Regionalization of Sonic hedgehog transcription along the anteroposterior axis of the mouse central nervous system is regulated by Hnf3-dependent and -independent mechanisms

Douglas J. Epstein1,*, Andrew P. McMahon2 and Alexandra L. Joyner1,3,‡

1Developmental Genetics Program and Howard Hughes Medical Institute, Skirball Institute of Biomolecular Medicine, and
2Department of Cellular, Molecular and Developmental Biology, Harvard University, 1600 Divinity Avenue, Cambridge MA, USA
3Departments of Cell Biology and Physiology and Neuroscience, NYU School of Medicine, 540 First Avenue, New York, NY 10016, USA
*Present address: Department of Genetics, University of Pennsylvania Medical School, Philadelphia, PA 19104-6145, USA (epsteind@mail.med.upenn.edu)
‡Author for correspondence (e-mail: joyner@saturn.med.nyu.edu)

Accepted 20 October; published on WWW 14 December 1998

SUMMARY

The axial midline mesoderm and the ventral midline of the neural tube, the floor plate, share the property of being a source of the secreted protein, Sonic hedgehog (Shh), which has the capacity to induce a variety of ventral cell types along the length of the mouse CNS. To gain insight into the mechanisms by which Shh transcription is initiated in these tissues, we set out to identify the cis-acting sequences regulating Shh gene expression. As an approach, we have tested genomic clones encompassing 35 kb of the Shh locus for their ability to direct a lacZ reporter gene to the temporally and spatially restricted confines of the Shh expression domains in transgenic mice. Three enhancers were identified that directed lacZ expression to distinct regions along the anteroposterior axis including the ventral midline of the spinal cord, hindbrain, rostral midbrain and caudal diencephalon, suggesting that multiple transcriptional regulators are required to initiate Shh gene expression within the CNS. In addition, regulatory sequences were also identified that directed reporter expression to the notochord, albeit, under limited circumstances. Sequence analysis of the genomic clones responsible for enhancer activity from a variety of organisms, including mouse, chicken and human, have identified highly conserved binding sites for the hepatocyte nuclear factor 3 (Hnf3) family of transcriptional regulators in some, but not all, of the enhancers. Moreover, the generation of mutations in the Hnf3-binding sites showed their requirement in certain, but not all, aspects of Shh reporter expression. Taken together, our results support the existence of Hnf3-dependent and -independent mechanisms in the direct activation of Shh transcription within the CNS and axial mesoderm.

Key words: Shh, Gene regulation, Central nervous system, Mouse, Hnf3

INTRODUCTION

The establishment of regional identity along the two major axes of the vertebrate nervous system involves the coordinated expression of signaling molecules localized in time and space to discrete organizing centers (Tanabe and Jessell, 1996; Lumsden and Krumlauf, 1996). Two of such embryonic territories, the axial mesoderm (prechordal plate and notochord) underlying the neural plate and the ventral midline of the spinal cord (floor plate) have been defined as organizing centers based on their ability to promote ventral CNS differentiation in tissue recombination assays (reviewed in Tanabe and Jessell, 1996).

The cloning of the secreted factor, Sonic hedgehog (Shh) (Echelard et al., 1993; Riddle et al., 1993; Krauss et al., 1993; Roelink et al., 1994; Chang et al., 1994) and the finding that it is expressed within organizing centers such as the notochord and floor plate prompted experiments to test its ability to function as an inductive signal during nervous system development. Gain- and loss-of-function experiments have revealed a crucial role for Shh in patterning the ventral CNS along the length of the anteroposterior neuraxis (Tanabe and Jessell, 1996). At spinal cord levels, Shh has been shown to be a sufficient and required signal for the induction of motor neuron and floor plate differentiation (Roelink et al., 1995; Marti et al., 1995a; Chiang et al., 1996). At more rostral axial levels, Shh has been shown to be dependent on the cooperation of additional factors, including members of the BMP and Fgf superfamilies, for the induction of distinct neuronal cell types (Dale et al., 1997; Ye et al., 1998).

Considerable attention has focussed on the roles of mediators in Shh signal transduction (reviewed in Hammerschmidt et al., 1997; Ruiz i Altaba, 1997; Kalderon, 1997). Stemming largely from studies of hh signaling in
**Drosophila**, the framework of a pathway is now emerging to explain how on a molecular level these conserved transducers of the Shh signal participate in the patterning of the ventral CNS. Shh, secreted from the notochord, is thought to be bound by its co-receptors Ptc and Smo in the overlying medial neural plate (Stone et al., 1996; Marigo et al., 1996a), presumably leading to the downregulation of PKA activity (Hammerschmidt et al., 1997) and activation of members of a family of transcriptional regulators, the Gli genes (Hui et al., 1994; Lee et al., 1997; Platt et al., 1997). Gli2 has been shown to be required for proper floor plate development (Ding et al., 1998; Matise et al., 1998) and Gli1 can positively regulate the transcription of two target genes in the ventral CNS, the winged helix transcriptional activator Hnf3β and Ptc (Sasaki et al., 1997; Hynes et al., 1997; Lee et al., 1997). Hnf3β, in turn, has been shown sufficient to activate Shh expression in the CNS (Echelard et al., 1993; Hynes et al., 1995a; Ruiz i Altaba et al., 1995) hence, completing a molecular cascade that culminates in the initiation and maintenance of floor plate gene expression. Subsequent to floor plate induction, the maintenance of floor plate gene expression is no longer dependent on Gli activity, as expression of all family members is downregulated from the floor plate (Lee et al., 1997; Platt et al., 1997).

While progress has been made in deciphering the pathway downstream of Shh, little is known about the mechanisms involved in activating Shh gene expression within the aforementioned organizing centers. Regulation of the Shh gene is critical for proper growth and development, as loss of a single copy can result in the human condition holoprosencephaly (Belloni et al., 1996; Roessler et al., 1996). Furthermore, inappropriate activation of Shh (Echelard et al., 1993; Fan et al., 1995; Oro et al., 1997) or components of the pathway (Epstein et al., 1996; Johnson, 1996; Dahmane et al., 1997; Xie et al., 1998) has been implicated in over-proliferation phenotypes including the formation of tumors, thus suggesting that the Shh pathway requires a tightly balanced level of regulation.

Hnf3β has been proposed to be an upstream activator of Shh based initially on the observation that it is expressed prior to and in the same cells as Shh in the axial mesoderm and floor plate (Sasaki and Hogan, 1993; Ang et al., 1993; Monaghan et al., 1993; Echelard et al., 1993; Marti et al., 1995b). That ectopic expression of Hnf3β, in a variety of developmental systems, can lead to the induction of floor plate gene expression including Shh, suggests that Hnf3β may be involved in initiating Shh transcription, however, it has yet to be determined whether this activation is direct (Echelard et al., 1993; Sasaki and Hogan, 1994; Ruiz i Altaba et al., 1995). As mice carrying loss-of-function alleles of Hnf3β result in the loss of a functional node, notochord and floor plate, they have proved uninformative with respect to addressing whether Hnf3β is necessary for the activation of Shh expression (Ang and Rossant, 1994; Weinstein et al., 1994).

Interestingly, another transcriptional component of the Shh pathway, Gli3, has been shown to negatively regulate the expression of Shh (Marigo et al., 1996b). The spontaneously arising extra-toes mouse mutant, which harbors a deletion of Gli3 (Vortkamp et al., 1992; Hui and Joyner, 1993), shows ectopic expression of Shh in both the anterior limb bud (Masuya et al., 1995; Buscher et al., 1997) and dorsal CNS

(Ruiz i Altaba, 1998), two sites where Shh is not normally found.

In an attempt to further characterize the transcriptional regulators that act upstream of mouse Shh in polarizing centers such as the floor plate and notochord, we initiated an in vivo reporter assay to identify the cis-acting sequences that regulate Shh gene expression. Several enhancers were detected that can direct the expression of a lacZ reporter gene to discrete regions of the ventral CNS in transgenic mice, suggesting that activation of Shh transcription is regionalized along the anteroposterior axis and that not one, but multiple, regulators are responsible for initiating Shh expression in the CNS. Consensus binding sites for Hnf3 class transcriptional regulators were identified in some, but not all, enhancers supporting the view that Hnf3-dependent and -independent mechanisms exist to activate Shh gene expression in the floor plate.

### MATERIALS AND METHODS

**Generation of Reporter Constructs**

Genomic clones overlapping the Shh locus were isolated from a λ Dash II library (Stratagene) as described in Echelard et al. (1993). Reporter constructs were generated by subcloning DNA restriction fragments from the purified λ clones into a reporter cassette modified from that previously described (Echelard et al., 1993). The reporter vector contained either a Shh (1.1 kb EcoRI-SacII) or hsp68 (Kothary et al., 1989) promoter, a lacZ DNA followed by SV40 polyadenylation sequences, and a multiple cloning site polylinker located either upstream or downstream of the lacZ cDNA. As the 1.1 kb EcoRI-SacII fragment proved to be toxic to bacteria when propagated in a high replication copy plasmid, the low copy pBR322 vector was used in the cloning of all reporter constructs. Genomic fragments were blunt ended with T4 DNA polymerase (NEB) and cloned into the SmalI restriction site of the reporter vector, situated either upstream or downstream of the promoter depending on its native orientation. The restriction sites listed on reporter constructs 1 through 19 (Figs 1, 3, 5) correspond to those used in subcloning.

Mutations in Hnf3-binding sites 4 and 5 (RC18 and RC19) were generated using a multipurpose PCR-based approach. For Hnf3-binding site 4, two sets of primers (J23-5′ GGCAGAGCCCCGGTCG-CATTCC 3′, J24-5′ TTTAACGCTTCCAGAAAGATCTTCTGC 3′, and J25-5′ TTTAACGCTTCCAGAAAGATCTTCTGC 3′, and J26-5′ TTTAACGCTTCCAGAAAGATCTTCTGC 3′) were used to amplify sequences on either side of the Hnf3 site, that were subsequently digested with Smal-HindIII and HindIII-XhoI, respectively, and cloned into the corresponding sites within the pBSKII (Stratagene) vector. This strategy effectively replaced the Hnf3 site 4 with a HindIII restriction site. The 350 Smal-XhoI fragment was then cloned into a reporter vector containing the 150 bp (EcoRI-Smal) fragment (RC18). Two point mutations that abolish the ability of Hnf3 site 5 to bind to DNA were also generated during the aforementioned mutagenesis of Hnf3-binding site 4 through the introduction of mismatches within the J26 primer (underlined bases). To generate RC19, RC18 was digested with BamHI/XhoI and the 500 bp fragment containing the mutated Hnf3 sites 4 and 5 was cloned into the corresponding sites of RC17.

The sequencing of all PCR-generated mutations, as well as the genomic fragments contained within RC1 (GenBank Accession Number AF098925), RC6 (GenBank Accession Number AF098926), RC14 (GenBank Accession Number AF098927) and RC15 was performed by dye terminator cycle sequencing (PE Applied Biosystems) using an ABI Prism 377 DNA sequencer.
Primer extension analysis
The start site of Shh transcription was mapped by primer extension. For this, an oligo complementary to nucleotides −211 to −230 (5′ CGACGGCTACGGTCCAGCGGC 3′) of the Shh cDNA was end labeled with T4 polynucleotide kinase (NEB) and annealed overnight (4 M NaCl, 100 mM Pipes pH 6.4) at 56°C to 3 μg of poly(A)+ RNA isolated from the lungs of 18.5 dpc embryos. As negative controls, both tRNA and poly(A)+ E18.5 liver RNA were used as primer extension templates. The annealed primer was extended using Superscript II reverse transcriptase (GIBCO BRL) under conditions recommended by the supplier. After phenol/chloroform extractions and subsequent precipitation, the primer extension products were electrophoresed on an 8% sequencing gel along side a sequencing reaction generated with the identical primer and a genomic clone overlapping the region. The gel was dried and exposed to autoradiography film (Kodak) for 7 days at −80°C using an intensifying screen.

Production and genotyping of transgenic mice
Outbred Swiss webster mice (Taconic) were used to produce transgenic embryos and mouse lines essentially as described (Hogan et al., 1994). Transgenes were prepared for microinjection as described (Epstein et al., 1996). The genotyping of transgenic embryos and mice was carried out by PCR using proteinase K-digested yolk sacs or tail biopsies as DNA templates. Upstream primers directed against either the Shh (5′ GACGGCGGGGGACA-GCTCAC 3′) or hsp68 (5′ GACGAACTTCCCCAGGACCATC 3′) promoters and a downstream primer directed against lacZ (5′ AAGGGCGATCGTGGCGGCG 3′) generating DNA fragments of (200 bp) and (220 bp), respectively, were used under the following PCR conditions: 94°C for 1 minute, 60°C for 1 minute, 72°C for 1 minute, for 30 rounds followed by a final extension at 72°C for 10 minutes. For staging embryos, the day of vaginal plug detection corresponded to 0.5 days post coitus (dpc).

Whole-mount β-galactosidase and in situ hybridization
β-galactosidase activity was detected in whole-mount embryos by using X-gal (Sigma) or Salmon-gal (Biosynth) as substrates according to Echelard et al. (1994). The embryos were stained from 20 minutes to overnight depending on the strength of transgene expression. Whole-mount RNA in situ hybridization of embryos was performed essentially as described (Matise et al., 1998) using an Shh riboprobe (Echelard et al., 1993).

Histological analysis
After whole-mount staining some embryos were sectioned at 50-75 μm on a vibrotome and counterstained with Nuclear Fast Red (Poly Sciences).

RESULTS
As an impetus towards the identification of the critical cis-acting sequences responsible for the regulation of Shh gene expression, three overlapping λ phage clones encompassing 35 kb of the Shh locus were isolated and a high resolution restriction map constructed (Fig. 1). Based on this map, genomic fragments surrounding the Shh gene were cloned into a reporter vector comprising of a minimal promoter, the bacterial β-galactosidase gene and an SV40 polyadenylation signal. All genomic fragments were initially cloned in the orientation found within the endogenous Shh locus, thus preserving its normal position with respect to the Shh promoter. The reporter constructs were then injected into fertilized mouse eggs and assayed for X-gal staining at either 9.5 or 10.5 dpc. These stages of development were chosen because Shh expression in tissues with polarizing activity is well developed. The pattern of Shh expression during mid-embryogenesis in the mouse has been documented (Echelard et al., 1993). Briefly, at 9.5 dpc, Shh mRNA can be detected throughout much of the ventral neuraxis including the floor plate of the spinal cord, hindbrain and midbrain. In the ventral midline of the midbrain and caudal diencephalon Shh expression is broader than in other regions whereas in the rostral diencephalon Shh is absent from the midline but found in two stripes of lateral expression that merge in the floor of the telencephalon (Echelard et al., 1993) and Fig. 2A). Expression of Shh can also be observed in tissues of non-ectodermal origin including the notochord, gut and pharyngeal endoderm and, at 10.5 dpc, the posterior region of the limb bud (Echelard et al., 1993). Although the analysis focussed at identifying regulatory sequences mediating Shh-like expression at these stages, relevant reporter expression was also assessed in stable mouse lines at earlier and later time points.

As gene regulation can be conferred by promoter-specific interactions with tissue-specific enhancers (Hansen and Tjian, 1995; Merli et al., 1996), all reporter constructs were initially tested utilizing a Shh promoter. The Shh transcriptional start site was mapped by primer extension analysis to a position 352 bp upstream of the initiator methionine where a consensus TATA box (TATAAT)-binding site was identified (data not shown). A Shh promoter comprising 1 kb of upstream sequence was cloned into a lacZ reporter construct [Fig. 1, Reporter Construct 1 (RC1)] and tested for its ability to recapitulate the pattern of Shh gene expression in transgenic mouse embryos. Although four of eight embryos carrying the RC1 transgene expressed lacZ at 10.5 dpc, suggesting that the promoter was functional, none of them displayed staining in a pattern consistent with that of Shh. Given this finding, the relevance of the previous report that maintenance of Shh expression in cultured limb bud cells is dependent on multiple consensus binding sites for Hoxd-12 situated in proximity of the Shh promoter is questionable (Knezevic et al., 1997). However, the possibility remains that binding sites for Hoxd-12 located elsewhere are still pertinent for Shh regulation.

An upstream Shh enhancer directs expression to the floor plate
Extending the enhancer search to include an additional 13 kb of genomic DNA upstream of the start of Shh transcription (Fig. 1, RC2) revealed lacZ expression exclusively in the ventral midline of the neural tube in transgenic embryos (Fig. 2C). In contrast to the pattern of Shh expression normally seen at 9.5 dpc, embryos carrying the RC2 transgene showed X-gal staining in a limited pattern along the anteroposterior axis (compare Fig. 1A and C). The rostral expression boundary of the reporter gene was detected at the mid-hindbrain junction and the caudal limit extended only to the hindlimbs. Vibratome sections cut in the transverse plane through several of these embryos also demonstrated that expression was restricted to the floor plate of the spinal cord and hindbrain and absent from the notochord and gut (compare Fig. 2A with C and Fig. 2B with D). Founder lines of mice were generated with the RC2 transgene and, other than the floor plate, no additional sites of reporter activity were revealed in embryos between 8.5 and 13.5 dpc.
In addition to assessing spatial control of the RC2 transgene, we also investigated whether temporal regulation was consistent with the onset of \textit{Shh} transcription in the floor plate. \textit{Shh} expression in the CNS is first detected at the 8-somite stage in the presumptive ventral midbrain and then progresses rostrally into the forebrain and, by the 12-somite stage, caudally into the floor plate of the hindbrain and spinal cord (Echelard et al., 1993). Interestingly, the onset of RC2 activity at the 12-somite stage coincided precisely with the initiation of \textit{Shh} expression in the floor plate of the hindbrain and spinal cord (data not shown).

A series of deletion constructs (Fig. 1, RC3-RC6) aided in better defining the location of the \textit{Shh} floor plate enhancer (SFPE1) to within a 1.1 kb genomic fragment approximately 8 kb upstream of the \textit{Shh} transcription start site. This 1.1 kb construct (RC6) displayed similar reporter activity compared to the 14 kb construct (RC2), as assessed by the consistency of expression as well as the minimal time required to initiate staining (20 minutes). Furthermore, the 1.1 kb fragment was found to possess the same activity when directing transcription from a heterologous (\textit{hsp68}) promoter, independent of its orientation, thus, fitting the criteria of a bone fide enhancer (data not shown).

**Intronic enhancers direct reporter gene expression to cranial and spinal cord regions**

Given that upstream sequences directed \textit{lacZ} to only a portion of the \textit{Shh} expression domain, sequences downstream of the start of \textit{Shh} transcription were also tested for reporter activity. A second enhancer complex was identified when constructs containing intronic sequences were used to generate transgenic embryos (Fig. 1, RC7-9). Once again, \textit{X}-gal staining was observed in a restricted pattern in comparison to the \textit{Shh} mRNA expression profile. Embryos carrying a transgene that contained all of intron 1 and a portion of intron 2 (RC7) displayed reporter gene activity in the ventral midline of the rostral midbrain and caudal diencephalon at 9.5 dpc (Fig. 2E). The location of this \textit{Shh} brain enhancer (SBE1) was further delineated to within a 2.2 kb segment within the second intron (Fig. 1, RC8, RC9). Interestingly, both RC8 and RC9 displayed...
Regulation of Shh transcription

Additional expression in the hindbrain and spinal cord, identifying a second Shh floor plate enhancer element (SFPE2; Fig. 2F and data not shown).

Stable lines of transgenic mice generated from RC8 showed X-gal staining at 10.5 dpc, within the zona limitans intrathalamica (ZLI), the boundary between the dorsal and ventral thalami, in addition to the midbrain and caudal diencephalon (Fig. 2G,H). 8-somite-stage embryos carrying the RC8 transgene showed staining in the presumptive ventral midbrain, hence marking the initial site of Shh expression in the CNS (compare Fig. 2JL). The RC8-mediated expression progressed into the hindbrain and spinal cord in similar fashion to Shh but never advanced further anterior than the caudal diencephalon or more posterior than the hindlimbs. An additional site of expression was also detected, albeit transiently, in the axial mesoderm of embryos at 8.5 dpc (Fig. 2J), however, by 9.5 dpc, there was no noticeable staining in the notochord. The two intronic enhancers, SBE1 and SFPE2, were also found to function in an orientation- and promoter-independent fashion (data not shown).

From this reporter gene analysis, we conclude that activation of Shh gene expression along the anteroposterior axis of the mouse CNS is under the transcriptional control of several discrete enhancer elements scattered throughout the Shh locus.

SFPE1 activity is regulated independent of Hnf3 function

To further explore the nature of the cis-acting sequences within RC6 that mediate floor plate expression, the 1.1 kb genomic fragment was sequenced and analyzed for the presence of consensus binding sites recognized by known transcription factors using the Signal Scan (Advanced Biosciences Computing Center, University of Minnesota) and Best Fit (GCG) algorithms. As expected, a number of general transcription factor binding sites were identified; however, attention was only directed at those suggestive of tissue specificity. Of particular interest was the identification of three 12 nucleotide stretches that matched the recognition site of the hepatocyte nuclear factor 3 (Hnf3) class of winged helix transcriptional regulators (Overdier et al., 1994; see Table 1).

Presently, three Hnf3 family members (α, β and γ) have been identified, all of which bind to similar target sequences (reviewed in Kaufmann and Knochel, 1996).

To determine whether Hnf3 proteins were able to bind to the identified Hnf3 recognition sequences, electromobility shift assays using liver extracts as the source of all three Hnf3 proteins were performed. Each of the three Hnf3 isoforms was shown to be capable of binding to the three identified Hnf3-binding sites (data not shown). Moreover, the three Hnf3-binding sites were also found to be 100% conserved with human sequences when a best fit alignment (GCG) of the 1.1 kb (RC6) mouse sequence was performed with the sequence of a 30 kb cosmid overlapping the human Shh locus (Table 1). The preservation of the sequence of the Hnf3-binding sites between mouse and human is suggestive of functional conservation.

In order to determine whether the 5′ segment of the 1.1 kb
(RC6) fragment, encompassing the cluster of Hnf3 sites, was sufficient for floor plate expression, embryos carrying a transgene containing only the proximal half of RC6 were generated (Fig. 3, RC10). Remarkably, strong X-gal staining was detected in the notochord, whereas floor plate staining was found to be relatively weak (Fig. 4A). Of the nine transgenic embryos carrying RC10, only two showed this pattern with the remainder being devoid of any Shh-like expression. This suggests that a Shh notochord enhancer (SNE) is present in RC10 and is particularly sensitive to the site of integration.

Given the unexpected finding that the proximal portion of the 1.1 kb fragment (RC10) is capable of directing reporter expression to the notochord, whereas the full fragment (RC6) is sufficient for only floor plate expression, we set out to determine to what degree this dual enhancer activity was dependent on Hnf3 function. For this, a series of constructs containing small deletions resulting in the sequential removal of the Hnf3-binding sites were tested in the reporter assay. The first construct generated contained an 80 bp deletion of RC6, overlapping Hnf3-binding site 1 (Fig. 3, RC11). 10.5 dpc embryos carrying the RC11 transgene displayed strong reporter activity in the floor plate and consistently patchy X-gal staining in the notochord (Fig. 4B). When embryos carrying a construct that removed the first and second Hnf3-binding sites of RC6 were assessed for reporter expression, X-gal staining was observed in the floor plate of all four transgenic animals generated, while patchy notochord expression was only detected in one embryo (Fig. 3, RC12). Embryos carrying a construct that deleted all three Hnf3-binding sites (Fig. 3, RC13) showed strong staining in the floor plate but were completely devoid of reporter activity in the notochord (Fig. 4C).

As no other sequences within the distal portion of RC13 were found to match the consensus for Hnf3-binding proteins as defined by Overdier et al. (1994), we conclude that SFPE1 activity occurs in an Hnf3-independent manner. Moreover, SNE activity may be dependent on Hnf3 function as the loss

**Fig. 3.** Dissection of the upstream Shh notochord and floor plate enhancers. The schematic map of the Shh locus is as described in Fig. 1, with the addition of a red balloon depicting the position of a Shh notochord enhancer (SNE). The 1.1 kb PstI-EcoRI genomic fragment (RC6) is enhanced to reveal the position of additional restriction enzyme sites (Bc, BclI; E, EcoRI; Nc, Ncol; P, PstI; Ss, SspI; SspBI) used to generate RC10-13. The position of Hnf3-binding sites 1-3 (magenta bars) is also shown. RC6 (from Fig. 1) serves as a reference point for comparison with RC 10-13. Sites of expression (fp, floor plate; nt, notochord).

**Fig. 4.** Shh notochord and floor plate enhancers are independently regulated. (A) RC10 expression in the notochord (arrow) showing robust expression compared to the floor plate. The section is from the lumbar region of a 9.5 dpc embryo stained for 60 minutes. After staining, the notochord was dissected away from other tissues to reveal extent of staining along a portion of its length (below). (B) RC11 expression in the notochord (arrow) showing weak and patchy (below) expression along its length compared to the strong floor plate staining. The section is from the thoracic region of a 9.5 dpc embryo stained for 3 hours. (C) RC 13 expression is absent from the notochord (arrow and below) yet persists in the floor plate. The section is from lumbar regions of a 9.5 dpc embryo stained for 3 hours.
Table 1. Sequence of mouse Hnf3-binding sites identified in Shh regulatory regions compared to human and chicken

<table>
<thead>
<tr>
<th>Hnf3 binding site</th>
<th>Mouse sequence</th>
<th>Human sequence</th>
<th>Chicken sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (RC10)</td>
<td>AAATATTAGTC</td>
<td>AAATATTAGTC</td>
<td>NA</td>
</tr>
<tr>
<td>2 (RC10)</td>
<td>CCTTATTTGATT</td>
<td>CCTTATTTGATT</td>
<td>NA</td>
</tr>
<tr>
<td>3 (RC10)</td>
<td>ACATATTATAC</td>
<td>ACATATTATAC</td>
<td>NA</td>
</tr>
<tr>
<td>4 (RC16)</td>
<td>CAATCTCTATT</td>
<td>NA</td>
<td>AGATCTCTGTATT</td>
</tr>
<tr>
<td>5 (RC16)</td>
<td>AGATGTGATT</td>
<td>NA</td>
<td>AGATGTGATT</td>
</tr>
<tr>
<td>6 (RC17)</td>
<td>GAATGTTGTC</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>7 (RC17)</td>
<td>GGATATTATCT</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>8 (RC17)</td>
<td>GGGTGGTGGCAT</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

HNF3 consensus binding site: VAWTRRTKRYTY

The Hnf3-binding sites were identified according to their sequence matching that of the consensus site for Hnf3β (Overdier et al., 1994). Best-fit alignment (GCG) determined the overlap between mouse and human as well as mouse and chicken sequences. Underlined bases in Hnf3-binding site 4 show differences between the mouse and corresponding chicken sequence yet, still match the consensus. Solid line refers to the fact that no homology was detected between the mouse and chicken sequence at the corresponding positions. NA, sequence not available for comparison. Mismatches between certain bases and the Hnf3 consensus were occasionally detected, however binding to the site persisted (Overdier et al., 1994; data not shown). V: A,C or G; W: A or T; R: G or A; K: G or T; Y: C or T.

Fig. 5. Dissection of the SBE1 and SFPE2 intronic enhancers. The schematic map of the Shh locus is as described in Fig. 1. The 1.3 kb EcoRI-BglII genomic fragment is enhanced and shows the position of Hnf3-binding sites 4-8 as well as the restriction enzymes used in the generation of RC14-19. RC 15 contains chicken genomic DNA isolated from a similar region of intron 2 compared to RC 14. Hnf3-binding sites 4 and 5 are mutated in RC 18 and 19 (magenta X). RC18 expression is absent from the ventral midline of the CNS in transgenic embryos but shows expression in lateral regions of the rostral midbrain and caudal diencephalon (lat rm, lat cd). Similarly, 2/5 embryos transgenic for RC19 showed absence of ventral midline expression in the CNS (lat rm, lat cd). In contrast, 3/5 embryos transgenic for RC19 also showed either very weak expression (2/5) or normal expression (1/5) in the ventral midline of the CNS (fp, rm, cd, zli).
of binding sites correlated with the loss of reporter activity in the notochord. This conclusion is not definitive as the possibility remains that, in addition to the Hnf3-binding sites, we have removed additional sequences required for SNE activity.

**SBE1 and SFPE2 rely on Hnf3 function for midline but not lateral Shh expression**

In an effort to delineate the region containing the critical sequences for SBE1 and SFPE2 activity, a 1.3 kb genomic fragment (RC14) was cloned downstream of the hsp68 lacZ reporter cassette and injected into fertilized zygotes (Fig. 5). At 9.5 dpc, transgenic embryos were found to stain in the same pattern as those carrying the RC8 transgene including the ventral midline of the spinal cord, midbrain and caudal diencephalon (Fig. 6A).

Shh has been cloned from most experimental organisms and shows a pattern of expression highly conserved across phyla (Hammerschmidt et al., 1997). The pattern of Shh expression within the CNS and notochord in the chick is similar to that in the mouse. To benefit from this high degree of conservation of expression, we generated a reporter construct containing DNA overlapping a similar region of intron 2 from the chicken Shh locus (Fig. 5, RC15). Interestingly, embryos carrying the chicken DNA construct showed X-gal staining in a similar pattern to those carrying the mouse DNA (compare Fig. 6A with B). The one difference in the staining of embryos carrying chicken versus mouse constructs related to the posterior extent of expression within the CNS. Embryos carrying the RC14 mouse DNA construct showed a posterior limit of lacZ expression at the level of the hindlimbs whereas embryos carrying the RC15 chicken DNA construct showed staining into the tail bud.

Relying on the premise that conservation of sequence underlies conservation of function, the 1.3 kb (RC14) intron fragment containing the cis-acting sequences required for SBE1 and SFPE2 activity was sequenced from mouse and compared to the comparable region of intron 2 from chicken. The overall sequence identity between the mouse and chicken fragments approached 70% including two conserved Hnf3-binding sites (Table 1) that were found capable of binding Hnf3 proteins (data not shown). Three other Hnf3-binding sites matching the consensus were also identified in the mouse DNA fragment but were not found to be conserved in the corresponding chicken DNA (Table 1).

Before determining the requirement of the Hnf3-binding sites for SBE1 and SFPE2 activity, the positions of the two enhancers were better delineated. The XhoI restriction site appeared to demarcate the boundary between SBE1 and SFPE2 given the results obtained with RC7 and RC9 (Fig. 1). Two constructs were therefore generated, RC16 and RC17 (Fig. 5), which contained sequences proximal and distal to the XhoI restriction site, respectively. Although RC16 conveyed SBE1 activity, RC17 showed no Shh-like expression, suggesting that SFPE2 activity depends on sequences both proximal and distal to the XhoI site and that on their own, Hnf3-binding sites 6-8 are not sufficient for floor plate expression (see Discussion).

Hnf3-binding sites 4 and 5 were then mutated in the context of the 532 bp fragment that mediates SBE1 expression (Fig. 5, RC18). Transgenic embryos carrying RC18 displayed X-gal staining that was excluded from the ventral midline of all embryos tested (n=4) but was found present in lateral cranial regions (compare Fig. 6C,E with D,F) including, at 10.5 dpc, the zli (not shown).

To determine whether SFPE2 activity was also dependent on Hnf3-binding sites 4 and 5, they were perturbed in the context of the 1.3 kb fragment (Fig. 5, RC19). Two of five embryos carrying the RC19 transgene showed a complete absence of X-gal staining in the ventral midline of the spinal cord, midbrain and caudal diencephalon (compare Fig. 6G with H). The staining observed lateral to the floor plate at spinal cord levels in Fig. 6H is derived from the hsp68 promoter which has endogenous β-gal activity and thus, serves as a positive control for staining (Song et al., 1996). In contrast, staining in lateral cranial regions, where Shh is normally found broader than in more posterior regions, was unaffected by the loss of Hnf3 sites 4 and 5 in these embryos. In the remaining three embryos carrying the RC19 transgene, expression in the midline was weak and patchy (2/5), or unaffected by the loss of the Hnf3 sites (1/5) (data not shown).

Taken together, these results suggest that Hnf3 function is required for consistent floor plate activation by SBE1 and SFPE2 but that an Hnf3-independent pathway exists to activate lateral SBE1 activity.

**DISCUSSION**

**CNS expression of Shh is regulated by multiple enhancers**

With the intent of identifying the transcriptional regulators that act upstream of Shh, we conducted a reporter gene assay in transgenic mice that assessed the ability of 35 kb of genomic DNA surrounding the locus to direct reporter gene expression in a Shh-like pattern. From this assay, several enhancers were identified and shown to direct transcription to discrete regions of the Shh expression domain. Within the CNS, where Shh is expressed along the length of the ventral neuraxis, two independent enhancers, Shh floor plate enhancer 1 and 2 (SFPE1 and SFPE2), were shown to direct reporter expression exclusively to the floor plate of the hindbrain and spinal cord, whereas a third, Shh brain enhancer 1 (SBE1), directed reporter expression to the ventral midbrain and caudal diencephalon. With respect to the genomic organization of the enhancers, SFPE1 was localized to a 500 bp fragment approximately 8 kb upstream of the Shh transcription start site, whereas, SFPE2 and SBE1 were identified on overlapping 1.3 kb fragments situated within intron 2. Transcriptional activity from these three enhancers, as judged by X-gal staining, covered the majority of the anteroposterior neuraxis with the exception of the rostral forebrain and the region caudal to the hindlimbs, which presumably fall under the control of additional enhancers lying outside of the genomic region analyzed.

Regarding expression in other tissues, the proximal half of RC6 was shown to occasionally direct expression to the notochord and RC8 transiently directed expression to the axial mesoderm at early somite stages. Due to the general inconsistency in the reporter gene activity at these sites of expression, it is likely that additional enhancers must exist that control the full extent of Shh expression in the axial mesoderm. Further support for this view emanates from our finding that a 100 kb BAC clone, overlapping the 35 kb of DNA tested in our
Regulation of Shh transcription

Aspects of Shh transcription are dependent on Hnf3 and cooperating factors

A common feature underlying the regulation of SBE1, SNE and SFPE2 activity is the requirement of Hnf3-binding sites contained within their critical cis-acting sequences. Perturbation of the Hnf3 sites in each of these three enhancer constructs alters reporter gene activity, thus confirming that Hnf3 transcriptional regulators are required for aspects of Shh gene expression in the floor plate and notochord. Of the three Hnf3 family members, Hnf3α and Hnf3β have been implicated in Shh signaling based on their overlapping expression with Shh, as well as the ability of Hnf3β to positively regulate Shh transcription in ectopic expression assays (Echelard et al., 1993; Sasaki and Hogan, 1994; Hynes et al., 1995a; Ruiz i Altaba et al., 1995). Relying on expression data alone has been difficult in assessing whether Hnf3α and/or Hnf3β is the activator of Shh transcription in the CNS. Interestingly, Hnf3β precedes Shh expression in the floor plate by approximately 12 hours (Echelard et al., 1993), while Hnf3α is expressed at approximately the same time as Shh in both the floor plate of

...
the ventral midbrain and spinal cord (Ang et al., 1993); (D. J. E. and A. L. J., unpublished observations). Why such a delay exists between the onset of Hnf3β expression and activation of Shh transcription could be explained by the presence of an inhibitory signal in the floor plate (Arkell and Beddington, 1997). Ectopic expression of BMP family members in the ventral midline of the CNS inhibits transcription of Shh even in the presence of Hnf3β (Arkell and Beddington, 1997; McMahon et al., 1998). Interestingly, BMP7 expression is normally found in the floor plate at cranial levels, at a time prior to the initiation of Shh transcription and its downregulation corresponds with the time when Shh is first detected (Arkell and Beddington, 1997). Whether this inhibition by BMP7 acts directly on Hnf3β, a cooperating transcriptional partner required for Shh activation (see below) or through the activation of an inhibitor is unclear.

Genetic data, on the other hand, do not support a required role for Hnf3α in the transcription of Shh, as mice carrying a targeted disruption of the Hnf3α gene still possess a floor plate and maintain Shh expression (Duncan et al., 1998) (Klaus Kaestner, personal communication). Moreover, Hnf3α appears to be a less potent transcriptional activator than Hnf3β, at least for targets of endodermal origin (Duncan et al., 1998). The targeted disruption of Hnf3β in mice results in the lack of an organized node and notochord, and the mice do not possess a floor plate (Ang and Rossant, 1994; Weinstein et al., 1994). As a result, these mice have not proved informative with respect to whether Hnf3β is required for Shh expression, because the tissues in which it is expressed are not specified. This reasoning holds true for two other genes, nodal and goosecoid, which have been shown to genetically interact with HNF3B (Varlet et al., 1997; Filosa et al., 1997). Mice homozygous for a mutation in goosecoid and heterozygous for a mutation in Hnf3β, as well as mice doubly heterozygous for mutations in nodal and Hnf3β, each show varying degrees of loss of the ventral midline CNS which result, in part, from a primary defect in the functioning of the node and, as a secondary consequence, the loss of cell types that express Shh. Only upon the generation of either a conditional Hnf3β mutant or chimeric mice that are wild type for Hnf3β expression in the node and notochord and mutant in the floor plate, will the requirement of Hnf3β for floor plate expression of Shh be fully determined.

Hnf3 family members have been shown to promote the expression of endoderm-derived genes either by activating transcription directly (Kaufmann and Knochel, 1996) and/or by positioning nucleosomes into transcriptionally stable states (Shim et al., 1998). Although we have not determined by which of these mechanisms Hnf3 is mediating SBE1 and SFPE2 activity, it is important to note that the presence of the Hnf3 sites alone is not sufficient for the transcription of the reporter transgenes in all tissues expressing Hnf3 family members. For instance, reporter constructs containing SBE1 and SFPE2 activity never displayed expression in the gut, where Shh and Hnf3β overlap (Echelard et al., 1993), even though Hnf3-binding sites were identified in these constructs. Although it is conceivable that Hnf3 family members are not involved in the regulation of Shh gut expression, another possible explanation for this observation is that Hnf3 requires cooperation with additional factors to regulate Shh expression and, therefore, only when DNA-binding sites for all the obligatory components are present is Shh transcribed in a particular tissue.

This postulation is further supported by the finding that RC16, which possesses SBE1 activity, does not activate reporter expression in the floor plate of the spinal cord or elsewhere in the embryo, even though it contains two required Hnf3-binding sites. Only when more 3' sequences were added (RC14) was floor plate expression in the spinal cord detected. Although these distal sequences also contained Hnf3 sites, when tested on their own (RC17), transgenic embryos did not reveal floor plate expression (Fig. 5). From this, we conclude that Hnf3-binding sites alone are not sufficient for floor plate expression in the spinal cord.

This conclusion is further supported by experiments that deleted selected Hnf3-binding sites. In the absence of Hnf3-binding sites 4 and 5 (in the context of RC19), lacZ expression in the floor plate of the spinal cord was dramatically reduced. In some transgenic embryos however, floor plate expression persisted, perhaps suggesting that, in the absence of functional Hnf3-binding sites, the site of transgene integration was important for consistent reporter activity. That RC19 expression was dependent on its site of integration could be indicative of its sensitivity to chromatin structure, with expression being revealed only when the transgene integrates within a transcriptionally permissive chromatin environment. We would predict then, that the role of Hnf3 in mediating SFPE2 activity is to ensure a transcriptionally suitable chromatin organization, permitting the binding and activation of Shh transcription by cooperating factors. In support of this, the crystal structure of Hnf3 has been shown to share considerable similarity with histone H5, which plays a fundamental role in chromatin organization (Clark et al., 1993). Moreover, similar conclusions have been proposed with respect to the regulation of albumin gene expression by Hnf3 (Shim et al., 1998).

The zebrafish Shh promoter has recently been characterized in co-transfection experiments using HeLa cells (Chang et al., 1997). Although the relevance of HeLa cells to the various tissues that express Shh is not clear, results from these assays suggested that axial, a closely related Hnf3β homologue in zebrafish and frogs, and retinoic acid receptors are direct regulators of Shh (Chang et al., 1997). In comparing the sequence surrounding the zebrafish and mouse Shh promoters, we identified a stretch of 73 nucleotides that show a sequence identity of 73%, including a perfectly conserved TATA box. Of particular interest was the identification of at least one Hnf3-binding site showing 100% sequence identity to that of mouse. Although this particular binding site was shown to be required for reporter activity in HeLa cells, the equivalent segment of DNA in mouse yielded no activity in vivo (Fig. 1, RC1). This finding suggests that not all Hnf3-binding sites are equivalent, and/or that there are different requirements for the activation of Shh in vitro as compared to in vivo.

**Shh can be activated by Hnf3-independent mechanisms**

We have shown that Hnf3 is necessary for consistent SFPE2 activity and the ventral most activity of SBE1, but that ventral-lateral expression found in the rostral midbrain, diencephalon and ZLI, where Shh expression is broader than the ventral midline, is independent of Hnf3 function. In addition, SFPE1 activity within RC13 also appeared to be independent of Hnf3 function, as no binding sites matching the
Hnf3 consensus were identified within the 500 bp sequence. From this, it is apparent that additional DNA-binding sites within the various enhancers are also required for activating Shh gene expression. A powerful strategy used for the identification of critical regulatory sequences relies upon the comparison of sequences regulating gene expression in divergent organisms (Aparicio et al., 1995; Manzanares et al., 1997). To this end, we have compared the sequence of mouse RC16 that contains SBE1 activity, with a similar fragment from the chicken Shh locus, as well as the mouse RC13 sequence (SFP1) with the comparable human sequences. In both instances, a high degree of sequence conservation was identified, 70% and 75%, respectively. With respect to RC16, several perfectly conserved binding sites matching the consensus of homeodomain and octamer binding proteins were identified. Comparison of RC13 sequences showed several conserved E-box-binding sites that serve as targets for bHLH transcriptional regulators. Future studies will be required to address the significance of these and other potentially critical cis-acting regulatory sequences directing Shh gene expression.

We would like to thank Randy Johnson for providing us with the chicken Shh genomic clone and Stephen Duncan for providing us with liver nuclear extracts. The assistance of the staff at the Skirball Institute Transgenic Core Facility was much appreciated. We are grateful to Brian Bai and Alex Schier for critically reading the manuscript. D. J. E. was supported by a Centennial Fellowship from the Howard Hughes Medical Institute. Work in A. P. M.’s laboratory was supported by a grant (NS 33642) from the NINDS. Work in A. L. J.’s laboratory was supported by a grant (R01 HD35768) from the NICHD.


