Neurogenin2 Expression in Ventral and Dorsal Spinal Neural Tube Progenitor Cells Is Regulated by Distinct Enhancers

Andrew D. Simmons, Sharon Horton, Andrew L. Abney, and Jane E. Johnson

Center for Basic Neuroscience, University of Texas Southwestern Medical Center, Dallas, Texas 75390-9111

The basic helix-loop-helix transcription factor Neurogenin2 (NGN2) is expressed in distinct populations of neural progenitor cells within the developing central and peripheral nervous systems. Transgenic mice containing ngn2/lacZ reporter constructs were used to study the regulation of ngn2 in the developing spinal cord. ngn2/lacZ transgenic embryos containing sequence found 5' or 3' to the ngn2 coding region express lacZ in domains that reflect the spatial and temporal expression profile of endogenous ngn2. A 4.4-kb fragment 5' of ngn2 was sufficient to drive lacZ expression in the ventral neural tube, whereas a 1.0-kb fragment located 3' of ngn2 directed expression to both dorsal and ventral domains. Persistent β-gal activity revealed that the NGN2 progenitor cells in the dorsal domain give rise to a subset of interneurons that send their axons to the floor plate, and the NGN2 progenitors in the ventral domain give rise to a subset of motor neurons. We identified a discrete element that is required for the activity of the ngn2 enhancer specifically in the ventral neural tube. Thus, separable regulatory elements that direct ngn2 expression to distinct neural progenitor populations have been defined.

Key Words: ngn2; Math4A; bHLH genes; transcription factors; spinal cord development; gene regulation; transgenic mice; motor neuron progenitors; interneuron progenitors.

INTRODUCTION

Members of the basic helix-loop-helix (bHLH) family of transcription factors are temporally and spatially expressed in precise patterns during neural development and have been shown to function in the determination and differentiation of multiple neuronal lineages (Guillemot, 1999; Hassan and Bellen, 2000). In Drosophila, the proneural genes of the achaete-scute (ac-sc) complex and atonal (ato) encode bHLH transcription factors that establish cell fate during external sensory and chordotonal organ formation, respectively (Hassan and Vaessin, 1996; Jan and Jan, 1994). Numerous vertebrate homologs of the Drosophila proneural genes have been identified (reviewed in Hassan and Bellen, 2000) and essential roles in neural development for many of these genes have been demonstrated using mouse mutants (Ben-Arie et al., 1997; Bermingham et al., 1999; Fode et al., 1998; Guillemot et al., 1993; Ma et al., 1998; Miyata et al., 1999; Naya et al., 1997). A subset of the vertebrate neural bHLH factors is expressed in restricted progenitor populations, and as these progenitors exit the cell cycle and differentiate, bHLH expression is extinguished (Lee, 1997). These “early” expressed bHLH factors include MASH1, MATH1, and the neurogenins (NGN1, NGN2, and NGN3) (Akazawa et al., 1995; Gradwohl et al., 1996; Lo et al., 1991; Ma et al., 1996; Sommer et al., 1996).

The early expressed bHLH transcription factors overlap in their temporal expression profiles, but differ in their spatial expression in the CNS and PNS (Gradwohl et al., 1996; Sommer et al., 1996). The bHLH transcription factors appear to function both as neuronal determinants (Cai et al., 2000; Farah et al., 2000; Ma et al., 1996) and as regulators of cell-type specification (Fode et al., 1998, 2000; Ma et al., 1998; Perez et al., 1999). In addition, misexpres-
sion of the bHLH factors is detrimental to the survival of the organism. Constitutive ectopic expression of ngn1 and ngn2 in the mouse forebrain resulted in neuronal cell death (Cai et al., 2000). Forced expression of Math1 in neural progenitors also resulted in extensive cell death and embryonic lethality (Isaka et al., 1999). And finally, overexpression and/or persistent expression of Math1 in its normal tissue distribution resulted in neonatal lethality and a disruption in cerebellar granule cell differentiation (J.E.J., unpublished data). These data demonstrate the importance of precisely controlling the expression of this set of transcription factors.

How are these precise expression patterns of bHLH factors controlled? The mechanisms regulating expression of multiple neural bHLH genes, including achaete, scute, and atonal in Drosophila and Mash1 and Math1 in mouse, have been studied (Gomez-Skarmeta et al., 1995; Helms et al., 2000; Helms and Johnson, 1998; Sun et al., 1998; Verma-Kurvari et al., 1996, 1998). The regulation of these genes is controlled by multiple regulatory cassettes, each directing expression to different subsets of the total expression pattern. The regulatory elements controlling bHLH gene expression frequently span large genomic regions (Gomez-Skarmeta et al., 1995; Verma-Kurvari et al., 1996). The idea that the activities of multiple upstream factors combine to result in the exquisitely patterned expression is supported by the extensive sequence homology between enhancers in the human, mouse, and chicken bHLH genes (Helms et al., 2000; J.E.J., unpublished data).

Autoregulation is another component of bHLH gene expression. Direct positive autoregulation has been shown for achaete, atonal, and Math1 (Cull and Modolell, 1998; Gómez-Skarmeta et al., 1995; Helms et al., 2000; Sun et al., 1998). In contrast, indirect negative feedback autoregulation has been observed for Mash1 (Meredith and Johnson, 2000). These differences in autoregulation may simply reflect the discrete elements examined in each study, or they may reflect mechanistic differences in how these genes interface with other regulatory pathways (Chitnis, 1999; Jan and Jan, 1994; Kageyama et al., 1997).

There has been no systematic study of neurogenin2 (ngn2) regulation. Our aim is to define the mechanisms that control the precise pattern of ngn2 expression in the developing nervous system. Here we address this question by assaying genomic sequences flanking the ngn2 coding region for the ability to direct lacZ reporter gene expression to the neural tube of E11.5 mouse embryos. These studies reveal a complex pattern of regulation derived from multiple cis-acting regulatory elements. A subset of ngn2 expression in the ventral spinal cord is controlled through a discrete element, providing the first identified putative binding site for a transcription factor(s) that directly regulates ngn2 expression. These studies also provide tools for identifying upstream regulators of ngn2 expression and for characterizing the neuronal subtypes generated from NGN2-expressing progenitor cells.

MATERIALS AND METHODS

Genomic Sequence Analysis and Transgenic Constructs

Genomic sequences spanning the ngn1 and ngn2 loci were derived from overlapping clones either provided by D. Anderson or directly identified by low-stringency hybridization with the coding regions of ngn1 and ngn2 using mouse 129 (Stratagene), human leukocyte EMBl3 (Clontech), or chicken spleen Lambda FIX II (Stratagene) genomic libraries. Fragments were subcloned into PBS SK (+) (Stratagene) and directly sequenced by primer walking using ABI 310 and 377 automated fluorescence sequencers (Applied Biosystems). The mouse, human, and chick ngn2 genomic sequences were submitted to GenBank as AF303001, AF303002, and AF303000. The reporter construct used for the transgenic studies was BG2A, which contains the β-globin minimal promoter, lacZ gene, and SV40 polyadenylation cassette (Helms et al., 2000; Yee and Rigby, 1993). Fragments flanking the ngn2 coding region were cloned into the multiple cloning site of BG2A (5′ of the β-globin minimal promoter) as described below and diagrammed in Fig. 1. TgN2-1: A 7.4-kb fragment 5′ of the ngn2 coding region was generated using a vector-derived 5′ NotI site and an endogenous 3′ XhoI site. TgN2-2: This is a 4.4-kb PstI fragment contained within TgN2-1. TgN2-3: This is a 2.2-kb Smal/BglII fragment located 3′ of ngn2. Due to a rearrangement in the parent clone containing these ngn2 sequences, this construct also contained 3.2 kb of genomic sequence >5 kb downstream of the ngn2 coding region. The 3.2-kb fragment was tested and shown to have no enhancer activity. TgN2-4: A 1.0-kb fragment was PCR amplified with the primers 5′-TAGGATCCATTTTTGTAATT TTTCTTAG-3′ and 5′-GAGATCTGACCTCGT TGAAGCAGTGA-3′. The BamHI site on the former primer is endogenous. Nucleotide sequence analysis of a ngn2-containing EST (GenBank Accession No. AW413187) derived from a mouse forebrain cDNA library revealed that the 5′ 886 bp of the TgN2-4 fragment (1024 bp) was contained within the 3′ UTR of ngn2. PCR strategies were used to generate the deletions and mutations in the next four constructs. Based on the 1047 bp in TgN2-4, bases 372–662 and 661–913 were deleted in transgenes TgN2-5 and TgN2-6, respectively. Transgene TgN2-7 replaced 34 bp (459-506) within the context of TgN2-2 using the primers 5′-TAGGATCCATTTTTGTAATT TTTCTTAG-3′ and 5′-GAGATCTGACCTCGT TGAAGCAGTGA-3′. TgN2-8 replaced 35 bp (600–634) using the primers 5′-GAGATCTGACCTCGT TGAAGCAGTGA-3′ and 5′-GAGATCTGACCTCGT TGAAGCAGTGA-3′. Transgene TgN2-9: A 290-bp fragment (372–662) was PCR amplified and multimerized to generate four head-to-tail copies using the restriction enzymes XbaI, SpeI, and SacI. Transgene TgN2-10 replaced 7 bp within the context of TgN2-4 using the primers 5′-GGGATCCATTTTTGTAATT TTTCTTAG-3′ and 5′-GGGATCCATTTTTGTAATT TTTCTTAG-3′. Transgene TgN2-11 replaced 14 bp within the context of TgN2-4 using the primers 5′-GGGATCCATTTTTGTAATT TTTCTTAG-3′ and 5′-AGGATCTGACCTCGT TGAAGCAGTGA-3′. Transgene TgN2-12 replaced 13 bp within the context of TgN2-4 using the primers 5′-GACTCGAGATTTTTGTAATT TTTCTTAG-3′ and 5′-TGTTCTGAGATTTTTGTAATT TTTCTTAG-3′. Transgene TgN2-13 replaced 6 bp within the context of TgN2-4.

Copyright © 2000 by Academic Press. All rights of reproduction in any form reserved.
using the primers 5’-ATGGATCCACATTGACCTTCTATCCCTGGG-3’ and 5’-ATGGATCCCTGAGCTTCTTTTGTGAAGTTT-3’. TgN120: A 254-bp fragment from the 3’ end of ngn1 was PCR amplified using the primers 5’-TCAGTCTAGATTTGATGGCAGAGGCT-3’ and 5’-CTAGACTAGTGGTTTTCACTG-3’ and 5’-TTACCTTGTTGGAGGAGAC-3’. With the exception of TgN2-3, studies were performed using founder embryos. TgN2-3 founder animals were outbred with B6D2F1 animals for all studies. Studies were performed as described (Birren et al., 1986) using fertilized eggs from B6D2F1 (C57BL/6 X DBA) crosses. Transgenic embryos were identified by PCR with the lacZ primer gene, and embryos were identified by PCR with the lacZ vector is inactive in the transgenic assay in the absence of an added enhancer (Helms et al., 2000; Yee and Rigby, 1993). A transgene containing a 7.4-kb fragment located 5’ of the ngn2 coding region (TgN2-1), as well as one containing a 4.4-kb subfragment (TgN2-2) of TgN2-1, drove expression of the lacZ reporter gene to a small region withing the ventral mesencephalon and the ventral neural tube (vnt) at E11.5 (Figs. 1A and 1B). For these studies, we chose to focus on the expression of the ngn2/lacZ reporter genes within the developing mouse spinal cord. Transverse sections of TgN2-2 embryos revealed lacZ expression within the ventricular zone of the ventral neural tube and persisting in motor neurons (Fig. 1H). This expression represents a subset of the expression pattern of endogenous ngn2 at this stage (Fig. 2). In contrast to the endogenous ngn2 expression pattern, lacZ expression from TgN2-2 (and other transgenes) persisted into differentiating neurons and axons that exit the ventral spinal cord. Persistent lacZ expression, which was also observed with MATH1/lacZ transgenes (Helms and Johnson, 1998), is at least partially due to the stability of β-gal and does not hinder the identification of tissue specific enhancers. These data clearly show that cis-acting elements within the 4.4-kb fragment in TgN2-2 are sufficient to recapitulate a subset of the endogenous ngn2 expression profile.

**Localisation of a 1.0-kb Enhancer 3’ of ngn2**

Since the 5’ enhancer identified above was sufficient to drive the expression of the reporter gene only in a subset of the cells normally expressing ngn2, we tested additional sequences located 3’ of the coding region. Transgene TgN2-3, which contains 90 bp of the ngn2 coding sequence and 2.1 kb 3’ of the coding region, was sufficient to drive expression of lacZ in the mesencephalon, rhombencephalon, and neural tube of E11.5 embryos (Fig. 1C). Expression within the neural tube was observed in both dorsal and ventral populations of cells, as well as in axons that extend ventrally from the neural tube (Fig. 1I). Distinct dorsal and partially overlapping ventral domains of lacZ expression were observed with the 5’ (TgN2-2) and 3’ (TgN2-3) ngn2 enhancers (Figs. 1H and 1I).

To study the 3’ enhancer in greater detail, a transgenic line was generated and used to derive a developmental time course of transgene expression (Fig. 2). Reporter gene expression was first detected at E9.5 in a broad stripe in the ventral neural tube (Figs. 2A and 2E) and persisted through E12.5 (Figs. 2B-2D and 2F-2H). A dorsal domain of expression, although faintly visible at E9.5, was readily apparent from E10.5 to E12.5. Overall, the lacZ expression profile for this transgene reflects the endogenous expression of ngn2 mRNA as seen by in situ hybridization (Figs. 2I-2K). For example, lacZ expression was consistently observed in the ventricular zone of the ventral neural tube from E9.5 to E12.5, parallelling ngn2 expression during neurogenesis. However, lacZ expression persisted in differentiated neurons, whereas ngn2 mRNA was observed only within the ventricular zone. Furthermore, the lacZ expression profile

**RESULTS**

**Identification of an Enhancer 5’ of the ngn2 Coding Region**

The bHLH transcription factor ngn2 is expressed in a precise spatial and temporal pattern during the development of the vertebrate spinal cord and brain (Gradwohl et al., 1996; Sommer et al., 1996). In order to identify cis-acting regulatory sequences that modulate the expression of this gene, we tested genomic sequences flanking ngn2 in transgenic mouse assays. Overlapping clones that contained the mouse ngn2 locus were subcloned into plasmid vectors and used to derive >10 kb of nucleotide sequence flanking the ngn2 coding region (Fig. 1). In all experiments, DNA fragments were tested for enhancer activity by placing them 5’ of the β-globin basal promoter in a reporter vector (BGZA) containing the β-globin promoter, lacZ reporter gene, and SV40 polyadenylation cassette. The BGZA vector is inactive in the transgenic assay in the absence of

**Generation and Analysis of Transgenic Mice**

Transgenes for injection were separated from vector sequences using 1% SeaPlaque agarose (FMC), purified on Etulip-D columns (Schleicher & Schuell), precipitated, and resuspended in injection buffer (10 mM Tris-HCl, pH 7.5, 0.1 mM EDTA). Transgenic mice were generated by standard procedures (Hogan et al., 1986) using fertilized eggs from B6D2F1 (C57BL/6 X DBA) crosses. Transgenic embryos were identified by PCR with the lacZ primers 5’-CGAGTGTGATCATCTGGTCG-3’ and 5’-ATGCTAGATTTGATGGCAGAGGCT-3’. With the exception of TgN2-3, studies were performed using founder embryos. TgN2-3 founder animals were outbred with B6D2F1 animals for all studies. Staged embryos were dissected from the uterus in ice-cold PBS and fixed for 30 min at room temperature in 4% formaldehyde/PBS (pH 7.2). Whole-mount β-gal staining and analysis of the embryos was performed as described (Verma-Kurvari et al., 1996). Reporter gene expression within the developing neural tube was examined using whole-mount or 200-μm transverse Vibratome sections.

**In Situ Hybridization**

Staged embryos were dissected from the uterus in ice-cold PBS and fixed for 4–6 h at 4°C in 4% formaldehyde (pH 7.2). Embryos were incubated in 30% sucrose/PBS, embedded in OCT (Tissue Tek), and cryosectioned at 30 μm. In situ hybridization was performed as described (Birren et al., 1993) using digoxigenin-labeled RNA probes. The ngn2 antisense probe was derived from a 950-bp Xhol-Xbal fragment containing the entire ngn2 coding region.
at both stages E10.5 and E11.5 more closely matched ngn2 expression at E10.5. Although these observations most likely result from differences in the stability of the two proteins, it is also possible that the activation of the lacZ reporter gene is delayed or that the accumulation of threshold levels of β-gal are required for detection. These apparent differences between N2 and N2-2 were more closely matched in mouse transgenic assays. Only restriction enzymes used to derive the ends of constructs are listed. The open box in the ngn2 genomic sequence (thick black bar) is the coding region. The DNA sequences (designated A and B) mutated in TgN2-7 and TgN2-8 correspond to the analogous regions in Figs. 3 and 4. Fragments were cloned 5′ to the β-globin basal promoter in the BGZA reporter construct. The first column on the right indicates the number of transgenic embryos (#TgM) and the second column indicates the number of embryos with β-gal activity (#expressing). Transgenic embryos were harvested at E11.5 and scored for expression in the areas of the brain and spinal cord indicated as determined by whole-mount β-gal staining and transverse sections of the neural tube at the forelimb level. Only two of the five TgN2-8 embryos expressed lacZ; in the ventral nt (asterisk). (B–G) Representative pictures of E11.5 whole-mount stained and cleared founder embryos containing the transgenes indicated. (H–M) Transverse 200-μm Vibratome sections at the forelimb level illustrating the β-gal staining in the spinal neural tube for the transgenes shown in B–G. ci, commissural interneurons; dnt, dorsal neural tube; mes, mesencephalon; mn, motor neurons; nt, neural tube; rho, rhombencephalon; vnt, ventral neural tube. Scale bar: 1 mm (B–G), 125 μm (H–M).
discrepancies notwithstanding, TgN2-3 drives the temporal and spatial expression of the lacZ reporter gene in dorsal and ventral domains similar to those that express endogenous ngn2.

Deletion analyses revealed that all of the functional regulatory activity of TgN2-3 was found within a smaller 1047-bp fragment (TgN2-4, Fig. 1D) almost entirely contained within the 3' UTR of ngn2 (see Materials and Methods). High levels of lacZ expression obtained with TgN2-4 also revealed what appeared to be axonal projections from the dorsal neural tube toward the floor plate, projections characteristic of commissural interneurons (Fig. 1J). In summary, ~1.0 kb of sequence 3' of ngn2 is sufficient to drive the expression of the lacZ reporter gene in distinct neural progenitor populations within the developing neural tube. Persistent reporter gene expression suggested that commissural interneurons are generated from the dorsal neural tube, and motor neurons are derived from the ventral neural tube. Combined with the data from the 5' enhancer, we have identified distinct enhancers that regulate ngn2 expression both 5' and 3' of the coding region. However, the combined 10-kb flanking sequence tested does not drive expression in all endogenous ngn2 domains of expression, suggesting that sequences outside this region carry additional regulatory elements. For example, we have not detected lacZ expression in the dorsal telencephalon or the dorsal root ganglia, two important domains of ngn2 expression.

Species Conservation of the 3' ngn2 Enhancer

Cis-acting regulatory sequences that modulate the expression of specific genes are frequently conserved between species (Helms et al., 2000; Rowitch et al., 1998).
Distinct Enhancers for Dorsal and Ventral ngn2 Expression Domains

Further delineation of the regulatory elements within the 3' ngn2 enhancer focused on the region of homology between the mouse, human, and chick sequences. This region was approximately divided in half and these regions were deleted in the context of TgN2-4 to derive transgenes TgN2-5 and TgN2-6 (Fig. 1A). Deletion of the 5' 290 bp (TgN2-5, 372–662Δ) resulted in the complete loss of reporter gene expression from the nine transgenic embryos obtained. In contrast, deletion of the 3' 252 bp (TgN2-6, 661–913Δ) resulted in embryos that expressed lacZ in the mesencephalon, rhombencephalon, and neural tube (Fig. 1E). However, analysis of transverse sections from TgN2-6 embryos (Fig. 1K) revealed a complete loss of reporter gene expression in the neural progenitor populations in the ventral neural tube compared to the parent TgN2-4 transgene (Fig. 1J). Ventral lacZ expression is retained in axons from putative commissural interneurons from the dorsal region that extend processes toward the floor plate. These data demonstrate that conserved sequences deleted from transgene TgN2-5 are absolutely required for enhancer activity in all domains, and sequences deleted from TgN2-6 are specifically required for the ventral neural tube expression seen with this enhancer. Thus, even within the 3' enhancer, multiple nonoverlapping fragments are required for the expression of ngn2 within the motor neuron progenitor populations in the ventral neural tube.

Functional Conservation of Regulatory Sequences between the ngn1 and the ngn2 Genes

Partially overlapping expression of ngn1 and ngn2 is observed during the development of the central nervous system (Gradwohl et al., 1996; Ma et al., 1996, 1997; Sommer et al., 1996). Although not confirmed at a cellular level, one apparently overlapping domain is the ventral neural tube. Comparison of the nucleotide sequences 3' of both ngn1 and ngn2 revealed a block of ~275 bp that is greater than 50% identical (Fig. 4A). This highly conserved region is entirely contained within the previously identified homology seen within the 3' UTR of human, mouse, and chick ngn2 (Fig. 3) and corresponds to the fragment deleted in transgene TgN2-5, which resulted in a complete loss of enhancer activity in all domains.

We next tested whether this ngn2 sequence (372–662), conserved between mouse ngn1 and ngn2, is sufficient to direct expression in the neural tube in the ngn2 expression pattern. Transgene TgN2-9, which contains four copies of the conserved element (372–662), was sufficient to drive lacZ expression in the mesencephalon, rhombencephalon, and neural tube (Fig. 4C). Transverse sections of the neural tube revealed lacZ expression that did not completely reflect the expression of the 1.0-kb enhancer (Fig. 4E). Reporter gene expression was mostly seen at the lateral edge of both the dorsal and the ventral neural tube and in motor neurons projecting ventrally from the ventral neural
tube (compare Figs. 4E and 1J). These data suggest that the 372–662 domain is both sufficient and necessary for driving ngn2-specific expression in the neural tube in these transgenic mouse assays, but sequences outside this region are required to refine expression to endogenous ngn2 domains.

The ngn1 sequence corresponding to the 372–662 domain of ngn2 was also multimerized to four copies (TgN1-20) and tested in transgenic mice (Fig. 4D). Although the rationale for comparing these ngn1 and ngn2 sequences was to identify potential common regulatory elements, the conserved sequences identified did not direct lacZ expression to the same domains within the neural tube. Whereas the ngn2 enhancer was observed primarily in the ventral and lateral edges of the neural tube, the ngn1 enhancer drove lacZ expression to intermediate regions of the ventricular zone, as well as to the lateral edge of the dorsal half of the neural tube (Fig. 4F). At least some of these dorsolateral cells appeared to be commissural interneurons, since neural tube (Fig. 4F). At least some of these dorsolateral zone, as well as to the lateral edge of the dorsal half of the lacZ

To identify potential transcription factor binding sites within region A that could suggest candidate upstream factors important for ventral neural tube expression of ngn2, MATInspector (Quandt et al., 1995) was used to search the TRANSFAC database (Heinemeyer et al., 1998). This revealed a potential binding site [A/T][A/T]C[AA(A/T)] associated with members of the SOX (or SRY-related) gene family, including SRY, Sox5, and Sox17 (Denny et al., 1992; Harley et al., 1994; Kanai et al., 1996). Members of the SOX gene family have previously been shown to regulate neural induction (Mizuseki et al., 1998), and overexpression of a dominant negative Sox2 in Xenopus resulted in downregulation of X-ngnr-1 (Kishi et al., 2000). Furthermore, Sox2 expression in the developing mouse at E10.5 is restricted to the ventricular zone, apparently overlapping the entire domain of ngn2 expression (Cheung et al., 2000). Given these observations, the putative SOX binding site was mutated within the context of TgN2-4 to generate transgene TgN2-10 (Fig. 5A). Reporter gene expression in the ventral neural tube was observed only in four of the six expressing embryos (Fig. 5D). This relatively subtle perturbation observed with the mutation of the SOX site, which is located approximately in the middle of region A, suggested that it is not the element within region A that is absolutely required for ventral neural tube expression. Thus the remaining 14 bp downstream and 13 bp upstream of the SOX site within region A were mutated in transgenes TgN2-11 and TgN2-12, respectively. Although lacZ expression in TgN2-12 embryos (Fig. 5F) was observed in the ventral neural tube, all six of the TgN2-11-expressing embryos (Fig. 5E) failed to show ventral neural tube expression. Detailed examination of the latter interval showed a classical 5′-TAAT-3′ homeodomain consensus binding site (Gehring et al., 1994). Direct mutation of the HOX site (TgN2-13) within the context of TgN2-4 demonstrated that this site is also not required for the expression of the lacZ reporter gene in the ventral neural tube (Fig. 5G). Taken together, these data infer that a 9-bp element comprising the sequence 5′-TGTAGGCTT-3′ is required for ventral neural tube expression (Fig. 5A). This sequence falls in a region at the 3′ end of region A that is not conserved with the ngn1 sequence (Fig. 4B). The absence of conservation is consistent with the absence of ventral neural tube expression when the ngn1 enhancer was multimerized (Figs. 4D and 4F). Searches of the TRANSFAC database with this se-

A Novel Sequence Is Required for ngn2 Enhancer Activity in the Ventral Neural Tube

Analyses of the conserved ngn1 and ngn2 sequences revealed two 24- to 28-bp regions, designated A and B, that are greater than 80% conserved (Fig. 4A). These regions are also conserved between the mouse, the human, and the chick ngn2 nucleotide sequences (Fig. 4B). Regions A and B are 100 and 83% identical, respectively, between the three species. To determine if these sequences are required for reporter gene expression, mutations of these regions within the context of the parent TgN2-4 construct were generated. When region A (475–508) was mutated (TgN2-7), lacZ expression in ventral neural tube was completely lost (Fig. 1L). Whole-mount analysis revealed no apparent change in lacZ expression in the mesencephalon, rhombencephalon, and dorsal neural tube of TgN2-7 embryos (Fig. 1F). This specific loss of the ventral neural tube domain was also seen in transgene TgN 2-6, from which sequences 661–913 were deleted (Figs. 1E and 1K), demonstrating that at least two distinct sequences are necessary for expression in this domain. Mutation of region B (600–634) resulted in a partial loss of reporter gene expression in the neural tube (Figs. 1G and 1M). In this case, overall lacZ expression was clearly attenuated and present only in the ventral neural tube in two of the five expressing embryos. These data clearly indicate that the nucleotide sequence in both regions A and B are required to recapitulate the strong levels of lacZ expression in the domains observed with the parent TgN2-4 construct, but the requirement for region A is absolute.

To identify potential transcription factor binding sites within region A that could suggest candidate upstream factors important for ventral neural tube expression of ngn2, MATInspector (Quandt et al., 1995) was used to search the TRANSFAC database (Heinemeyer et al., 1998). This revealed a potential binding site [A/T][A/T]C[AA(A/T)] associated with members of the SOX (or SRY-related) gene family, including SRY, Sox5, and Sox17 (Denny et al., 1992; Harley et al., 1994; Kanai et al., 1996). Members of the SOX gene family have previously been shown to regulate neural induction (Mizuseki et al., 1998), and overexpression of a dominant negative Sox2 in Xenopus resulted in downregulation of X-ngnr-1 (Kishi et al., 2000). Furthermore, Sox2 expression in the developing mouse at E10.5 is restricted to the ventricular zone, apparently overlapping the entire domain of ngn2 expression (Cheung et al., 2000). Given these observations, the putative SOX binding site was mutated within the context of TgN2-4 to generate transgene TgN2-10 (Fig. 5A). Reporter gene expression in the ventral neural tube was observed only in four of the six expressing embryos (Fig. 5D). This relatively subtle perturbation observed with the mutation of the SOX site, which is located approximately in the middle of region A, suggested that it is not the element within region A that is absolutely required for ventral neural tube expression. Thus the remaining 14 bp downstream and 13 bp upstream of the SOX site within region A were mutated in transgenes TgN2-11 and TgN2-12, respectively. Although lacZ expression in TgN2-12 embryos (Fig. 5F) was observed in the ventral neural tube, all six of the TgN2-11-expressing embryos (Fig. 5E) failed to show ventral neural tube expression. Detailed examination of the latter interval showed a classical 5′-TAAT-3′ homeodomain consensus binding site (Gehring et al., 1994). Direct mutation of the HOX site (TgN2-13) within the context of TgN2-4 demonstrated that this site is also not required for the expression of the lacZ reporter gene in the ventral neural tube (Fig. 5G). Taken together, these data infer that a 9-bp element comprising the sequence 5′-TGTAGGCTT-3′ is required for ventral neural tube expression (Fig. 5A). This sequence falls in a region at the 3′ end of region A that is not conserved with the ngn1 sequence (Fig. 4B). The absence of conservation is consistent with the absence of ventral neural tube expression when the ngn1 enhancer was multimerized (Figs. 4D and 4F). Searches of the TRANSFAC database with this se-

Copyright © 2000 by Academic Press. All rights of reproduction in any form reserved.
sequence, including additional sequences to account for the possibility that a transcription factor binding site was disrupted at the 3' boundary, revealed no putative binding sites for known vertebrate or invertebrate transcription factors. This suggests that either a novel transcription factor or a known transcription factor with an as yet unknown binding site is involved in the regulation of ngn1 and ngn2.

**FIG. 4.** Functional conservation of regulatory sequences within ngn1 and ngn2. (A) Analysis of 3' ngn1 and ngn2 sequences revealed a block of approximately 275 bp that is conserved. This region was aligned using the Clustal algorithm and is shown. The percentage identity between these two sequences is greater than 50%. Regions A and B, which were mutated in transgenes TgN2-7 and TgN2-8 (Fig. 1), are boxed. (B) Alignment of region A and region B sequences from mouse, human, and chick ngn2 and mouse ngn1. Nucleotides conserved between all four sequences are indicated in blue and were used to derive the consensus. Nucleotides in black are conserved between three of the four sequences, and the variant nucleotide is indicated in red. Regions where three or more of the sequences did not match are indicated in green. The essential VNT enhancer (see Fig. 5) is indicated by the black bar. (C-F) The entire 250- to 300-bp regions aligned in A were multimerized (four copies) and tested using transgenic mouse assays. (C and D) Founder embryos from these transgenes were harvested at E11.5, stained for β-gal activity, and cleared. Transgene TgN2-9 contains a 287-bp fragment of ngn2, whereas TgN1-20 contains a 202-bp fragment of ngn1. (E and F) Transverse 200-μm Vibratome sections of the embryos shown in (C) and (D), respectively. The lacZ patterns shown were observed in 7 of 10 (TgN2-9) and 3 of 5 (TgN1-20) transgenic embryos. ci, commissural interneurons; cg, cranial ganglia; dnt, dorsal neural tube; int, intermediate neural tube; nt, neural tube; mes, mesencephalon; mn, motor neurons; rho, rhombencephalon; vnt, ventral neural tube. Scale bar: 1 mm (C and D), 125 μm (E and F).
uncharacterized binding site binds to this element to regulate ngn2 expression in the ventral neural tube.

**DISCUSSION**

The bHLH transcription factor NGN2 functions in neuronal differentiation and is expressed in precise domains of the developing neural tube. To define the regulatory mechanisms and signal transduction pathways controlling the spatial and temporal expression of ngn2, transgenic mice were used to study ngn2 regulation within the developing spinal cord. Multiple enhancers were identified both 5' and 3' of the ngn2 coding region that drive expression to distinct but partially overlapping populations of neural progenitor
cells within this tissue. Extensive characterization of an ∼1.0-kb enhancer 3' of the ngn2 coding region defined the sequence 5'-TGTAGGCTT-3' that is required for lacZ expression specifically in the ventral neural tube. These data provide the first insights into the transcriptional regulation of ngn2 expression in the developing ventral neural tube and provide the basis for identifying upstream transcription factors controlling the expression of this essential gene.

**Multiple Enhancers Regulate Dorsal versus Ventral ngn2 Expression**

From E8.5 to E12.5, ngn2 is expressed in a relatively dynamic pattern within the developing neural tube (Somer et al., 1996). Transgenic analysis of the ngn2 gene revealed multiple enhancers controlling this expression. Transgenic analysis of DNA sequences 5' of the ngn2 coding region defined a 4.4-kb enhancer that is sufficient to drive the expression of a lacZ reporter gene in the ventral neural tube (with no activity in the dorsal neural tube or nonneural tissues). In contrast, sequences 3' of the ngn2 coding region identified a 1.0-kb enhancer that is sufficient to drive lacZ expression in both the dorsal and the ventral domains. Contained within this 1.0-kb enhancer are at least two elements for ventral expression and one for dorsal expression. Thus, taking the regulatory sequences both 5' and 3' of the ngn2 coding region, of approximately 10 kb of genomic sequence tested, one dorsal neural tube enhancer and three ventral neural tube enhancers were identified. Sequence analysis between these enhancers did not reveal any significant homologies. Given the relatively simple pattern of dorsoventral ngn2 expression, it was surprising to find such a complex set of regulatory elements controlling ngn2 expression in the ventral neural tube. A likely explanation for the complexity of the ventral neural tube expression is that each enhancer may mark, and be required for, a distinct population of neural progenitors. Consistent with this explanation, it was recently shown that the activity of different 5' and 3' ngn2 enhancers marks distinct ventral progenitor populations as defined by differences in coexpression of PAX6, NKX6.1, NKX2.2, and IRX3 (F. Guillemot, personal communication).

In addition to the expression domain of ngn2 in the spinal neural tube focused on in this study, ngn2 is prominently expressed in the dorsal telencephalon, ventral mesencephalon, and metencephalon in the rostral CNS, as well as the cranial and trunk sensory ganglia (Fode et al., 1998; Sommer et al., 1996). Although transgenic analysis of the 10 kb surrounding the ngn2 coding region revealed enhancers that directly express to the ventral mesencephalon and metencephalon, expression was not consistently observed in the remaining ngn2 expression domains. Thus, cis-acting elements required for ngn2 expression in the remaining ngn2 domains likely reside outside of the 10 kb screened.

**Dorsoventral Patterning of ngn2 Expression**

The dorsoventral axis of the neural tube is set up by competing signals, BMP factors in the dorsal neural tube and Sonic hedgehog in the ventral regions (Briscoe and Ericson, 1999; Lee and Jessell, 1999). These signals set up patterns of homeodomain transcription factor expression that are important in specifying the types of neurons that form along the dorsoventral axis (Briscoe et al., 2000; Ericson et al., 1996; Tanabe et al., 1998). The pattern of ngn2 expression in the neural tube cannot be simply explained by invoking responsiveness to one of these factors or the other since ngn2 has distinct dorsal and ventral expression domains. However, a reasonable hypothesis is that ngn2 has distinct enhancers for the dorsal and ventral domains of expression and these individually may respond to BMP or Sonic hedgehog signaling. In support of this idea, we have identified a novel regulatory element that is required for the ventral domain of ngn2 expression but is not essential for the dorsal expression domain. Whether the distinct dorsal- and ventral-acting enhancers respond directly or indirectly to BMP and Sonic hedgehog signals has not been determined. Identification of the transcription factors that bind these identified enhancers will go far toward answering this question.

Numerous transcription factors are expressed in the ventricular zone of the developing neural tube with expression overlapping that of the bHLH factors. Different classes of homeodomain-containing factors, including members of the Dbx, Irx, Nkx, Msx, Pax, and Sox gene families, are expressed in the progenitor cells in the developing neural tube (Bendall and Abate-Shen, 2000; Briscoe et al., 2000; Cheung et al., 2000; Ruiz i Altaba, 1997). A current paradigm in neural tube development is that the combinatorial expression of members of these gene families specifies cell fate (Briscoe et al., 2000; Goulding, 1998; Lee and Jessell, 1999). Furthermore, these genes may be patterned in response to extracellular signals such as Sonic hedgehog and BMPs in gradients along the dorsoventral axis of the neural tube (Briscoe et al., 2000; Lee and Jessell, 1999; Lee et al., 1998). The nature of how these signals, and how the homeodomain-containing transcription factors, directly or indirectly interface with the neural bHLH factors to regulate their expression remains to be elucidated.

We have attempted to identify candidates for the upstream factor(s) controlling the ngn2 ventral expression domain through region A. Identification of SOX and HOX consensus binding sites in region A suggested members of these families as candidates for regulators of ngn2 expression. These factors were intriguing possibilities since, as described above, multiple members of these families are expressed in the ventral neural tube (Briscoe et al., 2000; Cheung et al., 2000) and have been implicated in X-ngnr-1 regulation in Xenopus (Kishi et al., 2000). However, transgenes with mutations in these sites did little to abrogate the activity of the enhancer. The conclusion from the mutation studies is that the sequence 5'-TGTAGGCTT-3' is the
important element for the ventral neural tube expression of the 3' ngn2 enhancer. Extensive sequence analysis of the essential VNT enhancer revealed that the sequence 5' GCTTTTG-3', which spans the break point, was present in the 5' and both 3' ngn2 enhancers. Database searches with this element revealed no potential transcription factor binding sites, and the significance of this sequence within each enhancer for ventral neural tube expression remains to be elucidated. Identification of the upstream factor(s) that binds this site may reveal how ngn2 expression is regulated by known signaling pathways in the ventral neural tube.

NGN2-Expressing Progenitor Cells Generate Distinct Neuronal Cell Types

Overall, the lacZ expression profiles of the transgenes identified here reflect endogenous ngn2 expression. However, lacZ expression often appeared slightly delayed in its initial expression and persisted in postmitotic neurons. These characteristics of reporter gene expression relative to the endogenous gene have been seen in similar studies of Mash1, Math1, and ngn1 regulation (Helms et al., 2000; Helms and Johnson, 1998; Verma-Kurvari et al., 1996, 1998; J.E.J., unpublished data). Although the reasons for this are unclear, it is likely due in part to differences in protein stability between the reporter, β-gal, and the endogenous bHLH factors. However, this feature of reporter gene expression allows for the partial characterization of the types of cells derived from bHLH factor-expressing progenitor populations (Helms and Johnson, 1998). The ngn2 expression in the dorsal neural tube was primarily observed at the lateral edge of the ventricular zone. In high-expressing embryos, β-gal activity in axons extending from the dorsal domain toward the floor plate was observed (Figs. 1J and 1K). In the absence of the ventral expression domain (Fig. 1K), these axons can be seen to cross the floor plate. The characteristics of these axons suggest that at least some of the neurons generated from dorsal NGN2-expressing progenitors are sensory commissural interneurons. Subsets of dorsal commissural interneurons, likely distinct from the type observed here, are derived from MATH1- and NGN1-expressing progenitor cells in the dorsal neural tube as well (Helms and Johnson, 1998; J.E.J., unpublished data). In the ventral neural tube, persistent lacZ expression from all three ventral enhancers was observed in motor neurons and their axons. Although we have not defined motor neuron subtype(s) marked by each of the enhancers, it was recently shown that the different 5' and 3' ngn2 enhancers direct expression to partially overlapping populations of progenitors that generate ventral interneurons and motor neurons (F. Guillemot, personal communication).

Sequence Homology of ngn2 Enhancers

Nucleotide sequence analysis of the 3' mouse, human, and chick ngn2 sequences revealed a conserved region that was specifically tested for activity in mouse transgenic assays. This block of homology, which spans a relatively large region (~500 bp), was shown to correspond to a functional enhancer. Expression of other bHLH transcription factors, including Mash1 and Math1, are also regulated by conserved enhancers that span sizeable genomic regions (Helms et al., 2000; Verma-Kurvari et al., 1996, 1998; J.E.J., unpublished data). Just as was seen for these other bHLH genes, the percentage identity of 800 bp within the 3' UTR of human and mouse ngn2 (91%) is significantly greater than the percentage identity (77%) of the ngn2 coding region. The simplest explanation for why these regions are so highly conserved is that they contain binding sites for multiple upstream transcription factors that together, in a combinatorial fashion, bind the regulatory region and result in expression in the precise temporal and spatial domains required for correct neural development. Consistent with this explanation, dissection of a 450-bp sequence conserved between mouse and human that controls TrkA expression revealed the importance of multiple consensus binding sites for distinct classes of transcription factors (Ma et al., 2000). One unique feature of the 3' enhancer is that it is almost entirely contained within the 3' UTR of ngn2. Thus, the high level of sequence conservation could also reflect aspects of mRNA structure that could be important for RNA stability which are not tested in the assays here.

Since ngn1 and ngn2 appear to have similar patterns of expression in both the CNS and the PNS (Sommer et al., 1996), we searched for sequence conservation between these genes in noncoding regions. Nucleotide sequence analysis of the 3' UTR of ngn1 and ngn2 revealed conserved blocks of homology. We predicted that transgenes derived from these sequences (TgN2-9 and TgN1-20) would express in similar domains of the developing neural tube. Although both transgenes directed expression only to neural tissue, contrary to our prediction, precise patterns of expression within the neural tube were distinct. Thus, the conserved sequences between ngn1 and ngn2 likely bind some common upstream factors that direct their expression specifically to neural tissue. In contrast, the distinct patterns of expression, such as the ventral expression in motor neurons in the ngn2 transgene but not the ngn1 transgene, likely reflect regions of the sequence that are not conserved. Consistent with this idea, the ventral neural tube element defined for the ngn2 enhancer, while found within the region conserved between mouse, human, and chick, is itself not conserved with the mouse ngn1 sequence (Fig. 4B).

Comparing the gene sequences of ngn2 from several species (and ngn1 from mouse) greatly facilitated the identification and dissection of potential regulatory sequences controlling ngn2 expression. Given the recent advances in sequencing the human and mouse genomes, it seems likely that a comparison of the genomic sequence surrounding coding regions may rapidly identify multiple regulatory elements for a given gene and thus greatly facilitate studies of eukaryotic transcriptional regulation.
ACKNOWLEDGMENTS

We gratefully acknowledge Drs. Q. Ma, L. Sommer, and D. Anderson for genomic clones. We thank Dr. F. Guillemot and colleagues for communication of data prior to publication. We also thank T. Savage, K. Gowen, L. Task, and A. Agerson for expert technical assistance and Drs. K. Zimmerman, A. Meredith, A. Helms, and Mr. P. Ebert for critical evaluation of the manuscript. This work was supported by NIH Postdoctoral Training Grant NS10959 to A.D.S. and a March of Dimes grant and NIH Grant HD37932 to J.E.J.

REFERENCES


Copyright © 2000 by Academic Press. All rights of reproduction in any form reserved.


Received for publication September 13, 2000
Revised October 18, 2000
Accepted October 18, 2000
Published online December 13, 2000