Developmental Biology **229,** 31–43 (2001)
doi:10.1006/dbio.2000.9964, available online at http://www.idealibrary.com on IDEAL®

The Regulation of *forkhead/HNF-3β* Expression in the *Ciona* Embryo

Anna Di Gregorio,[1] Joseph C. Corbo,[2] and Michael Levine

Division of Genetics, Department of Molecular and Cellular Biology, University of California, 401 Barker Hall, Berkeley, California 94720

The *Ciona forkhead/HNF-3β* gene (*Ci-fkh*) is expressed in the primary axial tissues of the developing tadpole, including the notochord, endoderm, and rudimentary floor plate of the CNS. In an effort to determine the basis for this complex pattern of expression we have conducted a detailed analysis of the *Ci-fkh* 5′-regulatory region. Different 5′ sequences were attached to a *lacZ* reporter gene and analyzed in electroporated *Ciona* embryos. A short regulatory sequence (AS) located ~1.7 kb upstream of the transcribed region is shown to be essential for expression in all three axial tissues. The proximal 20 bp of the AS contains overlapping Snail repressor elements and a T-box motif. Deleting these sequences causes the loss of reporter gene expression in the endoderm, as well as expanded expression in the neural tube. These results suggest that a T-box gene such as *Ci-VegTR* activates *Ci-fkh* expression in the endoderm, while the Ci-Sna repressor excludes expression from the lateral ependymal cells and restricts the *Ci-fkh* pattern to the rudimentary floor plate in ventral regions of the neural tube. We also present evidence for *Ci-fkh* positive autofeedback, whereby the Ci-Fkh protein binds to critical activator sites within the *Ci-fkh* 5′-regulatory region and helps maintain high levels of expression. We discuss these results with respect to *forkhead/HNF-3β* regulation in vertebrates.

***Key Words: Ciona intestinalis;* ascidian; *forkhead; HNF-3β*; notochord; endoderm; spinal cord.**

INTRODUCTION

The *Ciona* tadpole represents one of the most primitive and simplified forms of the chordate body plan. We are interested in using the *Ciona* embryo as a means to investigate the evolutionary origins of axial tissues, particularly the floor plate, notochord, and hypochord. Previous studies have shown that a *Ciona* homolog of the vertebrate *HNF-3β* (*forkhead*) gene, *Ci-fkh*, is specifically expressed in the CNS, notochord, and endoderm (Corbo *et al.*, 1997b). Considered in cross section, the spinal cord of the tadpole CNS is composed of just four ependymal cells. *Ci-fkh* is specifically expressed in the ventral-most ependymal cell, which is in direct contact with the notochord. This observation prompted the proposal that a rudimentary floor plate was already present in the last shared ancestor of ascidians and vertebrates (Corbo *et al.*, 1997b; Olsen and Jeffery, 1997; Shimauchi *et al.*, 1997). *Ci-fkh* is also expressed in the notochord and endoderm, including both the gut primordium and the endodermal strand. The latter tissue is a closed tube that runs along the length of the tail, just beneath the notochord. It has been suggested that the endodermal strand represents a primitive version of the hypochord (Corbo *et al.*, 1997b), a transient endodermal structure found in lampreys, amphibians, and fish (e.g., Lofberg and Collazo, 1997). Thus, it would appear that *forkhead* expression patterns are homologous in *Ciona* and vertebrates.

HNF-3β is expressed in the node, notochord, floor plate, and gut of vertebrate embryos, including zebrafish, chicks, and mice (e.g., Strahle *et al.*, 1993; Ruiz i Altaba *et al.*, 1993). Mice lacking *HNF-3β* gene activity die *in utero*, lack an organized node and notochord, and also exhibit severe dorsoventral patterning defects in the neural tube, as well as disruptions in the foregut (Ang and Rossant, 1994; Weinstein *et al.*, 1994). Previous studies identified two enhancers that regulate the mouse *HNF-3β* gene; these are located 14 kb upstream and 6 kb downstream of the transcription start site (Sasaki and Hogan, 1996). The 5′ enhancer directs expression in the node and notochord, while the 3′ enhancer mediates expression in

[1] To whom correspondence should be addressed. Fax: (510) 643-5785. E-mail: annadg@uclink4.berkeley.edu.

[2] Current address: Department of Pathology, Brigham and Women's Hospital, Boston, MA 02115.

0012-1606/01 $35.00

the floor plate, posterior notochord, and dorsal hindgut. There is currently little information regarding the identities of the *trans*-acting factors that regulate these enhancers, although it has been proposed that HNF-3β might mediate autoregulation while one or more unknown repressors might restrict the activities of the 3′ enhancer to the floor plate (Sasaki and Hogan, 1994). A potential Gli binding site that is located within the 3′ enhancer was suggested to play a role in activating *HNF-3β* expression in the floor plate in response to Sonic Hedgehog signaling from the notochord (Sasaki *et al.*, 1997).

Ciona provides a simple system for analyzing *forkhead/HNF-3β* regulation since the genome is small and compact (just 5% the size of most vertebrate genomes), contains well-defined embryonic lineages, and permits the introduction of transgenic DNA via electroporation (reviewed by Di Gregorio and Levine, 1998). To identify the different *cis*-regulatory elements responsible for the complex *Ci-fkh* pattern we have analyzed the expression of *Ci-fkh/lacZ* transgenes in electroporated embryos. Evidence is presented that the *Ci-fkh* promoter region contains partially separable, but tightly linked, *cis*-regulatory elements controlling gene expression in all three axial tissues: the neural tube, notochord, and endoderm. These elements are clustered within a common distal regulatory sequence (the axial sequence, or AS), located ~1.7 kb upstream of the *Ci-fkh* transcription start site. This organization is quite distinct from that observed for the mouse *HNF-3β* gene.

The analysis of various truncated and deleted *Ci-fkh* promoter sequences provides evidence that transcriptional repression is an essential component of *Ci-fkh* regulation. Removal of repressor sites from the AS results in the persistence of the early *Ci-fkh* expression pattern in the lateral ependymal cells of the spinal cord. Moreover, a series of internal deletions allowed the uncoupling of a region responsible for notochord/endoderm expression from the *cis*-regulatory elements controlling expression in the CNS. Partial uncoupling of the notochord and endoderm expression patterns was obtained by mutating a T-box recognition sequence in the distal enhancer. This mutation specifically disrupts expression in the endoderm, whereas staining in the notochord is essentially unaffected. DNA binding assays show that a GST fusion protein containing the T-domain of the *Ciona* Brachyury (Ci-Bra) protein binds this T-box element, although we suggest that a different member of the T-box family is probably responsible for regulating *Ci-fkh* expression in the endoderm. Finally, DNA binding assays reveal a number of high-affinity Ci-Fkh recognition sequences in the 5′-regulatory region. Mutations in one of these sites, adjacent to the AS, diminish the expression mediated by an otherwise normal *Ci-fkh/lacZ* transgene, thereby providing evidence for autoregulation.

MATERIALS AND METHODS

Ascidians

Adults were collected from several locations in Northern California, purchased from Woods Hole, Massachusetts, and kindly provided by Professor Nori Satoh (Kyoto University, Japan). The animals were kept at 18°C in recirculating natural seawater. After *in vitro* fertilization and dechorionation with Pronase E (Sigma), embryos were kept at 15°C for either electroporations or *in situ* hybridization experiments.

Whole-Mount in Situ Hybridizations

Embryos were fixed in paraformaldehyde as described by Corbo *et al.* (1997a) and stored at −20°C. Digoxigenin-labeled RNA probes were synthesized from the *Ci-fkh* RACE product after digestion with appropriate restriction enzymes. Either 5′ (antisense) or 3′ (sense) RNAs were synthesized using T7 or SP6 RNA polymerase (Promega). *In situ* hybridizations on whole-mount staged embryos were done as described by Corbo *et al.* (1997a).

fkh-lacZ Transgenes and Electroporations

A 350-bp PCR product containing the winged-helix domain of *Xenopus Pintallavis* was used to screen a cDNA library from gastrula-stage *Ciona* embryos (kindly provided by Drs. Jamie Lee and Tom Meedel). The largest cDNA that was isolated contains ~1950 bp, and both strands were sequenced using fragments generated by exonuclease III digestion (Sambrook *et al.*, 1986). A previously identified 5′-RACE product contains nt 1–437 of the ORF and 76 nt of the 5′-UTR (Corbo *et al.*, 1997b). The remaining sequences of the *Ci-fkh* ORF (nt 438–1838) were derived from the original cDNA.

The 350-bp *Xenopus* PCR fragment was also used to screen a *Ciona* genomic DNA library. One of the clones that was isolated contains ~5.5 kb of 5′-flanking sequence, as well as the entire 5′ UTR and the first 86 codons of the ORF. This genomic DNA fragment was used to prepare all of the *lacZ* transgenes shown in Table 1. It was cloned into the p72-1.27 vector (Corbo *et al.*, 1997a) as an *Spe*I–*Not*I fragment. The *lacZ* reporter gene was fused in-frame with the *Ci-fkh* coding sequence (−5.6 kb transgene shown in Table 1). An internal *Eco*RI site allowed the creation of the −2.6 kb *Ci-fkh/lacZ* transgene. The −1.77 kb *Ci-fkh/lacZ* transgene was made by inserting the −2.6 kb *Ci-fkh* genomic DNA fragment into a pBlueScript II SK(+) vector, digesting with exonuclease III, and then excising the remaining fragment with *Kpn*I and *Not*I sites located in the polylinker. The −1.77 kb fragment was subsequently inserted into the *Kpn*I and *Not*I sites within the p72-1.27 vector. The −1.69, −1.67, −1.66, −1.65, and −1.63 kb transgenes, as well as the transgenes containing point mutations, were generated by PCR using *Pfu* DNA polymerase (Stratagene). The PCR-amplified regions were checked by sequencing both strands using the dideoxy chain-termination method (Sambrook *et al.*, 1989). The −1.49, −1.37, −1.1, −0.9, −0.54, −0.37, and −0.3 kb *Ci-fkh/lacZ* transgenes were generated using the restriction sites shown in Table 1.

All the internal deletions shown in Table 1 were generated by digesting the −1.77 kb *Ci-fkh/lacZ* transgene with restriction enzymes that recognize unique internal sites.

Different fragments from the *Ci-fkh* 5′-flanking region were attached to two different heterologous promoters, *Ci-Bra* and

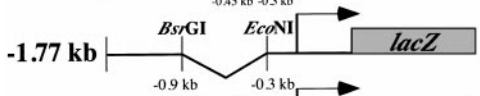
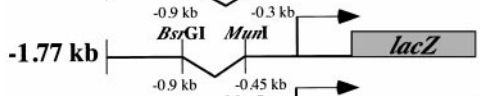
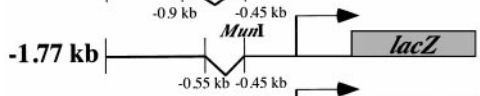
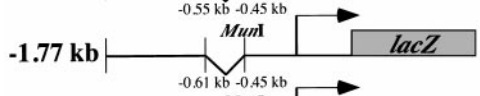
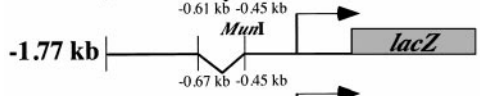
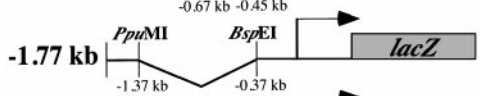
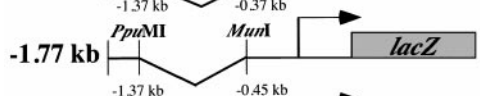
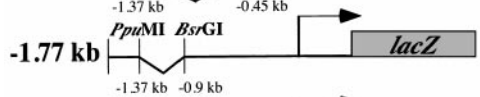
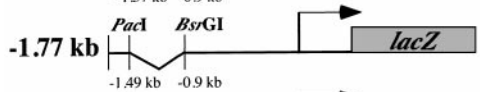
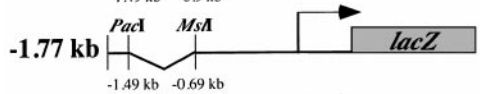
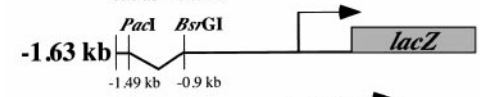
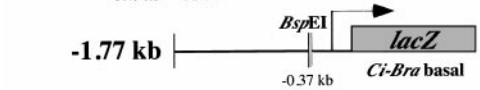
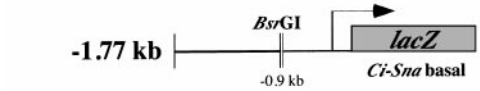
TABLE 1
Summary of *Ci-fkh/lacZ* Transgenes

	Notochord	Endoderm	CNS (V.E.C.)	Ectopic CNS (L.E.C.)	Trunk Epi.	Trunk Mesench.
*Spe*I -5.6 kb *lacZ*	++++	++++	+	+/-	++	+
*Eco*RI -2.6 kb *lacZ*	++++	++++	+	+/-	++	+
-1.77 kb *lacZ*	++++	++++	+/-	+/-	+++	++
-1.69 kb *lacZ*	++++	++++	+/-	+/-	+++	++
-1.67 kb *lacZ*	+++	++++	+/-	+/-	+++	++
-1.66 kb *lacZ*	++	++	+/-	+/-	++	+
-1.65 kb *lacZ*	++	++	-	+	++	+
-1.64 kb *lacZ*	++	+/-	-	++	++	+
-1.63 kb *lacZ*	+	-	-	++	++	+
*Pac*I -1.49 kb *lacZ*	+/-	-	-	++	++	+
*Ppu*MI -1.37 kb *lacZ*	+/-	-	-	+++	+++	+
*Cla*I -1.1 kb *lacZ*	+/-	-	-	++	+++	+/-
-1 kb *lacZ*	+/-	-	-	++	+++	+/-
-0.96 kb *lacZ*	+/-	-	-	+	+/-	+/-
-0.93 kb *lacZ*	-	-	-	+/-	-	+/-
*Bsr*GI -0.9 kb *lacZ*	-	-	-	-	-	+/-
*Mun*I -0.45 kb *lacZ*	-	-	-	-	-	+/-
*Eco*NI -0.3 kb *lacZ*	-	-	-	-	-	-

Ci-sna (some are summarized in Table 2). The minimal *Ci-Bra* promoter extends from -68 bp upstream of the transcription start site to a *Mun*I site corresponding to codon 15 of the ORF (Corbo *et al.*, 1997a). The *Ci-sna* promoter extends from −69 bp upstream of the transcribed sequence and includes the entire 66-bp 5′ UTR, 58 bp of coding sequence, a 1-kb intron, and another 50 bp of coding sequence from the second exon.

Electroporations were done as described by Corbo *et al.* (1997b). Aliquots containing 100 μg of purified plasmid DNA were used in each experiment. Each transgene was tested in at least six indepen-

TABLE 2
Internal Deletions in *Ci-fkh/lacZ* Transgenes

	Notochord	Endoderm	CNS (V.E.C.)	Ectopic CNS (L.E.C.)	Trunk Epi.	Trunk Mesench.
	+++	+++	-	++	+	++
	-	-	-	-	-	-
	+/-	-	-	-	-	-
	+++	++	-	+	+/-	+/-
	+++	++	-	-	+/-	+/-
	-	-	-	-	-	+/-
	-	-	-	-	-	-
	-	-	-	-	-	+/-
	+++	++++	-	-	-	++
	+++	++++	-	-	-	+
	+++	++++	-	-	-	+
	-	-	-	-	-	++
	+++	+	-	-	+/-	++
	-	-	-	-	-	-

dent experiments using different batches of embryos. The number of positive embryos scored in each experiment ranged from approximately 100 to 250.

Staining reactions for β-galactosidase activity were also done as described in Corbo *et al.* (1997a).

Preparation of a GST-Ci-Fkh Fusion Protein

A GST fusion protein was prepared by linking two DNA fragments containing different portions of the *Ci-fkh* coding region. A RACE fragment containing the first 138 codons of the *Ci-fkh* ORF

was obtained as a *Pst*I–*Msl*I fragment and ligated with a *Msl*I–*Eco*RI cDNA fragment containing codons 138–174. The two fragments were cloned into the pBlueScript II SK(+) vector (Stratagene) and then digested with *Bam*HI to obtain a 522-bp fragment that spans nt 344–866 (174 codons) and contains the entire winged-helix DNA binding domain as well as 64 flanking amino acid residues. The *Bam*HI fragment was cloned in frame into the pGEX-KG vector and expressed in *Escherichia coli* strain HB101. The fusion protein was induced and purified as described by Guan and Dixon (1991a). Approximate protein concentrations were determined by staining gels with Coomassie blue and comparing with a series of protein molecular weight standards (RPN 756; Amersham).

Gel Shift Assays

Gel shift assays were done with the GST-Ci-Fkh fusion protein and the following double-stranded oligonucleotides (only the 5′–3′ strand is shown): CIF1, 5′-GAGTAAATATTTGCGGTAGAATCA-3′; CIF1m, 5′-GAGTAAAGAGGTGCGGTAGAATCA-3′; and CIF3, 5′-TTAAGTTTTGTTTGTTCACTGAGA-3′.

Gel shift assays were done with a GST-Ci-Sna fusion protein (Fujiwara *et al.*, 1998), and the following double-stranded oligonucleotides: Sna Cryp, 5′-TTAAGGTGCTACATCCTATGCTTGATTG-3′, and Sna2, 5′-ACGTCACAATACACTTGGTGACGT-3′.

The double-stranded oligonucleotides were radiolabeled using T4 polynucleotide kinase (New England Biolabs) and [γ-^{32}P]ATP, purified on Bio-Gel P-10 (Bio-Rad) spin columns, and then recovered in TE buffer. GST fusion proteins were mixed with $3–5 \times 10^4$ cpm labeled DNA and then incubated at room temperature for 15 min in a final volume of 20 μl (Corbo *et al.*, 1998). Competition experiments were done by including a 5-, 50-, and 100-fold molar excess of the indicated unlabeled oligonucleotides in the binding reactions, after preincubating the mixtures 5 min at room temperature before adding the radiolabeled oligonucleotides. The reaction mixes were fractionated on 5% polyacrylamide/0.5× TBE gels and then analyzed by autoradiography.

RESULTS

Forkhead-Related Genes in Ciona

The winged-helix DNA binding domain of the *Xenopus Pintallavis* gene was used to screen a *Ciona* genomic DNA library. Two different classes of genomic clones were identified. One corresponds to the sequence contained in the *Ci-fkh* cDNA, the main focus of the present study. In a recently proposed unified nomenclature for the winged helix/forkhead proteins, Ci-Fkh has been named “FoxA5” (Kaestner *et al.*, 2000).

The second class, which we tentatively call *Ci-fkh-B*, differs by 10 aa residues within the winged-helix domain (A. Di Gregorio, unpublished results), although both domains fall into class 1 defined by Kaufmann and Knochel (1996). All members of this class contain the following conserved residues: A(9), L(43), Q(51), N(92), and C(98). Notwithstanding the high similarity of their DNA-binding domains, the *Ci-fkh* and *Ci-fkh-B* C-terminal sequences are quite divergent. Moreover, both sequences are only distantly related to a Forkhead-related ORF identified in a *Ciona* genomic survey (GenBank Accession No. Z83861). Although we cannot rule out the possibility that *Ci-fkh-B* represents an allelic variant of *Ci-fkh*, we propose that it represents a separate gene. This hypothesis is currently being investigated.

The predicted protein encoded by the *Ci-fkh* gene shares 43% amino acid identity with the mouse *HNF-3β* gene and 40% identity with the *Drosophila forkhead* gene. Particularly strong conservation is seen in the winged-helix DNA binding domain, which encompasses 110 amino acid residues (Kaufmann and Knochel, 1996). In this domain, Ci-Fkh shares ~90% amino acid identity with HNF-3β and 85% identity with *Drosophila* Fkh. In addition to the high conservation of the winged-helix domain, the putative Ci-Fkh protein sequence includes a conserved motif, HPFSI, that has been implicated in transactivation (Pani *et al.*, 1992b). The HNF-3β nuclear localization signals are also conserved in the *Ciona* protein (Qian and Costa, 1995).

Ci-fkh Expression Pattern

Northern assays suggest that *Ci-fkh* expression is activated quite early during embryogenesis, at the 16-cell stage (data not shown). Peak accumulations of the *Ci-fkh* mRNA are detected during gastrulation, and expression steadily declines during tail-bud and tadpole stages of development. *In situ* hybridization assays reveal that *Ci-fkh* transcripts are initially distributed in a broad pattern in 16- and 32-cell embryos (Figs. 1A–1D). Staining is detected in the progenitors of the trunk epidermis, palps, brain, and pigment cells (the “small a” lineages, see Fig. 1D). Expression is also observed in the progenitors of the notochord, spinal cord, and endoderm (the “capital A and B” lineages, see Figs. 1B and 1F). This broad staining pattern persists throughout the 32- and 64-cell stages (Figs. 1A–1F), but is rapidly refined during the early phases of gastrulation (Figs. 1G and 1H).

Ci-fkh RNAs are lost from the progenitors of the trunk epidermis and gradually diminish in the progenitors of the CNS and trunk mesenchyme during gastrulation and neurulation (Figs. 1E–1H). Transcripts persist at high levels in the progenitors of the notochord (Figs. 1G and 1H; red cells in 1H); weaker staining is detected in portions of the presumptive endoderm (Fig. 1H, yellow cells) and CNS (blue cells). Low levels of expression are detected in the B8.5 blastomeres, which give rise to trunk mesenchyme (light blue cells in Fig. 1H). During neurulation, a new site of staining appears in the developing neural tube (Figs. 1I and 1J; blue cells in 1J), while strong staining persists in the differentiating notochord (red cells in Fig. 1J). This basic pattern, expression in the notochord, endoderm, and CNS, persists during tail-bud formation (Figs. 1K and 1L). Staining in the CNS is restricted to the posteriormost region of the cerebral vesicle and to the ventral ependymal cells of the spinal cord (Fig. 1L, blue cells).

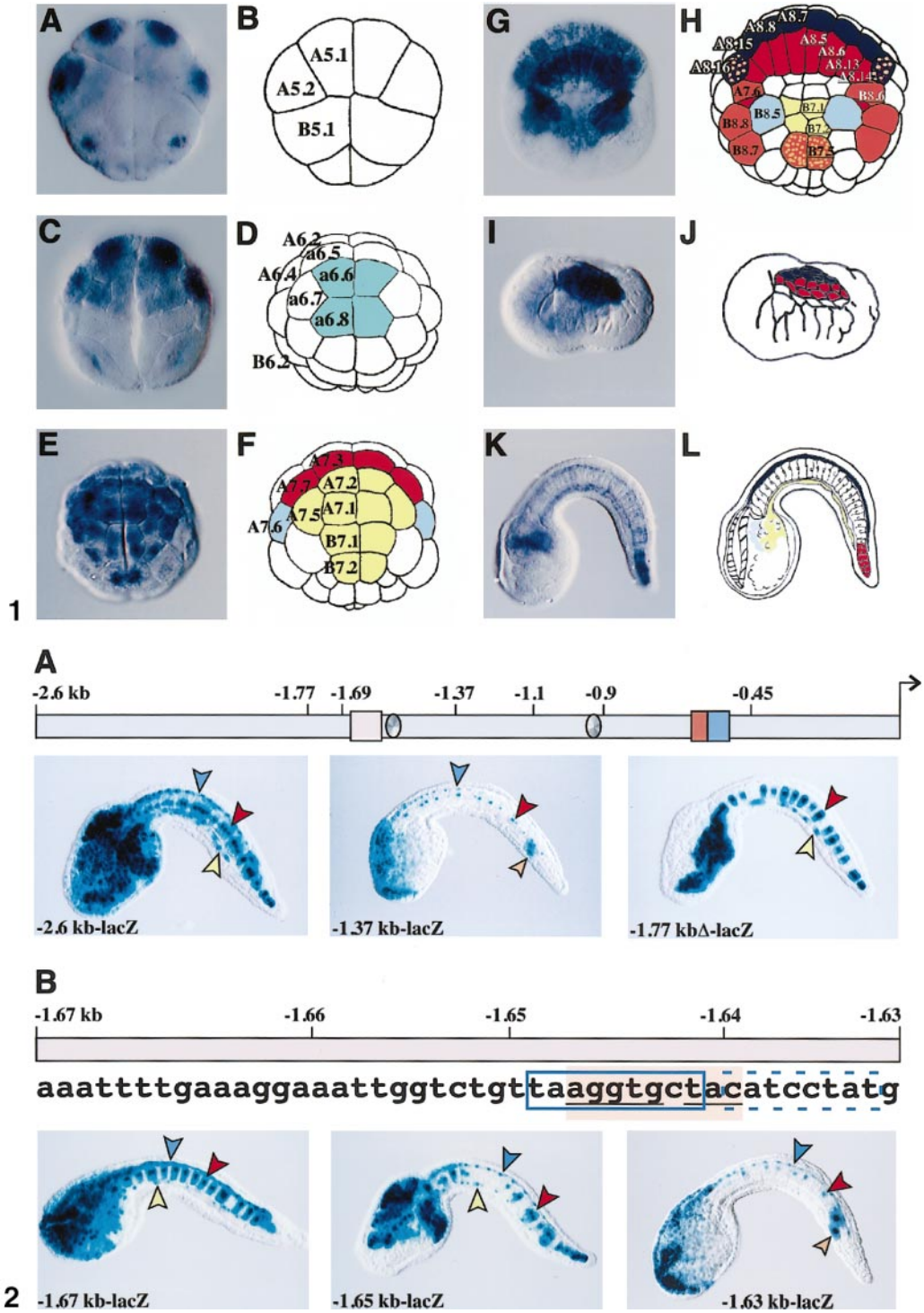


FIG. 1. *In situ* hybridization assays. A digoxigenin-labeled *Ci-fkh* antisense RNA probe was hybridized to whole-mount preparations of staged *Ciona* embryos. Only the blastomeres expressing *Ci-fkh* are labeled. (A) Vegetal view of a 16-cell embryo. (B) Schematic drawing indicating cell lineages. (C) Animal view of a 32-cell embryo. (D) The stained blastomeres will form epidermis and notochord. (E) Vegetal view of a 64-cell embryo. (F) The drawing shows the precursors of different larval tissues. Here and in later figures the following color code is used: red, notochord; yellow, endoderm; light blue, trunk lateral cells (mesenchyme); green, epidermis. Only the fate-restricted blastomeres are colored. (G) Vegetal view of a 110-cell embryo. (H) The drawing represents the definitive fate map. Neuroectodermal precursors are indicated in dark blue. The dotted labeling of the A8.16 pair indicates that these cells give rise to the lateral ependymal cells

Characterization of a Shared cis-Regulatory Sequence That Mediates Expression in Axial Tissues

Transgenes containing 5.6 or 2.6 kb of the *Ci-fkh* 5′-flanking region exhibit the complete staining pattern in tadpole-stage embryos, including expression in the CNS, notochord, endoderm (indicated by blue, red, and yellow arrowheads, respectively; left in Fig. 2A), and occasionally in the mesenchyme. The 1.77-kb transgene exhibits a similar expression pattern, although it sometimes directs abnormal expression in the trunk epidermis (data not shown). Further truncations of the promoter region disrupt the normal *Ci-fkh* expression pattern. A 1.37-kb region of the 5′-flanking sequence directs a highly abnormal *lacZ* staining pattern, including a severe reduction in notochord staining, and a complete loss of the endoderm pattern (middle, Fig. 2A). In addition, there is ectopic staining in the trunk epidermis and expanded expression in the CNS. Concomitant with the expanded CNS pattern is the ectopic expression of truncated transgenes in two pairs of secondary muscle cells (brown arrowhead) that derive from the same lineage as the lateral ependymal cells of the neural tube (the A7.8/A8.16 lineage). Serial truncations between −1.37 kb and −900 bp upstream of the *Ci-fkh* transcription start site eliminate expression in the CNS (see Table 1), thereby raising the possibility that there is a discrete CNS enhancer in this region. To test this idea we examined the staining pattern directed by a *Ci-fkh/lacZ* transgene that contains 1.77 kb of 5′-flanking sequence and an internal deletion between −1.37 kb and −900 bp (right in Fig. 2A). This transgene directs expression solely within the notochord and gut; no staining is detected in either the cerebral vesicle or the spinal cord. These results are consistent with the notion that there are separate, but closely linked, elements mediating expression in the CNS, notochord, and endoderm (see below).

The preceding results suggest that the region between −1.77 and −1.37 kb contains both positive and negative regulatory elements that are important for expression in all three axial tissues. Additional truncations (see Table 1) suggest that at least some of these elements are located in a 44-bp interval, between −1.67 and −1.63 kb (Fig. 2B). A *Ci-fkh/lacZ* transgene containing 1.67 kb of 5′-flanking sequence directs an essentially normal staining pattern (Fig. 2B, left). Expression is detected in the notochord (red arrowhead), endoderm (yellow arrowhead), and CNS (blue arrowhead). Truncating the distalmost 24 bp (middle, Fig. 2B) results in a consistent and severe reduction in notochord staining (indicated by a red arrowhead; compare with the left side). Expression in the endoderm may be slightly reduced, but is essentially normal (yellow arrowhead). The removal of another 20 bp (right in Fig. 2B) leads to a further reduction of the notochord pattern, a virtual loss of endoderm expression, and strong ectopic expression in the trunk epidermis and CNS. As shown previously (middle, Fig. 2A), the expanded staining in the CNS is accompanied by ectopic expression in the A7.8/A8.16 tail muscles (brown arrowhead).

The nucleotide sequence of this 44-bp region is presented in Fig. 2B. It includes a potential Brachyury/T-domain half-site (boxed in orange; key conserved residues are underlined) and a cluster of two potential Snail repressor sites

---

and secondary muscle cells. Similarly, the dotted labeling of the B7.5 pair indicates a mixed fate (endoderm and muscle derivatives). Staining is detected in all the notochord precursor cells (red), the endoderm (yellow), and mesenchyme (light blue). (I, J) Neurulating embryo, side view; anterior is to the left, dorsal is up. *Ci-fkh* transcripts are detected in notochord (red), CNS precursors (neural folds, blue), and endoderm (not visible). (K, L) Lateral view of a tadpole, ~14 h after fertilization. *Ci-fkh* expression is still strong in the spinal cord, endoderm, mesenchyme, and secondary notochord cells.

FIG. 2. Identification of distal and proximal *cis*-regulatory elements. (A) The horizontal bar on top represents the *Ci-fkh* 5′-regulatory region. The numbers refer to distances 5′ of the transcribed region. The pink box identifies the location of the AS element. The three larvae were electroporated with different fragments of the 5′-flanking sequence at the 1-cell stage (see Materials and Methods). Reporter gene expression was visualized by histochemical staining with X-gal. The larva to the left contains the "full-length" −1.77 kb *Ci-fkh* promoter region. Staining is detected in the notochord (red arrowhead), endoderm (yellow arrowhead), and CNS (blue arrowhead). The central larva contains the truncated, −1.37 kb *Ci-fkh/lacZ* fusion gene. Ectopic staining is detected in the CNS, A8.16 muscles (orange arrowhead), and trunk epidermis. Only residual expression is detected in the notochord, whereas staining in the endoderm is lost. The fusion gene to the right contains an internal, 470-bp deletion. *lacZ* staining is detected in the notochord (red arrowhead) and endoderm (yellow arrowhead), but is completely absent in the CNS and trunk epidermis. The orange box indicates a 40-bp sequence required for expression in notochord and endoderm (see Table 2). The adjacent blue box indicates a 60-bp region containing a lateral CNS activator. The ovals indicate two Ci-Fkh autoregulatory sites (see Fig. 4). (B) The horizontal bar on top represents a "blow-up" of the pink box in A. The numbers refer to distances 5′ of the transcribed region. The sequence of this 44-bp region is shown below the bar. The underlined region (boxed in orange) identifies the putative T box and the two putative Snail binding sites are boxed in blue. The three larvae were electroporated with sequentially truncated *Ci-fkh/lacZ* transgenes. The larva to the left contains −1.67 kb of 5′-flanking sequence. An essentially normal pattern is observed in notochord and endoderm variable staining in either the ventral or the lateral ependymal cells of the neural tube is observed. The central larva contains a fusion gene lacking just the distal-most 24 bp. There is reduced expression in the notochord (red arrowhead) and endoderm (yellow arrowhead). Occasionally, staining in the lateral ependymal cells is detected. The larva to the right contains a truncated fusion gene lacking another 20 bp. There is a dramatic loss of staining in the endoderm. In addition, strong ectopic expression is consistently observed in the A8.16 muscles (orange arrowhead) and lateral ependymal cells.

(indicated by solid and dashed blue boxes). Additional experiments were done to determine whether either sequence is important for *Ci-fkh* expression.

Transcriptional Repression Is Essential for the Normal Ci-fkh Pattern in the CNS

Ci-fkh is normally expressed in the ventral-most ependymal cells of the spinal cord, which have been proposed to represent a rudimentary floor plate (Corbo *et al.*, 1997b; Olsen and Jeffery, 1997; Shimauchi *et al.*, 1997). Optical sections of the tail region of hybridized tadpoles reveal the restricted expression of the endogenous gene in the ventral ependymal cell (Figs. 3A–3C). A similar expression pattern is observed for the "full-length" *Ci-fkh/lacZ* transgenes containing either 5 or 3 kb of the 5′-flanking region (data not shown). In contrast, truncated promoters containing less than 1.7 kb of the *Ci-fkh* 5′-flanking region exhibit abnormal patterns of gene expression in the CNS (Figs. 3D–3F). Staining is lost in the ventral ependymal cells and replaced by expression in lateral ependymal cells (Fig. 3E; summarized in Fig. 3F). These results suggest that distal regions of the *Ci-fkh* promoter contain one or more repressor elements responsible for restricting expression to the floor plate. Removal of these elements causes ectopic expression in the two lateral rows of ependymal cells. The same region also contains sequences that are important for the activation of *Ci-fkh* expression in ventral ependymal cells.

The extensive lineage information available in *Ciona* provides strong support for the notion that the truncated promoters direct ectopic expression in the spinal cord. The expanded expression patterns obtained with truncated *Ci-fkh* transgenes are associated with ectopic staining of a specific pair of secondary muscle cells (orange arrowhead, Fig. 2B, right). These muscles are derived from the A8.16 blastomere pair of the 110-cell embryo, which also gives rise to the lateral ependymal cells.

As shown previously (Fig. 2B, middle), an ~24-bp truncation within the AS, between −1.67 and −1.65 kb, alters the staining patterns directed by *Ci-fkh/lacZ* transgenes. The removal of an additional 20 bp (from −1.65 to −1.63) causes a loss of staining in the endoderm and expanded expression in the CNS. This 20-bp region contains a T-box recognition sequence partially overlapping with a cluster of divergent Ci-Sna binding sites (see Fig. 2B). Previous studies have shown that *Ci-sna* is transiently expressed in the lateral ependymal cells of the spinal cord (Corbo *et al.*, 1997b), and DNA binding assays were done to determine whether it might function as a repressor of *Ci-fkh* in these regions.

A GST/Ci-Snail fusion protein (Fujiwara *et al.*, 1998) was mixed with a labeled oligonucleotide containing the putative Sna sites from the *Ci-fkh* AS element, which is denoted "Sna Cryp" (Fig. 3G). A shifted complex is observed using the labeled Sna Cryp sequence as a probe. There is a reduction in the complexes upon addition of increasing amounts of cold Sna Cryp as a competitor DNA. Even more efficient inhibition is obtained with an unlabeled high-affinity *Drosophila* Snail recognition sequence (Ip *et al.*, 1992). The protein–DNA complex is not noticeably altered when a Ci-Fkh recognition sequence, CIF1, is used as a cold competitor. These experiments establish a close correlation between the deletion of two weak Ci-Snail binding sites between −1.65 and −1.63 kb (indicated by blue boxes in Fig. 2B) and expanded expression of *Ci-fkh/lacZ* transgenes in the neural tube. As noted earlier, this expanded staining is accompanied by ectopic expression in the A8.16 tail muscles.

As mentioned above, a *cis*-regulatory sequence located between −1.37 kb and −0.9 kb is required for maintaining the ectopic expression in the lateral regions of the neural tube (Fig. 2A, right). We narrowed this sequence interval down to 60 bp and found that it contains a Ci-Fkh autoregulatory site, CIF3 (see Figs. 4D and 4E). In addition to this, fusion genes lacking the region between −0.61 and −0.55 kb do not show CNS expression (cf. Table 2). Sequence analysis indicates the presence of a putative binding site for a Nkx-related factor.

The sequence interval between −1.77 and −1.65 kb contains the *cis*-regulatory sequence responsible for the normal expression in the ventral ependymal cells. In mice, it has been shown that a Gli binding site is responsible for the expression in the floor plate (Sasaki *et al.*, 1997). A putative Gli binding site (8/9 matches with the site found in the mouse HNF-3β promoter) is adjacent to the AS sequence, but it does not appear to be directly controlling *Ci-fkh* expression in the ventral ependymal cells.

cis-Regulation in Notochord and Endoderm

The proximal 20 bp of the 44-bp AS element are important both for repressing *Ci-fkh* expression in the lateral ependymal cells and for optimal expression in the endoderm (Fig. 2B). Truncations that remove a potential T-box recognition sequence (boxed in orange in Fig. 2B) essentially eliminate staining in the endoderm, but not in the notochord. This sequence shares 7 of 10 matches with the optimal Brachyury half-site (AGGTGTGAAA; Kispert and Herrmann, 1993) and binds a GST/Ci-Bra fusion protein (data not shown). It also shares an 8/10 match with an essential Ci-VegTR binding site (T2) within the minimal *Ci-sna* muscle enhancer (Erives and Levine, 2000). *Ci-VegTR* is expressed in vegetal regions of early embryos and might play a role in the specification of muscle and endoderm lineages (see Discussion). Point mutations in the T-box site abolish endodermal expression of a −1.65 kb fusion gene but do not affect notochord staining (data not shown). Notochord expression relies on a distal *cis*-regulatory sequence, located between −1.67 and −1.64 kb, and on a proximal sequence, located between −0.67 and −0.61 kb. The latter sequence is also necessary for normal endodermal expression and ectopic trunk epidermal staining (see Table 2) and its sequence contains matches for the endodermal marker CdxA (e.g., Frumkin *et al.*, 1991). The faint, residual notochord staining observed in the −1.49 to

−0.96 kb truncation series is completely lost in truncations from −0.96 kb on, thus indicating that this *cis*-regulatory region might be required to "stabilize" notochord expression.

Autoregulation

As a first step toward determining whether Ci-Fkh mediates autoregulation, DNA binding assays were performed with a GST/Ci-Fkh fusion protein and different regions of the *Ci-fkh* 5′-regulatory region. Efforts focused on distal sequences that encompass the AS, as well as more proximal DNA sequences located between −960 and −900 bp. This latter region is important for expression in the CNS and possibly in the notochord (see below; summarized in Table 1). Inspection of the DNA sequences in these regions identified two potential matches to the vertebrate HNF-3β recognition sequence, A(A/T)TRTT(G/T)RYTY (Qian and Costa, 1995, and references therein). One of the putative Ci-Fkh binding sites, CIF1, is located near the distal AS, while the other sequence, CIF3, is located within the proximal CNS regulatory region.

DNA binding assays were done with labeled CIF1 (Fig. 4A) and CIF3 (Fig. 4E) probes. Shifted complexes were formed upon addition of the GST/Ci-Fkh fusion protein. Addition of increasing amounts of unlabeled CIF1 or CIF3 (5-, 50-, and 100-fold molar excess) resulted in a substantial reduction in these complexes. In contrast, neither the Sna Cryp binding site (Figs. 4A and 4E) nor a mutagenized version of CIF1 (CIF1m) significantly diminished the amount of shifted complexes (Fig. 4A). Shifted complexes were not detected when the CIF1m sequence was used as a probe.

Coupled *in vitro* mutagenesis and electroporation assays were done to determine whether the CIF1 site, adjacent to the critical AS, is important for the expression of *Ci-fkh/lacZ* transgenes. The CIF1 binding site was mutagenized in the context of an otherwise normal −1.77 kb *Ci-fkh/lacZ* transgene, introducing the same point mutations previously tested *in vitro* (CIF1m). Normally, the −1.77 kb transgene directs the complete *Ci-fkh* staining pattern, including expression in the CNS, endoderm, and notochord. A low-magnification view of electroporated embryos shows the reproducibility of the staining pattern among a large number of embryos (Fig. 4B). The same *Ci-fkh/lacZ* transgene containing point mutations in the CIF1 site (the CIF1m sequence used in Fig. 4A) directed consistently weaker staining in all the tissues in which *Ci-fkh* is normally expressed (Fig. 4C). This result is consistent with the notion that *Ci-fkh* autoregulation is important for optimal expression of the gene during tail-bud and tadpole stages of development.

The analyses of additional truncations and internal deletions (see Tables 1 and 2) suggest that the CIF3 site is also important for *Ci-fkh* expression (Fig. 4D). A *Ci-fkh/lacZ* transgene containing 960 bp of the 5′-flanking region exhibited strong expression in the A8.16 muscle cells and lateral ependymal cells. Ectopic staining was also detected in the cerebral vesicle and trunk epidermis (presumably due to the loss of the repressor element in the AS). The removal of another ~60 bp from the 5′-regulatory region resulted in a complete loss of this staining pattern (Fig. 4D, bottom panel). The deleted region includes the CIF3 binding site, suggesting that this site might be important for Ci-Fkh autoregulation in the CNS.

DISCUSSION

We have presented evidence that the regulation of *Ci-fkh* expression in the notochord, endoderm, and floor plate is mediated by a compact region of the 5′-flanking sequence. A shared element, the AS, was shown to influence expression in all three tissues. This supports the proposal that the floor plate, notochord, and underlying gut arise from a common population of progenitor cells in early chick embryos (Catala *et al.*, 1996; Teillet *et al.*, 1998). DNA binding assays, along with coupled *in vitro* mutagenesis and electroporation assays, suggest that the Ci-Fkh protein might mediate autoregulation and that the Ci-Snail repressor might exclude expression in the lateral ependymal cells and restrict the pattern to the rudimentary floor plate. We also discuss the possibility that a T-box transcription factor regulates *Ci-fkh* expression in the endoderm.

Regulation of the CNS Pattern

Ci-fkh is specifically expressed in the ventral row of ependymal cells in the spinal cord. This observation prompted the suggestion that the ascidian CNS contains a rudimentary floor plate (Corbo *et al.*, 1997b; Olsen and Jeffery, 1997; Shimauchi *et al.*, 1997). Serial truncations of the *Ci-fkh* 5′-regulatory region result in an expanded expression pattern, whereby *Ci-fkh/lacZ* transgenes are ectopically expressed in lateral ependymal cells. Additional truncations in the *Ci-fkh* promoter region result in the specific loss of the ventral-most pattern, without affecting ectopic expression in the lateral ependymal cells. Thus, it is possible to separate, at least in part, expression in the ventral and lateral ependymal cells.

We suggest that one or more repressor elements, located in the 44-bp AS element, may be responsible for excluding *Ci-fkh* expression from the lateral ependymal cells. *Ci-fkh* is initially expressed in a broad pattern that includes the entire presumptive CNS and trunk epidermis. It is therefore possible that the repressor(s) shuts down aspects of the early pattern. It is also possible that separate repressors are required to keep *Ci-fkh* expression off in inappropriate lineages, such as the lateral ependymal cells.

The most likely location of a repressor element is between −1.65 and −1.63 kb upstream of the *Ci-fkh* transcription unit, within the AS element. The removal of this 20-bp sequence leads to expanded expression in the lateral ependymal cells and the staining of a characteristic pair of muscle cells which derive from the same lineage (Satoh,

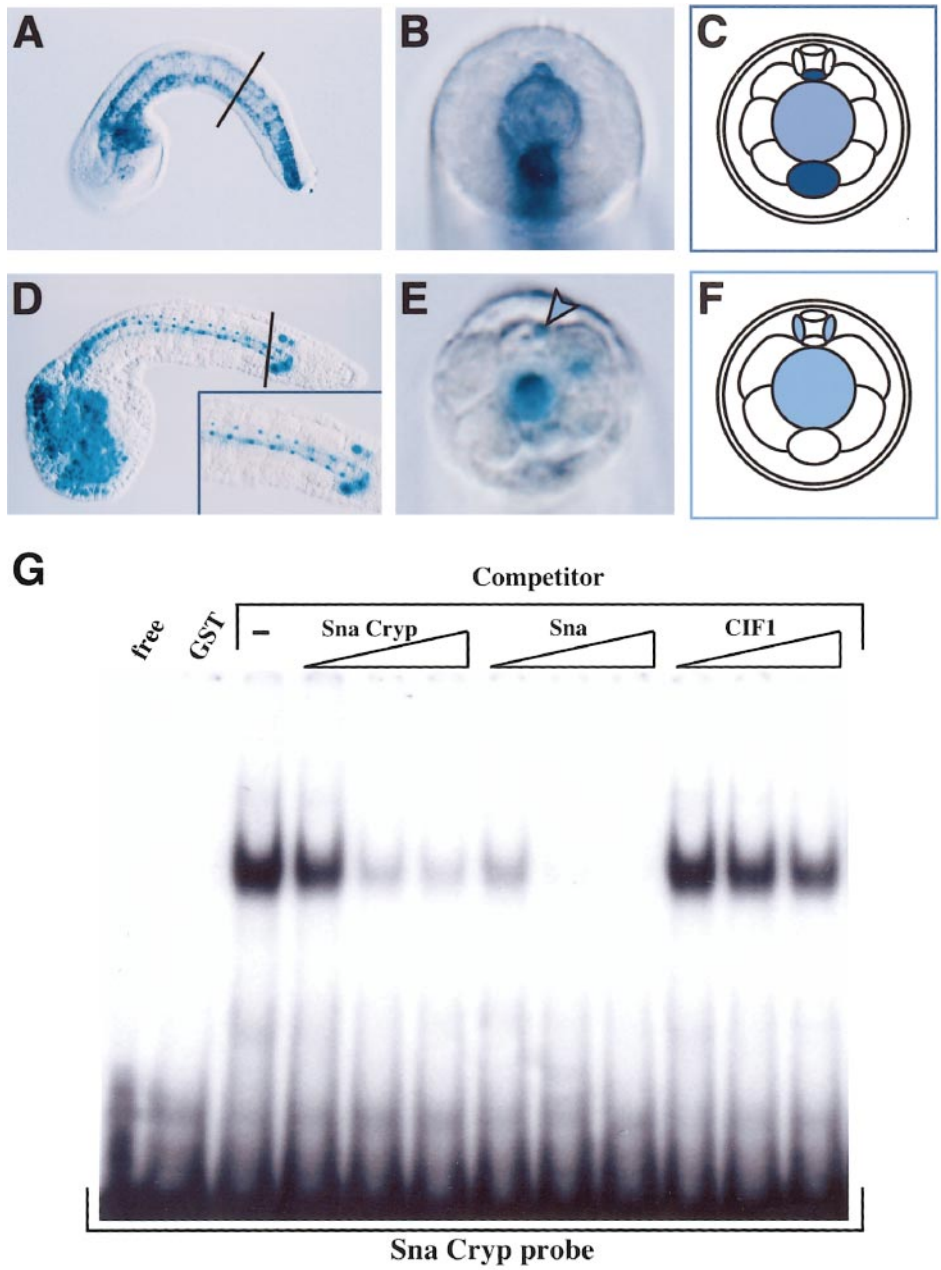


FIG. 3. Regulation of *Ci-fkh* in the CNS. (A–C) Tadpole-stage embryos were hybridized with a digoxigenin-labeled *Ci-fkh* antisense RNA probe. Staining is detected in the spinal cord, notochord, and endodermal strand (A). The line through the tail indicates the approximate position of the optical section shown in (B). Staining is restricted to the ventral-most row of ependymal cells within the spinal cord. Expression is also seen in the large central notochord and lower endodermal strand. (C) Diagram of the stained tissues. (D–F) Tadpole-stage embryos that were electroporated with a truncated *Ci-fkh/lacZ* fusion gene containing 1.37 kb of the 5′-flanking region. Staining is detected in the trunk epidermis, cerebral vesicle, spinal cord, and A8.16 tail muscles (D). The inset shows a high-magnification view of the ectopic staining in the tail muscles. The line indicates the approximate position of the optical section shown in (E). Staining is detected in the lateral ependymal cells. There is no obvious staining in the ventral row of cells (see diagram in F). (G) Gel shift assays using the Sna Cryp probe and a GST/Ci-Sna fusion protein (Fujiwara and Levine, 1998). A shifted complex is observed ("−" lane). This protein–DNA complex is diminished upon addition of increasing (5-, 50-, and 100-fold) molar excesses of cold Sna Cryp competitor DNA. There is a more rapid loss of the Ci-Sna/Sna Cryp protein–DNA complex upon addition of increasing amounts of the high-affinity s2 *Drosophila* Snail binding site ("Sna"). There is only a very slight diminishment in the amount of shifted complex upon addition of increasing amounts of the control CIF1 Ci-Fkh binding site.

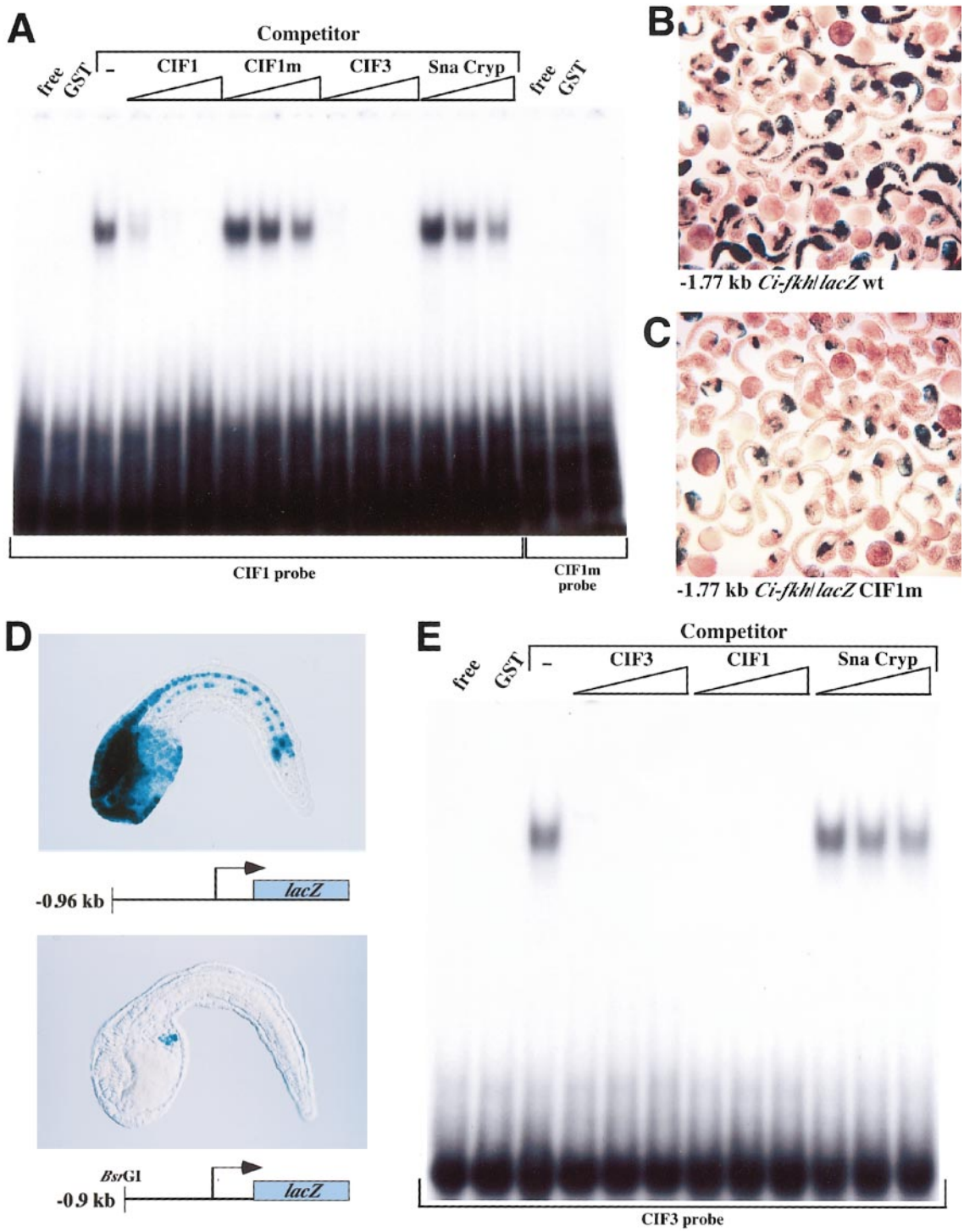


FIG. 4. *Ci-fkh* autoregulation. (A) Gel shift assays using a labeled CIF1 DNA probe and a GST/Ci-Fkh fusion protein. A shifted protein–DNA complex is obtained in the absence of competitor DNA ("−" lane). Increasing amounts of cold competitor CIF1 DNA lead to the progressive loss of the shifted complex. There is virtually no loss in the complex upon addition of either a mutant form of the cold CIF1 DNA (CIF1m) or the heterologous Sna Cryp binding site. There is a rapid loss in the complexes with increasing amounts of the CIF3 DNA competitor. (B) Low-magnification field of tadpole stage embryos that were electroporated with the −1.77 kb *Ci-fkh/lacZ* fusion gene. Staining is detected in the head and tail and includes the notochord, CNS, and endoderm. (C) Same as B except that the *Ci-fkh/lacZ* fusion gene contains point mutations in the CIF1 site (CIF1m; same sequence as used in the binding assays in A). There is reduced staining in all of the tissues in which *Ci-fkh* is normally expressed. This reduction is particularly evident in the tails. (D) Tadpoles that express the indicated *Ci-fkh/lacZ* transgenes. The embryo on top exhibits strong ectopic staining in the trunk epidermis, lateral ependymal cells of the spinal cord, and A8.16 tail muscles. In contrast, the deletion of just 60 bp results in a virtual loss of this staining pattern. The deleted region contains the CIF3 binding site. (E) Gel shift assays using the CIF3 DNA as a labeled probe. A protein–DNA complex is formed upon addition of the GST/Ci-Fkh fusion protein ("−" lane). This complex is rapidly lost with increasing amounts of cold competitor CIF3 or CIF1 DNA. The complex is not affected by increasing amounts of the heterologous Sna Cryp competitor DNA.

1994). This region contains two clustered, divergent Ci-Snail binding sites. Further evidence for interactions between the Ci-Snail repressor and the *Ci-fkh* 5′-regulatory region is the observation that *Ci-snail* is transiently expressed in the lateral ependymal cells during neurulation (Corbo *et al.*, 1997b). Alternatively, it is possible that another Snail-related repressor might regulate *Ci-fkh* expression.

Studies in vertebrate embryos suggest that the dorsal ectoderm signals the induction of *snail*-related genes, such as *slug*, in dorsal regions of the neural tube (e.g., Dickinson *et al.*, 1995; La Bonne and Bronner-Fraser, 1998). It is possible that Slug is important for restricting *HNF-3β* expression to the floor plate, similar to the situation we have encountered in *Ciona*.

Autoregulation

The *Ci-fkh* 5′-regulatory region contains a number of potential binding sites for Ci-Fkh. Point mutations in one of these sites, CIF1, cause a general reduction in all aspects of the staining pattern, with reduced expression in gut, notochord, and CNS. Moreover, a 60-bp deletion that removes the proximal CIF3 site causes a loss of the CNS and notochord patterns directed by a truncated *Ci-fkh/lacZ* fusion gene. Autoregulation may be the basis for the apparently heavy reliance on transcriptional repression. *Ci-fkh* is initially expressed in a broad pattern that encompasses most of the presumptive neurogenic ectoderm. Spatially (and temporally) localized repressors may be required to block autoregulation and restrict *Ci-fkh* expression to the rudimentary floor plate. The Sna Cryp repressor sites within the AS element map less than 40 bp from the distal CIF1 autoregulatory element, suggesting that Ci-Sna might function as a short-range repressor as seen for its counterpart in *Drosophila* (Gray *et al.*, 1994; Gray and Levine, 1996).

It is possible that the Ci-Fkh protein acts in concert with tissue-specific and lineage-specific regulatory factors to maintain the *Ci-fkh* expression pattern during development. For example, the distal CIF1 site maps near a putative T-box element within the AS, suggesting that Ci-Fkh works with a T-box transcription factor such as Ci-VegTR to mediate expression in the endoderm (Erives and Levine, 2000; see below).

Ci-VegTR is a newly identified maternal T-box gene that is related to *Xenopus VegT*. The latter gene is expressed in vegetal regions of early embryos and plays a key role in the specification of endodermal lineages (Clements *et al.*, 1999). Similarly, maternal Ci-VegTR mRNAs become localized within vegetal regions of fertilized *Ciona* eggs and are incorporated in blastomeres that give rise to muscle and endoderm lineages. Previous studies suggest that Ci-VegTR is important for the activation of target genes within B4.1 lineages that form most of the tail muscles in the tadpole. It is conceivable that Ci-VegTR is also important for the activation of target genes within endodermal lineages. In this regard we note that the T-box sequence located within the *Ci-fkh* AS element, AGGTGCTACA, is closely related to an essential Ci-VegTR element within the B4.1 enhancer of the *Ci-sna* gene (AGGTGCCAAA; Erives and Levine, 2000). We have shown that the deletion of the T-box motif within the AS results in the loss of *Ci-fkh/lacZ* expression in the endoderm (Fig. 2).

Comparison of HNF-3β and Ci-fkh Regulation

The mouse *HNF-3β* gene and the *Ciona Ci-fkh* gene exhibit strikingly similar patterns of expression. Both genes are expressed in all three axial tissues: the floor plate, notochord, and gut. In mice, separate enhancers mediate expression in the node/notochord and floor plate. These enhancers are located far upstream and downstream of the *HNF-3β* transcription unit. In contrast, the regulation of *Ci-fkh* expression depends on tightly linked *cis*-elements in the 5′-flanking region. Indeed, the 44-bp AS element is important for all aspects of the normal expression pattern.

Despite the compaction of the *Ci-fkh cis*-regulatory elements, there are similarities with the regulation of *HNF-3β* in mice. Expression in the endoderm and notochord depends on the distal AS element, whereas CNS expression is mediated, at least in part, by a separate element located in a more proximal region of the 5′-flanking sequence. Moreover, the detailed dissection of the −1.77 to −1.63 kb interval of the *Ci-fkh* promoter region resulted in the partial uncoupling of expression in the notochord and endoderm. The notochord/node and CNS enhancers are separated by nearly 20 kb in mice, whereas apparently comparable *cis*-regulatory elements are located within 1 kb of one another in *Ciona*. This streamlining in the organization of *cis*-elements is reminiscent of the situation encountered in the pufferfish genome (e.g., McLysaght *et al.*, 2000). The entire *Ciona* genome is only ~1/20 as large as mammalian genomes. One reason for this dramatic difference is that *Ciona* has not undergone the genome-wide duplication events seen in vertebrates, so there are many fewer genes (Simmen *et al.*, 1998). This study provides evidence that the small genome size also reflects the tight packaging of *cis*-regulatory elements. *Ciona* should be useful for the rapid identification and characterization of *cis* elements that specify basic chordate tissues.

ACKNOWLEDGMENTS

We thank Dr. Shigeki Fujiwara for his help in identifying the putative Ci-Sna binding sites. We thank Dr. Elena Silva Casey for critically reading the manuscript and Dr. Yutaka Nibu and Dave Keys for help in figure preparation and sequence analysis. This work was supported by a grant from the NSF (IBN-9514138).

REFERENCES

Ang, S., and Rossant, J. (1994). *HNF-3β* is essential for node and notochord formation in mouse development. *Cell* **78,** 561–574.

Catala, M., Teillet, M. A., De Robertis, E. M., and Le Douarin, N. M. (1996). A spinal cord fate map in the avian embryo: While

regressing, Hensen's node lays down the notochord and floor plate thus joining the spinal cord lateral walls. *Development* **122,** 2599–2610.

Clements, D., Friday, R. V., and Woodland, H. R. (1999). Mode of action of VegT in mesoderm and endoderm formation. *Development* **126,** 4903–4911.

Corbo, J. C., Levine, M., and Zeller, R. W. (1997a). Characterization of a notochord-specific enhancer from the *Brachyury* promoter region of the ascidian, *Ciona intestinalis*. *Development* **124,** 589–602.

Corbo, J. C., Erives, A., Di Gregorio, A., Chang, A., and Levine, M. (1997b). Dorsoventral patterning of the vertebrate neural tube is conserved in a protochordate. *Development* **124,** 2335–2344.

Corbo, J. C., Fujiwara, S., Levine, M., and Di Gregorio, A. (1998). Suppressor of Hairless activates *Brachyury* expression in the *Ciona* embryo. *Dev. Biol.* **203,** 358–368.

Dickinson, M. E., Selleck, M. A., McMahon, A. P., and Bronner-Fraser, M. (1995). Dorsalization of the neural tube by the non-neural ectoderm. *Development* **121,** 2099–2106.

Di Gregorio, A., and Levine, M. (1998). Ascidian embryogenesis and the origins of the chordate body plan. *Curr. Opin. Genet. Dev.* **8,** 457–463.

Duncan, S. A., Navas, M. A., Dufort, D., Rossant, J., and Stoffel, M. (1998). Regulation of a transcription factor network required for differentiation and metabolism. *Science* **281,** 692–695.

Erives, A., and Levine, M. (2000). Characterization of a maternal T-box gene in *Ciona intestinalis*. *Dev. Biol.* **225,** 169–178.

Frumkin, A., Rangini, Z., Ben-Yehuda, A., Gruenbaum, Y., and Fainsod, A. (1991). A chicken caudal homologue, CHox-cad, is expressed in the epiblast with posterior localization and in the early endodermal lineage. *Development* **112,** 207–219.

Fujiwara, S., Corbo, J. C., and Levine, M. (1998). The Snail repressor establishes a muscle/notochord boundary in the *Ciona* embryo. *Development* **125,** 2511–2520.

Gray, S., Szymanski, P., and Levine, M. (1994). Short-range repression permits multiple enhancers to function autonomously within a complex promoter. *Genes Dev.* **8,** 1829–1838.

Gray, S., and Levine, M. (1996). Short-range transcriptional repressors mediate both quenching and direct repression within complex loci in *Drosophila*. *Genes Dev.* **10,** 700–710.

Guan, K. L., and Dixon, J. E. (1991). Eukaryotic proteins expressed in *Escherichia coli:* An improved thrombin cleavage and purification procedure of fusion proteins with glutathione *S*-transferase. *Anal. Biochem.* **192,** 262–267.

Ip, Y. T., Park, R., Kosman, D., and Levine, M. (1992). The *dorsal* gradient morphogen regulates stripes of *rhomboid* expression in the presumptive neuroectoderm of the *Drosophila* embryo. *Genes Dev.* **6,** 1728–1739.

Kaestner, K. H., Knochel W., and Martinez, D. E. (2000). Unified nomenclature for the winged helix/forkhead transcription factors. *Genes Dev.* **14,** 142–146.

Kaufmann, E., and Knochel, W. (1996). Five years on the wings of fork head. *Mech. Dev.* **57,** 3–20.

Kispert, A., and Herrmann, B. G. (1993). The *Brachyury* gene encodes a novel DNA binding protein. *EMBO J.* **12,** 3211–3220.

La Bonne, C., and Bronner-Fraser, M. (1998). Neural crest induction in *Xenopus:* Evidence for a two-signal model. *Development* **125,** 2403–2414.

Lai, E., Prezioso, V. R., Tao, W. F., Chen, W. S., and Darnell, J. E., Jr. (1991). Hepatocyte nuclear factor 3 alpha belongs to a gene family in mammals that is homologous to the *Drosophila* homeotic gene *fork head*. *Genes Dev.* **5,** 416–427.

Lofberg, J., and Collazo, A. (1997). Hypochord, an enigmatic embryonic structure: Study of the Axolotl embryo. *J. Morphol.* **232,** 57–66.

McLysaght, A., Enright, A. J., Skrabanek, L., and Wolfe, K. H. (2000). Estimation of synteny conservation and genome compaction between pufferfish (*Fugu*) and human. *Yeast* **17,** 22–36.

Olsen, C. L., and Jeffery, W. R. (1997). A forkhead gene related to *HNF-3β* is required for gastrulation and axis formation in the ascidian embryo. *Development* **124,** 3609–3619.

Pani, L., Qian, X., Clevidence, D., and Costa, R. H. (1992a). The restricted promoter activity of the liver transcription factor hepatocyte nuclear factor 3*β* involves a cell-specific factor and positive autoactivation. *Mol. Cell. Biol.* **12,** 552–562.

Pani, L., Overdier, D. G., Porcella, A., Qian, X., Lai, E., and Costa, R. H. (1992b). Hepatocyte nuclear factor 3 beta contains two transcriptional activation domains, one of which is novel and conserved with the *Drosophila* forkhead protein. *Mol. Cell Biol.* **12,** 3723–3732.

Qian, X., and Costa, R. H. (1995). Analysis of hepatocyte nuclear factor-3 beta protein domains required for transcriptional activation and nuclear targeting. *Nucleic Acids Res.* **23,** 1184–1191.

Ruiz i Altaba, A., Prezioso, V. R., Darnell, J. E., and Jessell, T. M. (1993). Sequential expression of HNF-3 beta and HNF-3 alpha by embryonic organizing centers: The dorsal lip/node, notochord and floor plate. *Mech. Dev.* **44,** 91–108.

Sambrook, J., Fritsch, E. G., and Maniatis, T. (1989). "Molecular Cloning: A Laboratory Manual." Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sasaki, H., and Hogan, B. L. M. (1994). *HNF-3β* as a regulator of floor plate development. *Cell* **76,** 103–115.

Sasaki, H., and Hogan, B. L. M. (1996). Enhancer analysis of the mouse *HNF-3β* gene: Regulatory elements for node/notochord and floor plate are independent and consist of multiple subelements. *Genes Cells* **1,** 59–72.

Sasaki, H., Hui, C., Nakafuku, M., and Kondoh, H. (1997). A binding site for Gli proteins is essential for *HNF-3β* floor plate enhancer activity in transgenics and can respond to Shh in vitro. *Development* **124,** 1313–22.

Satoh, N. (1994). "Developmental Biology of Ascidians." Cambridge Univ. Press, New York.

Shimauchi, Y., Yasuo, H., and Satoh, N. (1997). Autonomy of ascidian *fork head/HNF-3β* gene expression. *Mech. Dev.* **69,** 143–154.

Simmen, M. W., Leitgeb, S., Clark, V. H., Jones, S. J., and Bird, A. (1998). Gene number in an invertebrate chordate, *Ciona intestinalis*. *Proc. Natl. Acad. Sci. USA* **95,** 4437–4440.

Strahle, U., Blader, P., Henrique, D., and Ingham, P. W. (1993). *Axial*, a zebrafish gene expressed along the developing body axis, shows altered expression in cyclops mutant embryos. *Genes Dev.* **7,** 1436–1446.

Teillet, M.-A., Lapointe, F., and Le Douarin, N. M. (1998). The relationships between notochord and floor plate in vertebrate development revisited. *Proc. Natl. Acad. Sci. USA* **95,** 11733–11738.

Weinstein, D. C., Ruiz i Altaba, A., Chen, W. S., Hoodless, P., Prezioso, V. R., Jessell, T. M., and Darnell, J. E., Jr. (1994). The winged-helix transcription factor *HNF-3β* is required for notochord development in the mouse embryo. *Cell* **78,** 575–588.

Received for publication August 17, 2000
Revised September 25, 2000
Accepted September 25, 2000
Published online December 5, 2000