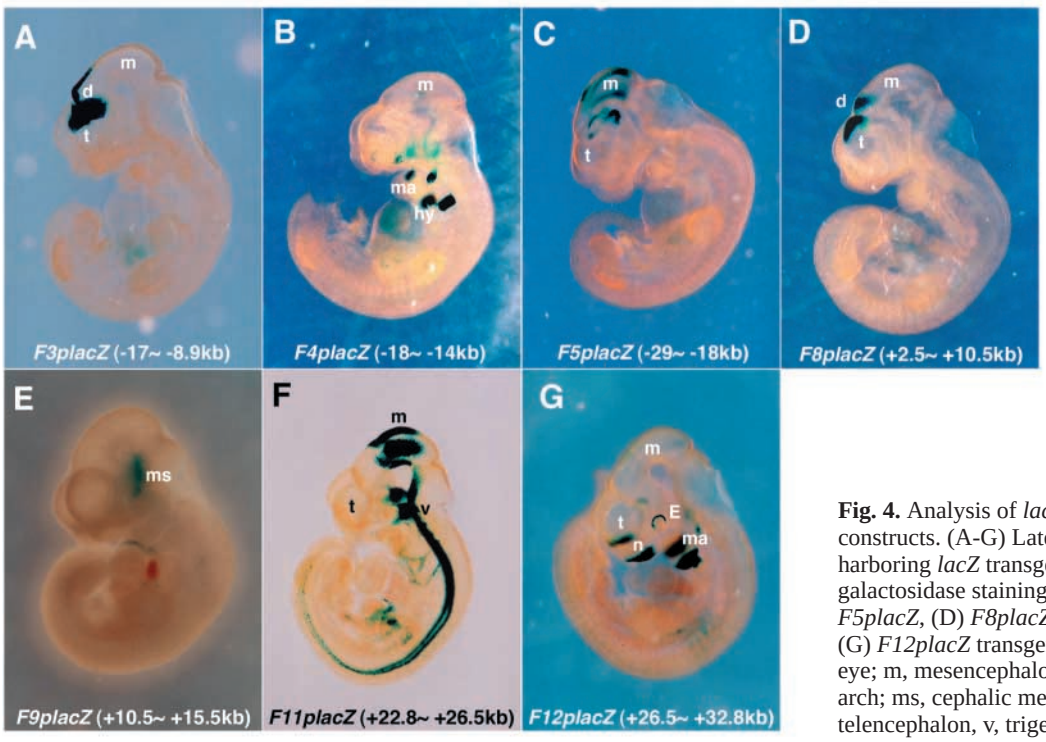


Fig. 4. Analysis of *lacZ* expression with seven transgene constructs. (A-G) Lateral views of 10.5 dpc embryos harboring *lacZ* transgene constructs, following β-galactosidase staining. (A) *F3placZ*, (B) *F4placZ*, (C) *F5placZ*, (D) *F8placZ*, (E) *F9placZ*, (F) *F11placZ* and (G) *F12placZ* transgenic embryos. d, diencephalon; E, eye; m, mesencephalon; hy, hyoid arch; ma, mandibular arch; ms, cephalic mesenchyme; n, nasal pits; t, telencephalon, v, trigeminal ganglion.

At 9.5 dpc, *lacZ* expression mediated by the F11 (–22.8 to –26.5 kb) cis-region was detected initially in the dorsal portion of the mesencephalic neuroepithelium, corresponding to the mesencephalic trigeminal ganglions, oculomotor nerves and a rostral aspect of trigeminal nerves (Fig. 8A). The active domains were extended both rostrally and caudally by 10.5 dpc in these structures (Fig. 8B,D-G). Additionally, expression was also evident in the dorsal region of the pretectum (Fig. 8B). With respect to branches of the trigeminal nerve, the F11 region drives expression exclusively in the opthalmic and maxillar branches, but not in the mandibular branch (Fig. 8B). Precise histological examination indicated that *lacZ* expression was present throughout the trigeminal ganglions and dorsal mesencephalic neural tube at 10.5 dpc (Fig. 8D-G). Subsequently, most *lacZ* expression in cranial nerves was diminished considerably by 11.5 dpc (Fig. 8C). Consequently, *lacZ* activity covered the dorsal pretectum, and the inferior and superior collicullus, by 13.5 dpc (Fig. 8H-J). Moreover, *lacZ* expression was observed consistently, from 9.5 dpc, in ventral domains of the neural tube at the level of the hindbrain, where Otx2 protein is also localized (Fig. 8A-C,E-G) (Mallamaci et al., 1996). Additionally, *lacZ* expression at the spinal cord level was observed ectopically in the transgenic embryos (Fig. 8A,B).

F12 (+26.5 to +32.8 kb) cis-regulatory activity was detected initially in the rostral first branchial groove and nasal portions as early as 9.5 dpc (Fig. 9A). *lacZ* activity in the retina commenced by 10.5 dpc (Fig. 4B,C). Subsequently, at 11.5 dpc, *lacZ* activity in the caudal portion of the mandibular arch extends to the cranial portion of the hyoid arch; moreover, activity was evident in both epithelium and mesenchyme, which corresponds to the auricular hillocks (Fig. 9E). This activity remained detectable in the external acoustic meatus of the ears at 12.5 dpc (Fig. 9F-I). Close histological examination revealed the presence of *lacZ* activity in the olfactory epithelium, as well as the in the outer layer of the optic cup corresponding to the prospective pigment layer of the retina at 12.5 dpc (Fig. 9J,K). Moreover, *lacZ* activity was observed in the developing inner ear, cochlea and saccule (Fig. 9H,I).

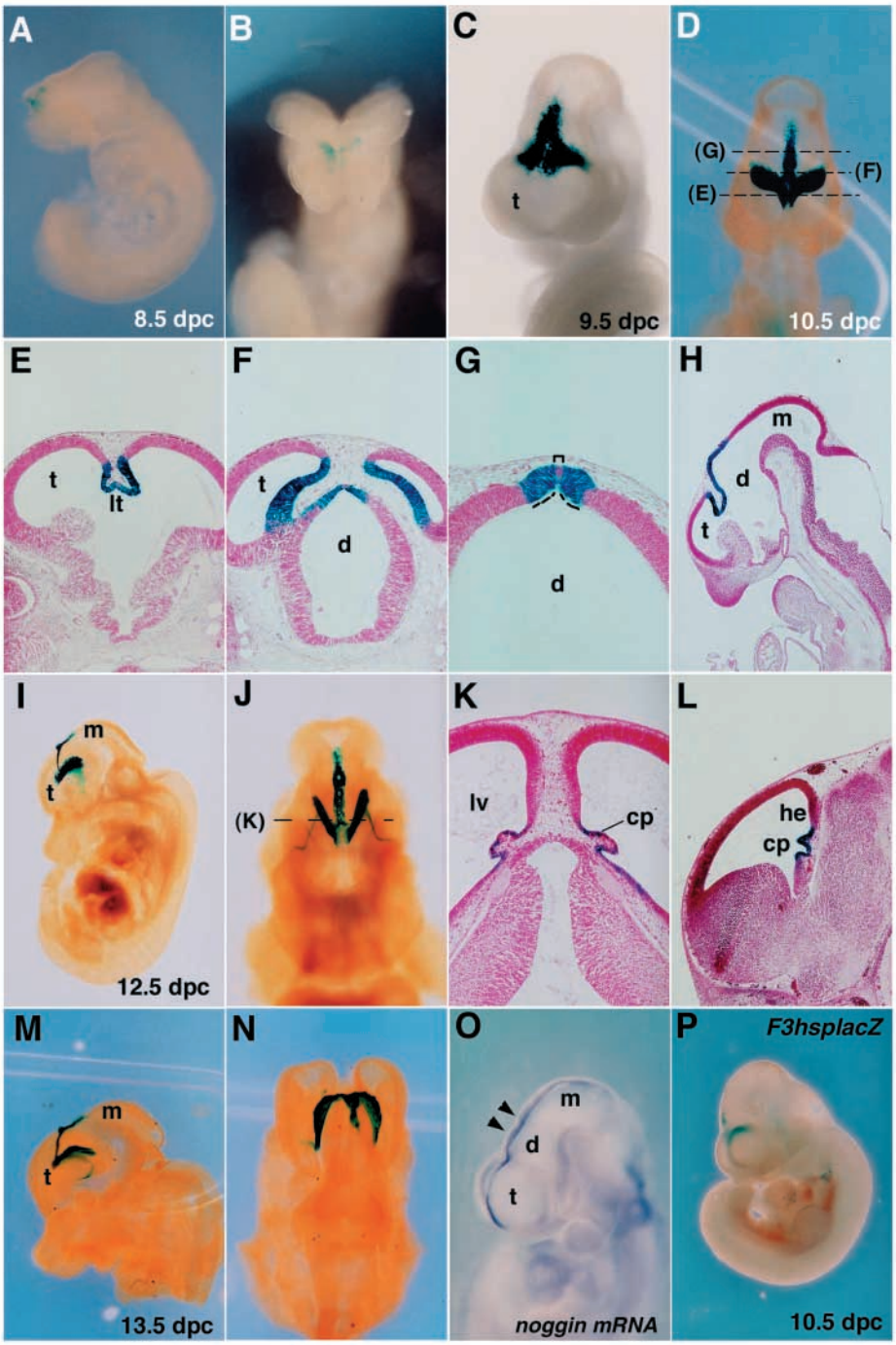


Fig. 5. Developmental changes in *lacZ* expression with the *F3placZ* construct. Lateral and frontal views of the *F3placZ* transgenic embryos at 8.5 (A,B), 9.5 (C), 10.5 (D), 12.5 (I,J) and 13.5 (M) dpc, and a dorsal view at 13.5 dpc (N), following β-galactosidase staining. Transverse sections of transgenic embryos at 10.5 (E-H) and 12.5 (K,L) dpc. Sagittal sections of embryos at 10.5 (H) and 12.5 (L) dpc. Whole-mount in situ hybridization analysis of noggin mRNA indicates that mouse noggin expression is evident in the roof of the neural tube, including at the level of the diencephalon (O, arrowheads). Lateral view of the *F3hsplacZ* transgenic embryo at 10.5 dpc, following β-galactosidase staining (P). cp, choroid plexus; d, diencephalon; he, cortical hem; lt, lamina terminalis; lv, lateral ventricle; m, mesencephalon; t, telencephalon.

These expression patterns indicate that the aforementioned cis-regulators, with the exception of F4, mediate *Otx2* expression in spatiotemporally distinct domains during rostral head development. Many cis-regulatory regions of genes have been shown to govern overlapping expression domains in a redundant manner; however, the seven cis-regulators currently identified displayed no redundant expression patterns (Figs 4-10). Furthermore, we concluded that the *Fotx2* cis-regions function in a domain-specific fashion, rather than in a developmental stage-specific fashion. Their activities most likely depend on the expression of position-specific transactivators, which are expressed over time, spanning several subdomains within a given lineage.

Fotx2 enhancers can drive transgene expression in zebrafish embryos

In order to test whether the cis-regulators identified here can direct transgene expression in fish, cis-activity in zebrafish embryos was investigated. Among the seven cis-regulators, F3 and F8 regions faithfully govern expression in distinct domains of the dorsal diencephalon and mediocaudal telencephalon (Figs 4-6, 10). Thus, the activity of these two cis-regulatory regions was transiently examined in transgenic zebrafish. All constructs were tested via injection of DNA into zygotes; subsequently, the resulting mosaic zebrafish embryos were analyzed for GFP activity. First, the *Fugu Otx2* promoter fused with the EGFP vector (*pGFP*) alone afforded no GFP signals throughout zebrafish embryogenesis (data not shown; *n*=0/74). The *F3pGFP* construct, which harbors the –17 to –8.9 kb genomic fragment fused with the *pGFP* vector, was initially expressed specifically at the anterior ectoderm, and in the posterior portion of zebrafish embryos ectopically from 80% epiboly onward (Fig. 11A). At early somites stage, GFP activity was observed specifically in the prospective forebrain region (Fig. 11B). Subsequently, at the 18 somites stage, GFP activity was consistently detected in the diencephalon and the telencephalon (Fig. 11C,D; *n*=38/51). At the later pharyngula, 32 hpf, stage, GFP activity remained in evidence in the dorsal diencephalon (Fig. 11E).

In contrast to the *F3pGFP* transgene, *F8pGFP*, which contains the –17 to –8.9 kb genomic fragment fused with the *pGFP* vector, was initially expressed from 12 to 14 hpf, and from the 6 to 10 somites stage (data not shown). Subsequently, GFP activity directed by *F8pGFP* was detected consistently in the diencephalon and in the telencephalon at 18 hpf (18 somites stage) (Fig. 11F; *n*=20/69). Later, at 30 hpf, an early pharyngula stage, *F8pGFP* expression remained in the telencephalon and the diencephalon (Fig. 11G). These expression patterns indicated that the activity of *F3pGFP* consistently occurs earlier, and is stronger and more widely distributed, than that of *F8pGFP*, suggesting that each fragment mediates distinct expression in the forebrain regions of the transient transgenic zebrafish embryos (Figs 5, 6, 11). These data indicate that the spatiotemporal expression governed by these two cis-regions is conserved between transgenic mouse and zebrafish, which suggests that an evolutionarily conserved mechanism controls these two cis-regions of *Otx2* gene.

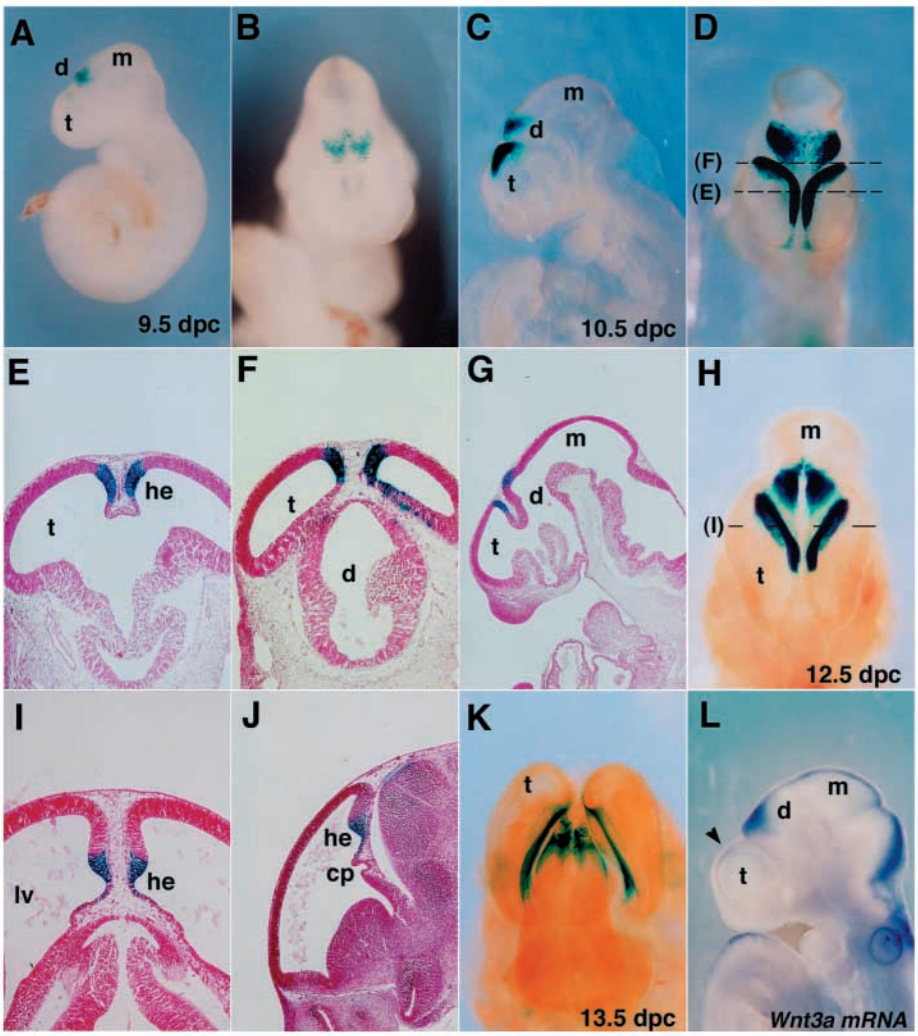


Fig. 6. Developmental changes in *lacZ* expression with the *F8placZ* construct. Lateral (A,C), frontal (B,D,H) and dorsal (K) views of *F8placZ* transgenic embryos following β-galactosidase staining. Transverse and sagittal sections of the transgenic embryos at 10.5 (E-G) and 12.5 (I,J) dpc. (A,B) 9.5, (C-G) 10.5, (H-J) 12.5 and (K) 13.5 dpc transgenic embryos. Whole-mount in situ hybridization analysis of *Wnt3a* mRNA indicates that mouse *Wnt3a* expression is detected in the prospective dorsal diencephalon and cortical hem (arrowhead; L). cp, choroid plexus; d, diencephalon; he, cortical hem; m, mesencephalon; t, telencephalon.

Identification and characterization of sequence elements conserved between mouse and pufferfish

Recent reports of large-scale sequence

analysis, combined with transgenic mouse approaches, reveal that evolutionarily conserved non-coding genomic sequence elements tend to control gene expression (Hardison, 2000). Thus, in order to identify highly conserved, putative regulatory sequence elements, a comprehensive comparative analysis of the intergenic regions flanking the mouse and pufferfish *Otx2* genes was performed by PipMaker analysis (Schwartz et al., 2000). Consequently, we identified eight highly conserved sequence elements within the pufferfish cis-regulators (Fig. 12). The F4, F9 and F12 cis-regions possess conserved elements, E4, E5 and E8, respectively. The F5 cis-region possesses three conserved elements, E1, E2 and E3. F11 contains two conserved regions, E6 and E7. In addition, the corresponding eight conserved elements in mouse were located in an order similar to that of pufferfish; however, these elements were within an approximately eightfold extended region up to 550 kb. These data further support the notion that genetic mechanisms of *Otx2* gene regulation are evolutionarily conserved among vertebrates. Moreover, as homologous regulatory elements present in different species represent direct descendants of a common ancestral regulatory element, individual cis-acting regions appear to have developed their distinct regulatory elements independently.

In order to determine whether these evolutionarily conserved elements play crucial roles with respect to cis-activity, a mutant F11 construct, in which both elements, E6 and E7, were specifically deleted (*E6/7dplacZ*) was generated (Fig. 3). The pufferfish E6 and E7 elements are located at the 3′ end of the *Otx2* gene, at +23.5 and +25 kb, respectively (Fig. 12B). The corresponding mouse E6 and E7 sequences, although also located at the 3′ end, are at +175 and +196 kb, respectively, which is distant from the *Otx2* coding region. Subsequently, *lacZ* expression of transient transgenic embryos carrying the *E6/7dplacZ* construct was analyzed as described. In contrast to *F11placZ* transgenic embryos, *lacZ* expression was not detected in *E6/7dplacZ* transgenic embryos at 10.5 dpc (n=0/9; Fig. 3, Fig. 12C,D). This finding demonstrates that conserved elements between mouse and pufferfish, E6 and E7, are essential for the transgene expression in mouse embryos. Moreover, this conserved function of regulatory elements between mouse and pufferfish suggests that the compact *Fugu* genome could be a potentially powerful and useful system for the rapid analysis of transgene expression.

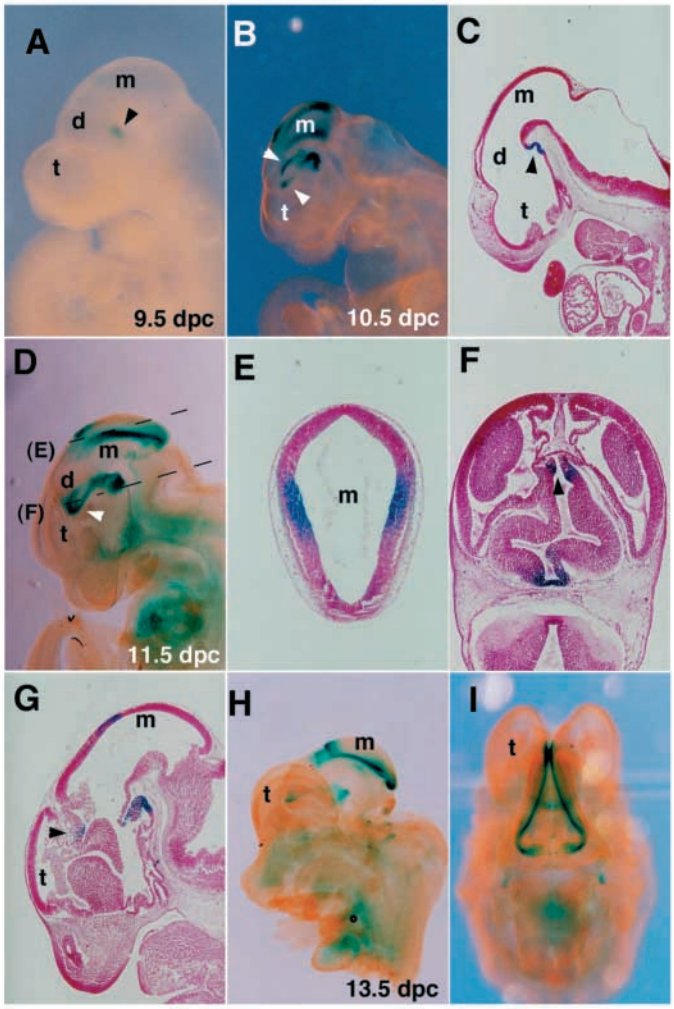


Fig. 7. Developmental changes in *lacZ* expression with the *F5placZ* construct. Lateral (A,B,D,H) and dorsal (I) views of the transgenic embryos following β-galactosidase staining. Sagittal (C,G) and transverse (E,F) sections of the transgenic embryos. (A) 9.5, (B,C) 10.5, (D-G) 11.5 and (H,I) 13.5 dpc transgenic embryos. *lacZ* expression is detected in the ventral diencephalon at 9.5 dpc (arrowhead; A). At 10.5 dpc, *lacZ* expression is present in the ZLI (white arrowheads; B). *lacZ* activity occurs in the mammillary recess (arrowhead; C). At 11.5 dpc, transgene activity remains in the lateral mesencephalon, ZLI (arrowheads) and the ventral diencephalon, including in the retromammillary region (D-G). d, diencephalon; m, mesencephalon; t, telencephalon.

Discussion

We have described the isolation and characterization of the pufferfish *Otx2* gene. Furthermore, we have identified multiple cis-regulators of the *Fotx2* locus in transgenic embryos. The current data indicate that *Fotx2* is regulated by distinct and physically separable cis-acting sequences that reside in a sequence of more than 60 kb surrounding the *Fotx2* coding region. Each transgene harboring different cis-acting regions is reproducibly expressed in subdomains of *Otx2* expression in mouse and fish embryos. Given the relatively simple pattern of *Otx2* expression, it was surprising to find such a complex set of regulatory elements controlling *Otx2* expression in the developing rostral head. A likely explanation regarding the complexity of regulation of *Otx2* expression is that each cis-regulator might mark a distinct domain related to local signaling centers and to the development of specific organs. Moreover, the expression patterns of *Fugu* cis-regions in transgenic zebrafish, and the presence of conserved sequence elements between the pufferfish and mouse, suggest that the genetic mechanism of vertebrate head specification is highly conserved between fish and mouse.

Fugu genome as a model system for the identification of cis-regulators

Fugu has the most compact genome among vertebrates. Comparative genomic analysis of human, mouse and pufferfish genome sequences identifies the coding region of the genome. However, the function of non-coding regions for gene expression remains elusive. The compact nature of the

pufferfish genome is primarily due to the smaller size of the non-coding region; on average there is one gene per 10 kb in the pufferfish genome (Aparicio et al., 2002). Therefore, the *Fugu* genome may provide an ideal system with which to analyse the function of non-coding regions; in particular, to analyse those regions with a pivotal role in gene expression. In the present investigation, in excess of 69 kb of the entire *Fotx2* locus (the *Otx2* locus occupies ~700 kb in the mouse genome) was examined. Coincidentally, cis-regulatory activities were not observed within the 70 kb spanning from –50 to +20 kb of the mouse *Otx2* locus, with the exception of the 5′ 1.8 kb promoter sequence (C.K.-Y. and I.M., unpublished). It is evident from these findings that cis-regions regulating *Otx2* expression are scattered over extremely large genomic distances from the coding region.

Coincidentally, genes immediately upstream and downstream of *Fotx2* include the Brain secretory protein Sec10 gene located at –38 kb, and the *pellino 2* gene located at +43 kb (http://fugu.hgmp.mrc.ac.uk/). Comparison with the mouse *Otx2* locus reveals that these immediate gene orders are conserved between pufferfish and mouse (http://www.ensembl.org/Mus_musculus/). By contrast, the mouse ortholog Brain secretory protein Sec10 gene is located approximately –350 kb, and *pellino 2* is located approximately +500 kb, downstream of the mouse *Otx2* locus. These genome sequence data indicate that the *Otx2* locus is an extremely gene-bare region when compared with the average of one gene per 10 kb in the pufferfish genome. Together, these suggest that the 69 kb *Fugu* sequence from –30.5 to +38.5 kb may encompass most of the cis-regulatory regions of the *Fugu Otx2* locus.

Consistently, the *Fugu* cis-acting regions we identified possess sequence elements that are highly conserved with those that are located within the 550 kb region of the mouse *Otx2* locus (Fig. 12). Moreover, the conserved elements E6 and E7 were shown to be crucial for F11 cis activity (Fig. 12D). Curiously, most cis-regulators identified in the *Fotx2* locus do not direct redundant expression patterns, which implies that cis-regulatory regions for redundant or overlapping expression may be exclusively abolished in the pufferfish so as to reduce its genome to as small a size as possible.

We have described pufferfish *Otx2* gene expression during wild *Fugu* embryogenesis. Developmental patterns of the *Fugu* teleost are essentially identical with those of other teleosts such as zebrafish and medaka. Accordingly, *Fotx2* expression appears to be substantially similar to zebrafish *Otx2* expression (Li et al., 1994); briefly, a common aspect of vertebrate *Otx2* gene expression in anterior region of the body is shared. Additionally, we have also identified multiple pufferfish cis-acting regions in transient transgenic zebrafish embryos. This finding implies that the cis-elements and the trans-acting factors, which are involved in *Otx2* expression, had fully evolved in the common ancestor of teleosts and mammals, and have been conserved during 400 million years of evolution.

Expression domains by *Fotx2* cis-regulators coincide with local signaling centers of the neuroectoderm

The F3, F5 and F8 cis-regulatory regions identified in this study direct transgene expression that nearly coincides with

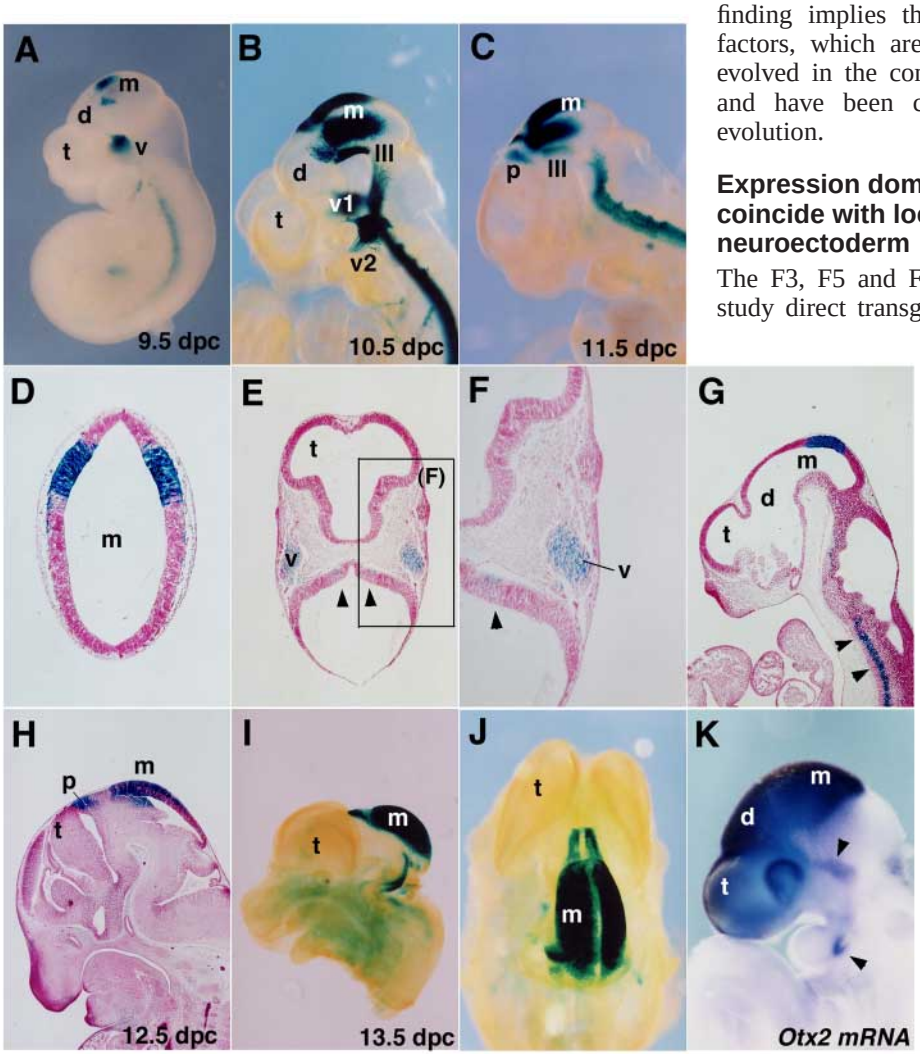


Fig. 8. Developmental changes in *lacZ* expression with the *F11placZ* construct. Lateral (A-C,I) and dorsal (J) views of the transgenic embryos following β-galactosidase staining. Transverse and sagittal sections of the transgenic embryos at 10.5 (D-G) and 12.5 dpc (H). (A) 9.5, (B,D-G) 10.5, (C) 11.5, (H) 12.5 and (I,J) 13.5 dpc transgenic embryos. (E,F,G) Transverse section through the cephalic region of a 10.5 dpc transgenic embryo, showing *lacZ* expression in the trigeminal ganglions (V) and the ventral portion of the spinal cord (arrowheads). (K) Whole-mount in situ hybridization analysis of mouse *Otx2* expression at 9.5 dpc. *Otx2* expression is also detected in the trigeminal nerves and the first branchial groove (arrowheads). d, diencephalon; m, mesencephalon; p, p1 diencephalon; t, telencephalon; V1, opthalmic branch of the trigeminal nerve; V2, maxillar branch of the trigeminal nerve.

local signaling centers in the rostral brain, as defined by the expression Bmps, Shh and Wnts, respectively (Figs 4-7, 10). The F3 cis-acting region governed expression in the most medial and caudal aspects of the telencephalon, and in the roof of the diencephalon, coincident with the prospective site of choroid plexus development (Figs 5, 10). The expression domains afforded by the F3 cis-region are strikingly co-localized with Bmps, noggin, *Msx1* and *transthyretin* (*Ttr*) (Furuta et al., 1997). Indeed, *Otx2* is expressed in the dorsocaudal telencephalon and in the dorsal diencphalon, including in the prospective choroid plexus and the 'cortical hem' (see below) in mouse embryos (Simeone et al., 1993; Boncinelli et al., 1993; Stoykova et al., 1996). In order to establish whether the F3 cis-region is controlled directly by Bmp signaling, experiments were performed involving neural plate explants with BMP beads as described (Furuta et al., 1997). The neuroectoderm was isolated at the level of the forebrain from *F3placZ* transgenic embryos at 10.5 dpc. Recombinant BMP2-coated beads were transplanted into these forebrain explants, and the explants were then cultured for 24 hours. However, no ectopic *lacZ* expression induced by the BMP2-coated beads was detected (C.K.-Y. and I.M., unpublished). These data suggest that the F3 cis-regulator may not be regulated by BMP signaling. Alternatively, the *Otx2* expression from this F3 cis-regulator might participate in the expression of Bmp molecules in the prospective choroid plexus. By contrast, the F8 cis-acting region directed transgene expression in the mediocaudal telencephalon where Wnt molecules are co-expressed, defining a zone termed 'cortical hem' (Figs 6, 10) (Parr et al., 1993; Grove et al., 1998). Therefore, Wnt signals might control the *Otx2* expression mediated by the F8 cis-region in the cortical hem and dorsal telencephalon. Additional transgenic zebrafish studies have indicated that F3 and F8 cis-regions directed considerably conserved expression in the forebrain (Fig. 11). Concomitant with this finding, zebrafish *Wnt8b* is expressed in the dorsal forebrain (Kelly et al., 1995). These data suggest that local signaling centers play an essential role in forebrain development in zebrafish embryos and are evolutionarily conserved among vertebrates.

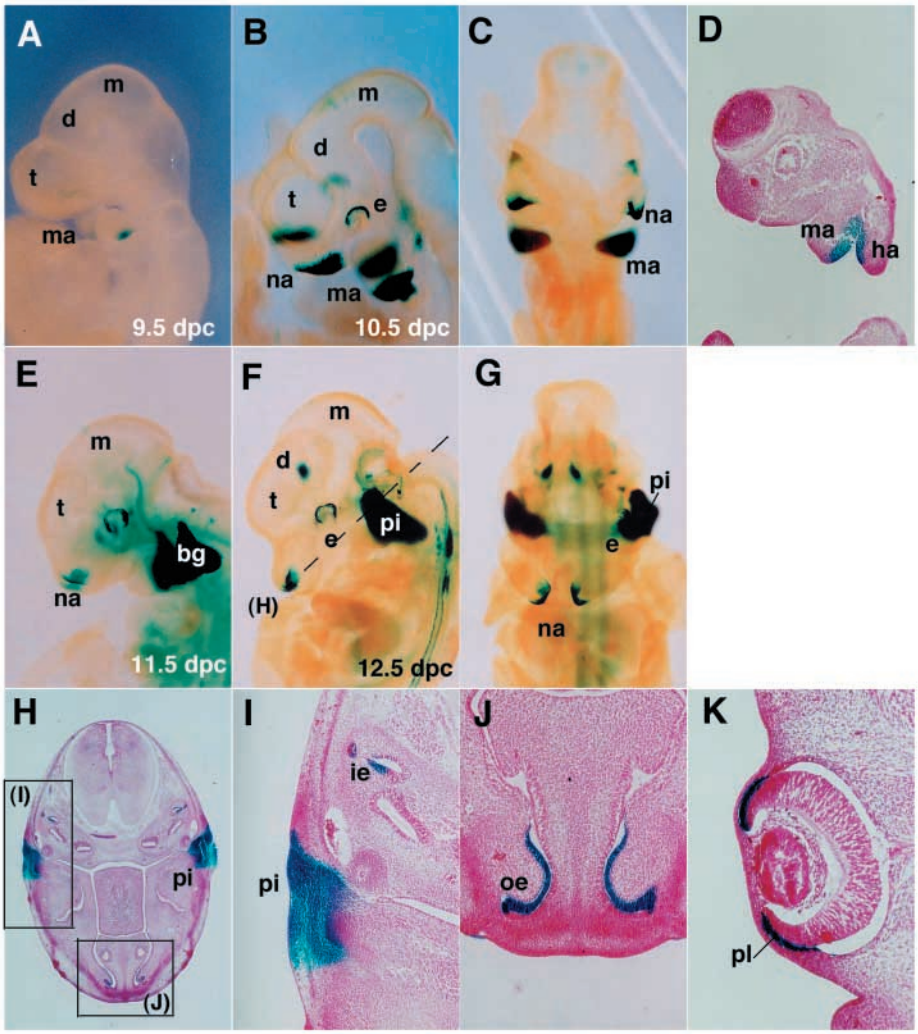


Fig. 9. Developmental changes in *lacZ* expression with the *F12placZ* construct. Lateral (A,B,E,F) and frontal (C,G) views of transgenic embryos following β-galactosidase staining. Transverse and sagittal sections of the transgenic embryos at 11.5 (D) and 12.5 (H-K) dpc. (A) *lacZ* expression is detected initially in the cranial portion of the first branchial groove and nasal regions (arrowhead) at 9.5 dpc. bg, branchial groove; d, diencephalon; e, eye; m, mesencephalon; ma, mandibular arch; ha, hyoid arch; ie, inner ear; na, nasal pits; oe, olfactory epithelium; pi, pinna; pl, pigment layer of retina; t, telencephalon.

The F5 cis-acting region directs expression in the ZLI, and in ventral portions of the diencephalon and the lateral mesencephalon that may be related to longitudinal columns termed 'midbrain arcs' (Figs 7, 10) (Agarwala et al., 2001; Sanders et al., 2002). Notably, mouse *Otx2* mRNA and Otx2 protein are expressed in the ZLI, and in the ventral diencephalon and mesencephalon in mouse embryos (Simeone et al., 1993; Mallamaci et al., 1996; Stoykova et al., 1996). Importantly, expression in the ZLI, ventral diencephalon and midbrain arcs are proposed to be closely related to Shh signaling, which is essential for the dorsoventral patterning of the ventral neural tube. Therefore, these expression patterns suggest the possibility that Shh signaling controls the *Otx2* expression mediated by the F5 cis-acting region.

Fotx2 cis-regulators direct expression in the developing organs

lacZ activity directed by the F9, F11 and F12 cis-regions appears to be closely related to the developing organs. Notably, it has been suggested that *Otx2* is required for the development and specification of these specific organs. The F9 cis-region governs expression in the cephalic mesenchyme (Kimura et al.,

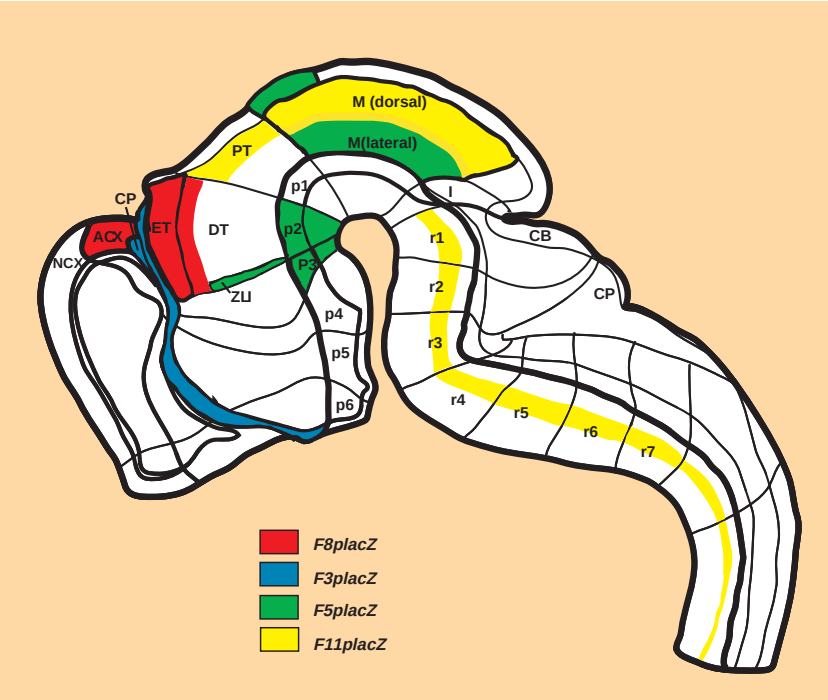


Fig. 10. Schematic representation of *Otx2-lacZ* reporter expression in the mouse brain, based on the neuromeric model of Puelles and Rubenstein (Puelles and Rubenstein, 1993) and Bulfone et al. (Bulfone et al., 1993). The expression patterns of the transgenes depicted in this figure are reconstructed from the analysis of whole-mount and serial sections of β-galactosidase-stained transgenic embryos, and are represented in the following colors: *F8placZ*, red; *F3placZ*, blue; *F5placZ*, green; *F11placZ*, yellow. ACH, archicortex; CB, cerebellum; CP, choroid plexus; DT, dorsal thalamus; ET, epithalamus I, isthmus; M, mesencephalon; NCX, neocortex; PT, pretectum; r1-7, rhombomere 1-7; ZLI, zona limitans intrathalamica.

1997). Consistently, *Otx2* is expressed in the cephalic mesenchyme, and functions in the formation of premandibular and mandibular skulls (Matsuo et al., 1995; Kimura et al., 1997; Hide et al., 2002). The F11 region governs expression in the dorsal pretectum and mesencephalon, including in the posterior commissure, the mesencephalic trigeminal neurons, the oculomotor nerve, the first and second branches of the trigeminal nerve, and the trigeminal ganglions (Figs 8, 10). In actuality, *Otx2* mRNA is expressed in the dorsal pretectum and mesencephalon, in the trigeminal ganglions and opthalmic branch (Simeone et al., 1993; Ang et al., 1994) (Fig. 8M). Furthermore, as *Otx2* heterozygous-mutant mice display anomalies of these neurons and cranial nerves, *Otx2* may be required for the proper development of these tissues cell-autonomously (Matsuo et al., 1995). The F12 cis-region drives expression in the developing sense organs, the olfactory epithelium, the pigment layer of the retina, and the developing inner and outer ears. *Otx2* expression is consistently found in these developing sense organs (Simeone et al., 1993; Mallamaci et al., 1996; Morsli et al., 1999) (Fig. 8K). Moreover, *Otx2* is required for normal development of the inner ear and pigment epithelium, cooperatively with the *Otx1* gene (Morsli et al., 1999; Cantos et al., 2000; Martinez-Morales et al., 2001).

Genetic control of *Otx2* expression during head development

We have surveyed the entire pufferfish *Otx2* locus; furthermore, we observed the regulation of *Otx2* expression by multiple, yet

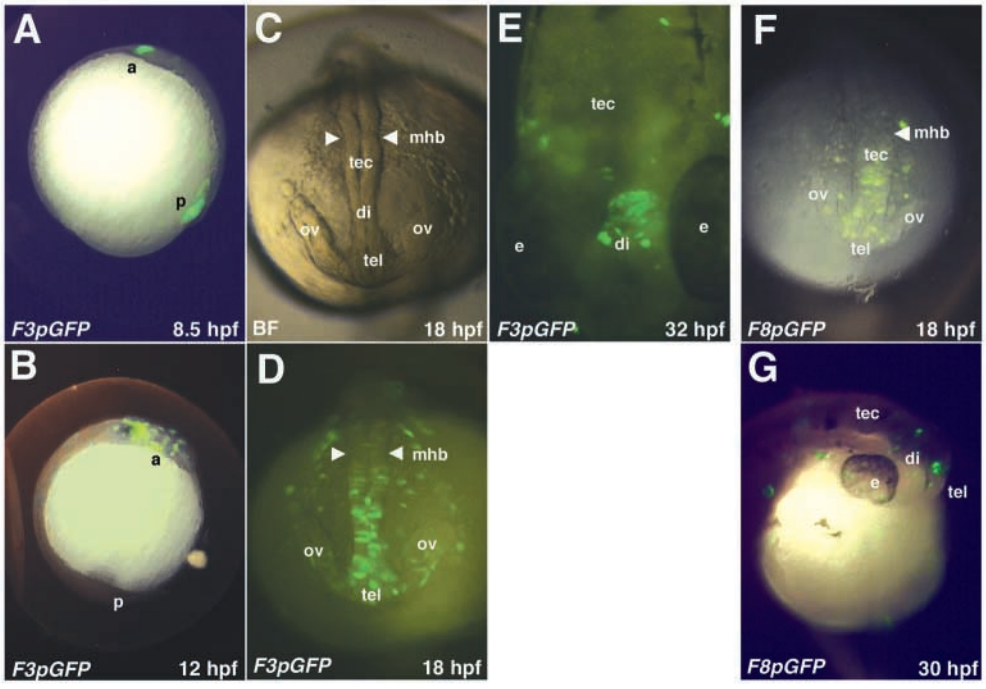


Fig. 11. Transient expression analysis of F3 and F8 cis-regions fused with the GFP reporter gene in transgenic zebrafish. Lateral and dorsal views of transgenic fish embryos transiently harboring the *F3pGFP* (A-E) and *F8pGFP* (F,G) constructs. GFP activity with *F3pGFP* at 8.5 (A), 12 (B) and 32 (E) hpf. Bright (C) and dark (D) field observations of *F3pGFP* transgenic fish at 18 hpf. *F8pGFP* expression at 18 (F) and (G) 30 hpf. a, anterior; di, diencephalon; e, eye; mhb, mid-hindbrain boundary; ov, optic vesicle; tec, tectum; tel, telencephalon; p, posterior.

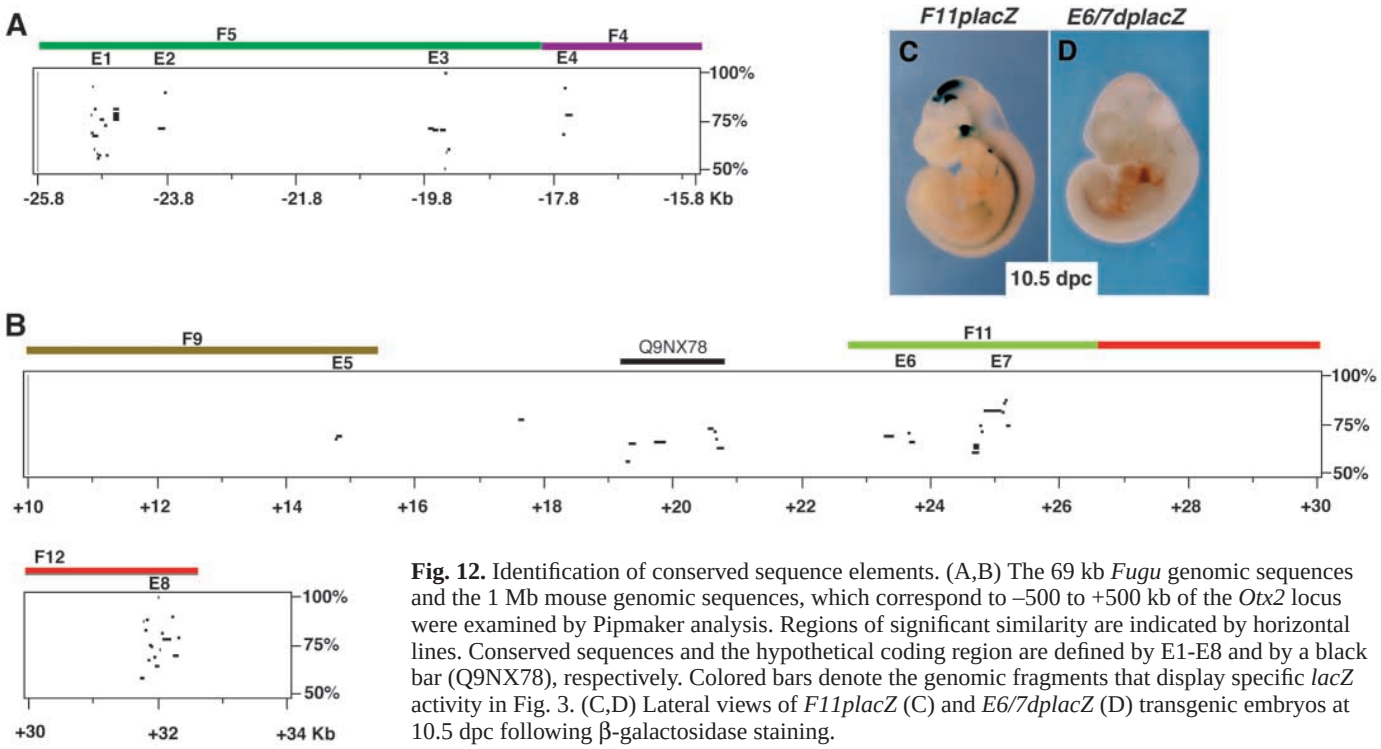


Fig. 12. Identification of conserved sequence elements. (A,B) The 69 kb *Fugu* genomic sequences and the 1 Mb mouse genomic sequences, which correspond to –500 to +500 kb of the *Otx2* locus were examined by Pipmaker analysis. Regions of significant similarity are indicated by horizontal lines. Conserved sequences and the hypothetical coding region are defined by E1-E8 and by a black bar (Q9NX78), respectively. Colored bars denote the genomic fragments that display specific *lacZ* activity in Fig. 3. (C,D) Lateral views of *F11placZ* (C) and *E6/7dplacZ* (D) transgenic embryos at 10.5 dpc following β-galactosidase staining.

distinct, cis-acting regions. The regulatory mechanism of this gene is much more complex than previously anticipated. It is certain that these regions transactivate *Otx2* expression independently. One possible assumption regarding the presence of such long-distance and multiple cis-acting regions entails the ability of each cis-region to transactivate *Otx2* expression upon adjacent cis-elements without mutual interference. Current reports have shown that Hox genes are clustered into complexes; moreover, their positions within the cluster correlate with their time of expression and with the position of the AP boundaries of their expression domains (Duboule and Morata, 1994). Similarly, *β-globin* gene clusters are controlled in order of their expression during ontogeny (Grosveld, 1999). In the present study, *Fotx2* cis-regulators are clustered into the 70 kb genomic sequence of the *Fotx2* locus; however, their positions relative to the coding region, and the order of the 5′ to the 3′ side, do not correlate with their time of expression, or with the expression domains along the anteroposterior or dorsoventral axis. This observation suggests that these cis-regulators may direct multiple promoters located within this genomic locus. Consistent with this notion, a novel gene, annoted as NM172600, which is located over +150 kb from the mouse *Otx2* coding region, is a homolog of *Fugu* Q9NX78 (Fig. 12B), and is expressed in the rostral brain, in which *Otx2* expression occurs during embryogenesis (C.K.-Y. and I.M., unpublished). This finding supports our assumption that genes located over 500 kb from this locus may be coincidentally controlled by these distant cis-regulators. Curiously, the corresponding conserved sequence elements in mouse are also located in an order similar to that in pufferfish; however, these elements are within an approximately eightfold extended region, up to 550 kb (Fig. 12). Thus, we propose that, during evolution in the common ancestor of teleosts and mammals, the *Otx2* promoter might have acquired multiple region-specific cis-regulators in order to control the highly coordinated processes of rostral head development. Consequently, the *Otx2* homeobox gene could play multiple and pivotal roles in rostal head development by means of these cis-elements. Investigation of the mechanism of non-vertebrate *Otx2* gene expression is of keen interest (Oda-Ishii and Saiga, 2003).

We are grateful to the Animal Resources and Genetic Engineering Team at the RIKEN Center for Developmental Biology for housing the mice. We thank Drs K. Shimamura and C. Ragsdale for helpful comments on the manuscript; Drs A. McMahon and R. Harland for in situ probes; and H. Nakano for his assistance. C.K.-Y. is supported by a special postdoctoral researchers' program. This work was supported in part by grants-in-aid for Scientific Research on Priority Areas from the Ministry of Education, Culture, Sports Science and Technology, Japan.

References

Acampora, D., Mazan, S., Lallemand, Y., Avantaggiato, V., Maury, M., Simeone, A. and Brûlet, P. (1995). Forebrain and midbrain regions are deleted in *Otx2*$^{-/-}$ mutants due to a defective anterior neuroectoderm specification during gastrulation. *Development* **121**, 3279-3290.

Acampora, D., Avantaggiato, V., Tuorto, F. and Simeone, A. (1997). Genetic control of brain morphogenesis through *Otx* gene dosage requirement. *Development* **124**, 3639-3650.

Agarwala, S., Sanders, T. A. and Ragsdale, C. W. (2001). Sonic hedgehog control of size and shape in midbrain pattern formation. *Science* **291**, 2147-2150.

Ang, S.-L., Conlon, R. A., Jin, O. and Rossant, J. (1994). Positive and negative signals from mesoderm regulate the expression of mouse *Otx2* in ectoderm explants. *Development* **120**, 2979-2989.

Ang, S.-L., Jin, O., Rhinn, M., Daigle, N., Stevenson, L. and Rossant, J. (1996). A targeted mouse *Otx2* mutation leads to severe defects in gastrulation and formation of axial mesoderm and to deletion of rostral brain. *Development* **122**, 243-252.

Aparicio, S., Morrison, A., Gould, A., Gilthorpe, J., Chaudhuri, C., Rigby, P., Krumlauf, R. and Brenner, S. (1995). Detecting conserved regulatory

elements with the model genome of the Japanese puffer fish, *Fugu rubripes*. *Proc. Natl. Acad. Sci. USA* **92**, 1684-1688.

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

Boncinelli, E., Gulisano, M. and Broccoli, V. (1993). *Emx* and *Otx* homeobox genes in the developing mouse brain. *J. Neurobiol*. **24**, 1356-1366.

Brenner, S., Elgar, G., Sandford, R., Macrae, A., Venkatesh, B. and Aparicio, S. (1993). Characterization of the puffer fish (*Fugu*) genome as a compact model vertebrate genome. *Nature* **366**, 265-268.

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

Bulfone, A., Puelles, L., Porteus, M. H., Frohman, M. A., Martin, G. R. and Rubenstein, J. L. (1993). Spatially restricted expression of *Dlx-1*, *Dlx-2* (*Tes-1*), *Gbx-2*, and *Wnt-3* in the embryonic day 12.5 mouse forebrain defines potential transverse and longitudinal segmental boundaries. *J. Neurosci.* **13**, 3155-3172.

Camacho-Hubner, A., Rossier, A., and Beermann, F. (2000). The *Fugu rubripes* tyrosinase gene promoter targets transgene expression to pigment cells in the mouse. *Genesis* **28**, 99-105.

Cantos, R., Cole, L. K., Acampora, D., Simeone, A. and Wu, D. K. (2000). Patterning the mammalian cochlea. *Proc. Natl. Acad. Sci. USA.* **97**, 11707-11713.

Duboule, D. (1998). Vertebrate *hox* gene regulation: clustering and/or colinearity? *Curr. Opin. Genet. Dev.* **8**, 514-518.

Duboule, D. and Morata, G. (1994). Colinearity and functional hierarchy among genes of the homeotic complexes. *Trends Genet.* **10**, 358-364.

Eng, S. R., Gratwick, K., Rhee, J. M., Fedtsova, N., Gan, L. and Turner, E. E. (2001). Defects in sensory axon growth precede neuronal death in Brn3a-deficient mice. *J. Neurosci.* **21**, 541-549.

Finkelstein, R. and Perrimon, N. (1991). The molecular genetics of head development in *Drosophila melanogaster*. *Development* **112**, 899-912.

Furukawa, T., Morrow. E. M. and Cepko, C. L. (1997). Crx, a novel otx-like homeobox gene, shows photoreceptor-specific expression and regulates photoreceptor differentiation. *Cell* **91**, 531-541.

Furuta, Y., Piston, D. W. and Hogan, B. L. (1997). Bone morphogenetic proteins (BMPs) as regulators of dorsal forebrain development. *Development* **124,** 2203-2012.

Grosveld, F. (1999). Activation by locus control regions? *Curr. Opin. Genet. Dev.* **9**, 152-157.

Grove, E. A., Tole, S., Limon, J., Yip, L. and Ragsdale, C. W. (1998). The hem of the embryonic cerebral cortex is defined by the expression of multiple *Wnt* genes and is compromised in *Gli3*-deficient mice. *Development* **125**, 2315-2325.

Hardison, R. C. (2000). Conserved noncoding sequences are reliable guides to regulatory elements. *Trends Genet.* **16**, 369-372.

Hebert, J., Mishina, Y. and McConnell, S. K. (2002). BMP signaling is required locally to pattern the dorsal telencephalon midline. *Neuron* **35**, 1029-1041.

Hide, T., Hatakeyama, J., Kimura-Yoshida, C., Tian, E., Takeda, N., Ushio, Y., Shiroishi, T., Aizawa, S. and Matsuo, I. (2002). Genetic modifiers of otocephalic phenotypes in *Otx2* heterozygous mutant mice. *Development* **129**, 4347-4357.

Hogan, B., Beddington, R., Costantini, F. and Lacy, E. (1994). *Manipulating the Mouse Embryo: A Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Jowett, T. and Lettice, L. (1994). Whole-mount in situ hybridisation on zebrafish embryos using a mixture of digoxigenin and fluorescein-labeled probes. *Trends Genet.* **10**, 73-74.

Kammandel, B., Chowdhury, K., Stoykova, A., Aparicio, S., Brenner, S. and Gruss, P. (1999). Distinct *cis*-essential modules direct the time-space pattern of the *Pax6* gene activity. *Dev. Biol.* **205**, 79-97.

Kelly, G. M., Greenstein, P., Erezyilmaz, D. F. and Moon, R. T. (1995). Zebrafish *Wnt8* and *Wmt8b* share a common activity but are involved in distinct developmental pathways. *Development* **121**, 1787-1799.

Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B. and Schilling, T. F. (1995). Stages of embryonic development of the zebrafish. *Dev. Dyn.* **203**, 253-310.

Kimura, C., Takeda, N., Suzuki, M., Oshimura, M., Aizawa, S. and Matsuo, I. (1997). *Cis*-acting elements conserved between mouse and pufferfish *Otx2* genes govern the expression in mesencephalic neural crest cells. *Development* **124**, 3929-3941.

Kimura, C., Yoshinaga, K., Tian, E., Suzuki, M., Aizawa, S. and Matsuo, I. (2000). Visceral endoderm mediates forebrain development by suppressing posteriorizing signals. *Dev. Biol.* **225**, 304-321.

Kimura, C., Shen, M. M., Takeda, N., Aizawa, S. and Matsuo, I. (2001). Complementary functions of *Otx2* and *Cripto* in initial patterning of mouse epiblast. *Dev. Biol.* **235**, 12-32.

Kobayashi, M., Nishikawa, K. and Yamamoto, M. (2001). Hematopoietic regulatory domain of *gata1* gene is positively regulated by GATA1 protein in zebrafish embryos. *Development* **128**, 2341-2350.

Kothary, R., Clapoff, S., Darling, S., Perry, M. D., Moran, L, A. and Rossant, J. (1989). Inducible expression of an *hsp68-lacZ* hybrid gene in transgenic mice. *Development* **105**, 707-714.

Krumlauf, R. (1994). *Hox* genes in vertebrate development. *Cell* **78**, 191-201.

Lee, S. M., Tole, S., Grove, E. and McMahon, A. P. (2000). A local Wnt-3a signal is required for development of the mammalian hippocampus. *Development* **127**, 457-467.

Li, Y., Allende, M. L., Finkelstein, R. and Weinberg, E. S. (1994). Expression of two zebrafish *orthodenticle*-related genes in the embryonic brain. *Mech. Dev.* **48**, 229-244.

Logan, C., Khoo, W. K., Cado, D. and Joyner, A. L. (1993). Two enhancer regions in the mouse *En-2* locus direct expression to the mid/hindbrain region and mandibular myoblasts. *Development* **117**, 905-916.

Mallamaci, A., Di Blas, E., Briata, P., Boncinelli, E. and Corte, G. (1996). OTX2 homeoprotein in the developing central nervous system and migratory cells of the olfactory area. *Mech. Dev.* **58**, 165-178.

Marshall, H., Studer, M., Aparicio, S., Kuroiwa, A., Brenner, S. and Krumlauf, R. (1994). A conserved retinoic acid response element required for early expression of the homeobox gene *Hoxb-1*. *Nature* **370**, 567-571.

Martinez-Morales, J. R., Signore, M., Acampora, D., Simeone, A. and Bovolenta, P. (2001). Otx genes are required for tissue specification in the developing eye. *Development* **128**, 2019-2030.

Matsuo, I., Kuratani, S., Kimura, C., Takeda, N. and Aizawa, S. (1995). Mouse *Otx2* functions in the formation and patterning of rostral head. *Genes Dev.* **9**, 2646-2658.

Morsli, H., Tuorto, F., Choo, D., Postiglione, M. P., Simeone, A. and Wu, D. K. (1999). *Otx1* and *Otx2* activities are required for the normal development of the mouse inner ear. *Development* **126**, 2335-2343.

Nonchev, S., Maconochie, M., Gould, A., Morrison, A. and Krumlauf, R. (1997). Cross-regulatory interactions between *Hox* genes and the control of segmental expression in the vertebrate central nervous system. *Cold Spring Harb. Symp. Quant. Biol.* **62**, 313-323.

Oda-Ishii, I. and Saiga, H. (2003). Genomic organization and promoter and transcription regulatory regions for the expression in the anterior brain (sensory vesicle) of *Hroth*, the *otx* homologue of the ascidian, *Halocynthia roretzi*. *Dev. Dyn.* **227**, 104-113.

Parr, B. A., Shea, M. J., Vassileva, G. and McMahon, A. P. (1993). Mouse *Wnt* genes exhibit discrete domains of expression in the early embryonic CNS and limb buds. *Development* **119**, 247-261.

Pöpperl, H., Bienz, M., Studer, M., Chan, S.-W., Aparicio, S., Brenner, S., Mann, R. S. and Krumlauf, R. (1995). Segmental expression of *Hoxb-1* is controlled by a highly conserved autoregulatory loop dependent upon *exd/pbx*. *Cell* **81**, 1031-1042.

Puelles, L. and Rubenstein, J. L. R. (1993). Expression patterns of homeobox and other putative regulatory genes in the embryonic mouse forebrain suggest a neuromeric organization. *Trends Neurosci.* **16**, 472-479.

Rubenstein, J. L. R., Shimamura, K., Martinez, S. and Puelles, L. (1998). Regionalization of the prosencephalic neural plate. *Annu. Rev. Neurosci*. **21**, 445-477.

Sambrook, J. and Russell, D. W. (2001). *Molecular Cloning: A Laboratory Manual*, 3rd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Sanders, T. A., Lumsden, A. and Ragsdale, C. W. (2002). Arcuate plan of chick midbrain development. *J. Neurosci.* **22**, 10742-10750.

Schwartz, S., Zhang, Z., Frazer, K. A., Smit, A., Riemer, C., Bouck, J., Gibbs, R., Hardison, R. and Miller, W. (2000). PipMaker – a web server for aligning two genomic DNA sequences. *Genome Res.* **10**, 577-586.

Simeone, A., Acampora, D., Gulisano, M., Stornaiuolo, A. and Boncinelli, E. (1992). Nested expression domains of four homeobox genes in developing rostral brain. *Nature* **358**, 687-690.

Simeone, A., Acampora, D., Mallamaci, A., Stornaiuolo, A., D'Apice, M. R., Nigro, V. and Boncinelli, E. (1993). A vertebrate gene related to *orthodenticle* contains a homeodomain of the *bicoid* class and demarcates

anterior neuroectoderm in the gastrulating mouse embryo. *EMBO J.* **12**, 2735-2747.

Stoykova, A., Fritsch, R., Walther, C. and Gruss, P. (1996). Forebrain patterning defects in *Small eye* mutant mice. *Development* **122**, 3453-3465.

Suda, Y., Matsuo, I., Kuratani, S. and Aizawa, S. (1996). *Otx1* function overlaps with *Otx2* in development of mouse forebrain and midbrain. *Genes Cells* **1**, 1031-1044.

Suda, Y., Matsuo, I. and Aizawa, S. (1997). Cooperation between *Otx1* and *Otx2* genes in developmental patterning of rostral brain. *Mech. Dev.* **69**, 125-141.

Suda, Y., Hossain, Z. M., Kobayashi, C., Hatano, O., Yoshida, M., Matsuo, I. and Aizawa, S. (2001). *Emx2* directs the development of diencephalon in cooperation with *Otx2*. *Development* **128**, 2433-2450.

Suzuki, T., Kurokawa, T., Hashimoto, H. and Sugiyama, M. (2002). CDNA sequence and tissue expression of *Fugu rubripes* prion protein-like: a candidate for the teleost orthologue of tetrapod PrPs. *Biochem. Biophys. Res. Commun.* **284**, 912-917.

Tian, E., Kimura, C., Takeda, N., Aizawa, S. and Matsuo, I. (2002). *Otx2* is required to respond to signals from anterior neural ridge for forebrain specification. *Dev. Biol.* **242**, 204-223.

Westerfield, M. (1995). *The Zebrafish Book*, 3rd edn. Eugene, OR: University of Oregon Press.

Zerucha, T., Stuhmer, T., Hatch, G., Park, B. K., Long, Q., Yu, G., Gambarotta, A., Schultz, J. R., Rubenstein, J. L. and Ekker, M. (2000). A highly conserved enhancer in the *Dlx5*/*Dlx6* intergenic region is the site of cross-regulatory interactions between *Dlx* genes in the embryonic forebrain. *J. Neurosci.* **20**, 709-721.