MOLECULAR AND CELLULAR BIOLOGY, Oct. 1996, p. 5792–5800 Vol. 16, No. 10
0270-7306/96/$04.00+0

neuroD2 and *neuroD3*: Distinct Expression Patterns and Transcriptional Activation Potentials within the *neuroD* Gene Family

MARY B. MCCORMICK,[1] RULLA M. TAMIMI,[1] LAUREN SNIDER,[1] ATSUSHI ASAKURA,[1] DONALD BERGSTROM,[1] AND S. J. TAPSCOTT[1,2]*

Fred Hutchinson Cancer Research Center, Seattle, Washington 98104,[1] and Department of Neurology, University of Washington Medical Center, Seattle, Washington 98109[2]

Received 16 April 1996/Returned for modification 29 May 1996/Accepted 25 July 1996

We have identified two new genes, *neuroD2* and *neuroD3*, on the basis of their similarity to the neurogenic basic-helix-loop-helix (bHLH) gene *neuroD*. The predicted amino acid sequence of neuroD2 shows a high degree of homology to neuroD and MATH-2/NEX-1 in the bHLH region, whereas neuroD3 is a more distantly related family member. *neuroD3* is expressed transiently during embryonic development, with the highest levels of expression between days 10 and 12. *neuroD2* is initially expressed at embryonic day 11, with persistent expression in the adult nervous system. In situ and Northern (RNA) analyses demonstrate that different regions of the adult nervous system have different relative amounts of *neuroD* and *neuroD2* RNA. Similar to neuroD, expression of neuroD2 in developing *Xenopus laevis* embryos results in ectopic neurogenesis, indicating that *neuroD2* mediates neuronal differentiation. Transfection of vectors expressing neuroD and neuroD2 into P19 cells shows that both can activate expression through simple E-box-driven reporter constructs and can activate a reporter driven by the *neuroD2* promoter region, but the *GAP-43* promoter is preferentially activated by neuroD2. The noncongruent expression pattern and target gene specificity of these highly related neurogenic bHLH proteins make them candidates for conferring specific aspects of the neuronal phenotype.

Transcription factors of the basic-helix-loop-helix (bHLH) family are implicated in the regulation of differentiation in a wide variety of cell types, including trophoblast cells (10), pigment cells (31), B cells (29, 42), chondrocytes and osteoblasts (11, 33), and cardiac muscle cells (8, 15). Perhaps the most extensively studied subfamilies of bHLH proteins are those that regulate myogenesis and neurogenesis. The myogenic bHLH factors, (MyoD, myogenin, Myf5, and MRF4), appear to have unique as well as redundant functions during myogenesis (37, 39). It is thought that either Myf5 or MyoD is necessary to determine myogenic fate, whereas, myogenin is necessary for events involved in terminal differentiation (14, 24, 28, 36). Recent work on neurogenic bHLH proteins suggests parallels between the myogenic and neurogenic subfamilies of bHLH proteins. Genes of the *Drosophila melanogaster achaete-scute* complex and the *atonal* gene have been shown to be involved in neural cell fate determination (2, 9, 17), and the mammalian homologs *MASH1* and *MATH1* are expressed in the neural tube at the time of neurogenesis (1, 22). Two related vertebrate bHLH proteins, neuroD and NEX-1/MATH-2, are expressed slightly later in central nervous system development, predominantly in the marginal layer of the neural tube and persisting in the mature nervous system (6, 21, 30). neuroD was also cloned as a factor that regulates insulin transcription in pancreatic beta cells and was named Beta2 (25). Constitutive expression of neuroD in developing *Xenopus* embryos produces ectopic neurogenesis in the ectodermal cells, indicating that neuroD is capable of regulating a neurogenic program. A *neuroD* homolog has been identified in *Caenorhabditis elegans* (21), suggesting that this molecular mechanism of regulating neurogenesis may be conserved between vertebrates and invertebrates.

We have isolated two novel genes related to *neuroD*: *neuroD2* and *neuroD3*. Similar to *MATH1*, the expression of *neuroD3* peaks during embryonic development and is not detected in the mature nervous system. *neuroD2* shows a high degree of sequence similarity to both *neuroD* and *NEX-1*/*MATH2* and is similarly expressed both during embryogenesis and in the mature nervous system, demonstrating an expression pattern that partially overlaps that of *neuroD*. Like *neuroD*, *neuroD2* expression in *Xenopus* embryos induces neurogenesis in ectodermal cells. Transfection of expression vectors for *neuroD* and *neuroD2* indicates that these highly similar transcription factors demonstrate some target specificity, with the *GAP-43* promoter activated by neuroD2 and not by neuroD. The partially overlapping expression pattern and target specificity of neuroD and neuroD2 suggest that this group of neurogenic transcription factors may contribute to the establishment of neuronal identity in the nervous system by acting on an overlapping but noncongruent set of target genes.

MATERIALS AND METHODS

Library screening. The human *neuroD2* and *neuroD3* genes were cloned from a genomic library in lambda FIX II made from human fibroblasts (Stratagene). LE392 cells were infected and plated in top agar to yield approximately 10^6 colonies on 20 15-cm-diameter plates. The plates were overlaid with Hybond filters, which were then dried, denatured in 0.5 M NaOH–1.5 M NaCl, neutralized in 1.5 M NaCl–0.5 M Tris (pH 7.2), and then cross-linked with UV. Membranes were probed with a 1-kb fragment from the murine *neuroD1* cDNA (22) labeled with [^{32}P]dCTP according to the instructions of the Random Priming Kit (Boehringer Mannheim). Membranes were prehybridized in FBI buffer (10% [wt/vol] polyethylene glycol 800, 7% sodium dodecyl sulfate [SDS], 1.5× SSPE [1× SSPE is 0.18 M NaCl, 10 mM Na_2HPO_4, and 1 mM EDTA; pH 7.7]) at 65°C for 10 min and hybridized in FBI buffer with 10 μg of denatured salmon sperm DNA per ml and probe at (0.25 to 0.5) × 10^7 cpm/ml at 65°C for 8 to 12 h. Membranes were washed at a final stringency of 0.1× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate)–0.1% SDS at 55°C. To clone the mouse *neuroD2* and *neuroD3* genes, a similar protocol was used to screen a 129Sv mouse

* Corresponding author. Mailing address: 1124 Columbia St., Seattle, WA 98104. Express mail address: Room C3-168, 1100 Fairview Ave. N., Seattle, WA 98109. Phone: (206) 667-4499. Fax: (206) 667-6524. Electronic mail address: stapscot@fred.fhcrc.org.

genomic library (provided by Glenn Friedrich) cloned in lambda-Dash II (Stratagene). To clone the mouse *neuroD2* cDNA, a cDNA library made from 16-day mouse embryos in the lambda-SH/lox (Novagen) vector was screened. Attempts to clone the mouse *neuroD3* cDNA from this library were not successful.

RNA preparation and Northern (RNA) blot analysis. Total RNA was isolated from whole mouse embryos, adult mouse tissues, and P19 cells with RNazol B according to the protocol provided (Cinna/Biotecx CS-105B). RNA was size fractionated on 1.5% agarose gels and transferred to Hybond-N membranes. Hybridization was carried out in 7% SDS–0.25 M Na_2PO_4–10 mg of bovine serum albumin per ml–1 mM EDTA at 65°C for at least 5 h and then washed in 0.1× SSC–0.1% SDS at 55 to 60°C. Northern blots were probed with fragments representing the region 3′ of the bHLH domain to avoid cross-hybridization between genes. The probe for *neuroD1* was made from a 350-bp *Pst*I fragment from the mouse *neuroD* cDNA (21) that encompasses the region coding for amino acids 187 to 304; the probe for *neuroD2* was made from a 635-bp *Pst*I fragment from the mouse *neuroD2* cDNA that encompasses the region from amino acid 210 through the 3′ nontranslated region; and the probe for *neuroD3* was made from a 400-bp *Apa*I-*Bam*HI fragment from the *neuroD3* genomic region that is 3′ prime to the region coding for the bHLH domain.

In situ hybridization. Probes for in situ hybridization were made from the mouse *neuroD1* and *neuroD2* cDNA fragments used for the Northern blotting with T3- and T7-generated transcripts for sense and antisense probes, incorporating ^{35}S-UTP label. Frozen 4- to 5-μm sagittal sections of adult mouse brain were cut, placed on Fisher Superfrost slides, and frozen at −80°C. Hybridization to ^{35}S-UTP-labeled probes and autoradiography were performed according to the method of Masters et al. (23).

Expression vectors. Expression vectors were made in pCS2+ or pCS2+MT (35), both of which contain the simian cytomegalovirus promoter, and MT contains six copies of the myc epitope recognized by the 9e10 monoclonal antibody (ATCC CRL1729) cloned in frame upstream of the insert. The 1.75-kb full-length human *neuroD1* cDNA (32) from plasmid phcnd1-17a was cloned into the *Eco*RI site to make pCS2-hND1-17s (hereafter referred to as pCS2-hND1). The 1.53-kb genomic region containing the entire coding sequence of the human *neuroD2* gene was cloned into the *Stu*I-*Xba*I site to make pCS2-hND2-14B1 (hereafter referred to as pCS2-hND2). The mouse 1.95-kb *neuroD2* cDNA was cloned into the *Eco*RI-*Xho*I sites to make pCS2-mND2-1.1.1 (hereafter referred to as pCS2-mND2). For the myc-tagged construct, a synthetic oligonucleotide-mediated mutagenesis was used to introduce an *Eco*RI site adjacent to the initial ATG codon to result in the myc tag and *neuroD2* coding regions being in-frame to make pCS2MT-mND2.

RNA microinjection and analysis of *Xenopus laevis* embryos. Synthesis of capped RNA for the *X. laevis* injections was done essentially as described previously (19) with SP6 transcription of pCS2-hND2, pCS2-hND1, and pCS2MT-mND2. The capped RNA was phenol-chloroform extracted, followed by separation of unincorporated nucleotides with a G-50 spin column. Approximately 350 pg of capped RNA was injected into one cell of a two-cell-stage albino *X. laevis* embryo in a volume of approximately 5 nl, as described previously (35). Embryos were allowed to develop in 0.1× modified Barth's saline and staged according to the method of Nieuwkoop and Faber (27). Embryos were fixed in 0.1 M morpholinepropanesulfonic acid (MOPS) (pH 7.4)–2 mM ethylene glycol-bis(β-aminoethyl ether)-*N*,*N*,*N*′,*N*′-tetraacetic acid (EGTA)–1 mM $MgSO_4$–3.7% formaldehyde for 2 h at room temperature and stored in methanol. Embryos were hydrated through a graded series of methanol–phosphate-buffered saline solutions and prepared for immunohistochemistry as described previously (35). The embryos were stained with an anti-neural cell adhesion molecule (NCAM) antibody (5) diluted 1:500 (gift of Urs Rutishauser), followed by a goat anti-rabbit alkaline phosphatase-conjugated secondary antibody, or stained with the monoclonal anti-myc tag 9e10 antibody. The presence of the antibody was visualized by nitroblue tetrazolium–5-bromo-4-chloro-3-indolylphosphate toluidinium salt color reaction according to the protocol provided (Gibco).

Tissue culture and transfections. The mouse teratocarcinoma cell line P19 (18) was cultured in minimal essential medium alpha supplemented with 10% fetal bovine serum. Transfections were performed as previously described (34) with an *N*,*N*,-bis-(2-hydroxyethyl)-2-aminoethanesulfonic acid (BES)-buffered calcium phosphate precipitation mixture. P19 cells were harvested 48 h after transfection for luciferase and *lacZ* assays. Expression vectors pCS2-hND1 and pCS2-hND2 are described above. The *GAP43* 760-bp promoter region driving luciferase in a pGL2 vector modified to contain a poly(A) site upstream of the multiple cloning site was the generous gift of Pate Skene and Joseph Weber. The pND2-luciferase construct was made by cloning a 1-kb fragment of mouse *neuroD2* sequence terminating in the first exon cloned into the pGL3 luciferase vector. The p4RTK-luciferase construct was made by placing the 4RTK region from *Hin*dIII to *Xho*I of the p4RTK-CAT vector (38) into the promoterless luciferase vector. Luciferase assays were performed according to methods described in *Current Protocols in Molecular Biology* (7).

Nucleotide sequence accession numbers. The GenBank accession numbers for the proteins described here are as follows: human neuroD2, U58681; mouse neuroD2, U58471; human neuroD3, U63842; and mouse neuroD3, U63841.

```
          BASIC REGION      HELIX I            HELIX II
position  abcdefghijklmnopqrstuvwxyz
NEUROD1   RRMKANARERNRMHGLNAALDNLRKVVPCYSK TQKLSKIETLRLAKNYIWALSEILR
NEUROD2   --Q----------D---------------- ------------------------
MATH2     --QE----------D---------------- ------------------------
NEUROD3   --V---D--------N------A--S-L-SFPD DT--T------F-Y------A-T--
MATH1     --LA------R------H-F-Q--N-I-SFNN DK----Y---QM-QI--N----L-Q
BETA3     L-LNI-----R---D--D---E--A-I-YAHSPSVR-----A--L------LMQAQAL
D.ATONAL  --LA------R--QN--Q-F-R--Q-L-
DAS-l'sc  --    -------VKQV-NGFV---QHL-
MASH1     --    -E-----VKLV-LGFAT--EH--
XASH3     --    -E-----VKLV-LGFAK--QH--
```

FIG. 1. Alignment of the bHLH region of neurogenic bHLH proteins. The neuroD1, neuroD2, and neuroD3 amino acid sequences do not vary between mice and humans in this region. D. Atonal, *D. melanogaster* atonal; DAS-l'sc, *Drosophila* lethal of scute; MASH1, mouse achaete-scute homolog 1; XASH3, *X. laevis* achaete-scute homolog 3.

RESULTS

Predicted protein sequence of neuroD2 and neuroD3. A human fibroblast genomic library (Stratagene) was screened at moderate stringency with a probe to mouse *neuroD* for the purpose of isolating the human *neuroD* gene. In addition to isolating the human *neuroD* gene (32), six of the genomic clones that were isolated showed a restriction pattern that was distinct from human *neuroD*, and subsequent sequence analysis demonstrated that the six clones represented two different genes that shared similarity with *neuroD* in the bHLH region (Fig. 1). We will refer to the original *neuroD* clone as *neuroD1* and the two new clones as *neuroD2* and *neuroD3.*

A fragment of the human *neuroD2* gene was used to screen both a mouse genomic library and an embryonic day-16 mouse cDNA library. Five independent cDNAs were mapped by restriction endonucleases and demonstrated identical restriction sites and sequences. Comparison with the corresponding mouse genomic sequence demonstrates that the entire coding region of *neuroD2* is contained in the second exon (data not shown). The sequence showed a predicted protein of 382 amino acids that differs from the major open reading frame (ORF) in the human *neuroD2* gene at only 9 residues, all in the amino-terminal portion of the protein (Fig. 2). neuroD2 has 98% similarity to neuroD1 and MATH2 in the bHLH region and 90% similarity in the 30 amino acids immediately carboxy terminal to the bHLH region (Fig. 2). Similar to neuroD1 and MATH2, neuroD2 contains an amino-terminal region rich in glutamate residues that may constitute an acidic activation domain, and has other regions of similarity to neuroD1 throughout the protein (Fig. 2).

Several attempts to isolate a cDNA for human and mouse *neuroD3* were unsuccessful. However, since all identified members of the family of genes related to *neuroD1* have their entire coding sequence in a single exon, we determined the major ORF encoded in the genomic DNA from mouse and human *neuroD3*. The genomic sequence of mouse *neuroD3* contains a major ORF of 244 amino acids, and the human *neuroD3* gene contains an ORF of 237 amino acids that differs from the predicted mouse protein at 30 positions (Fig. 2). In contrast to neuroD2, neuroD3 does not contain significant regions of homology to neuroD1 or MATH2/NEX-1 outside of the bHLH region and does not have an amino-terminal region rich in glutamates or acidic amino acids.

Overlapping and separate expression patterns during mouse development. Northern analysis was done to determine in what mouse tissues *neuroD2* and *neuroD3* were expressed. Fragments from the divergent carboxy-terminal regions of *neuroD1* and *neuroD2* and the predicted carboxy-terminal region of *neuroD3* were used to probe Northern blots containing RNAs prepared from various tissues of newborn (data not shown) and adult (Fig. 3A) mice. Both *neuroD1* and *neuroD2*

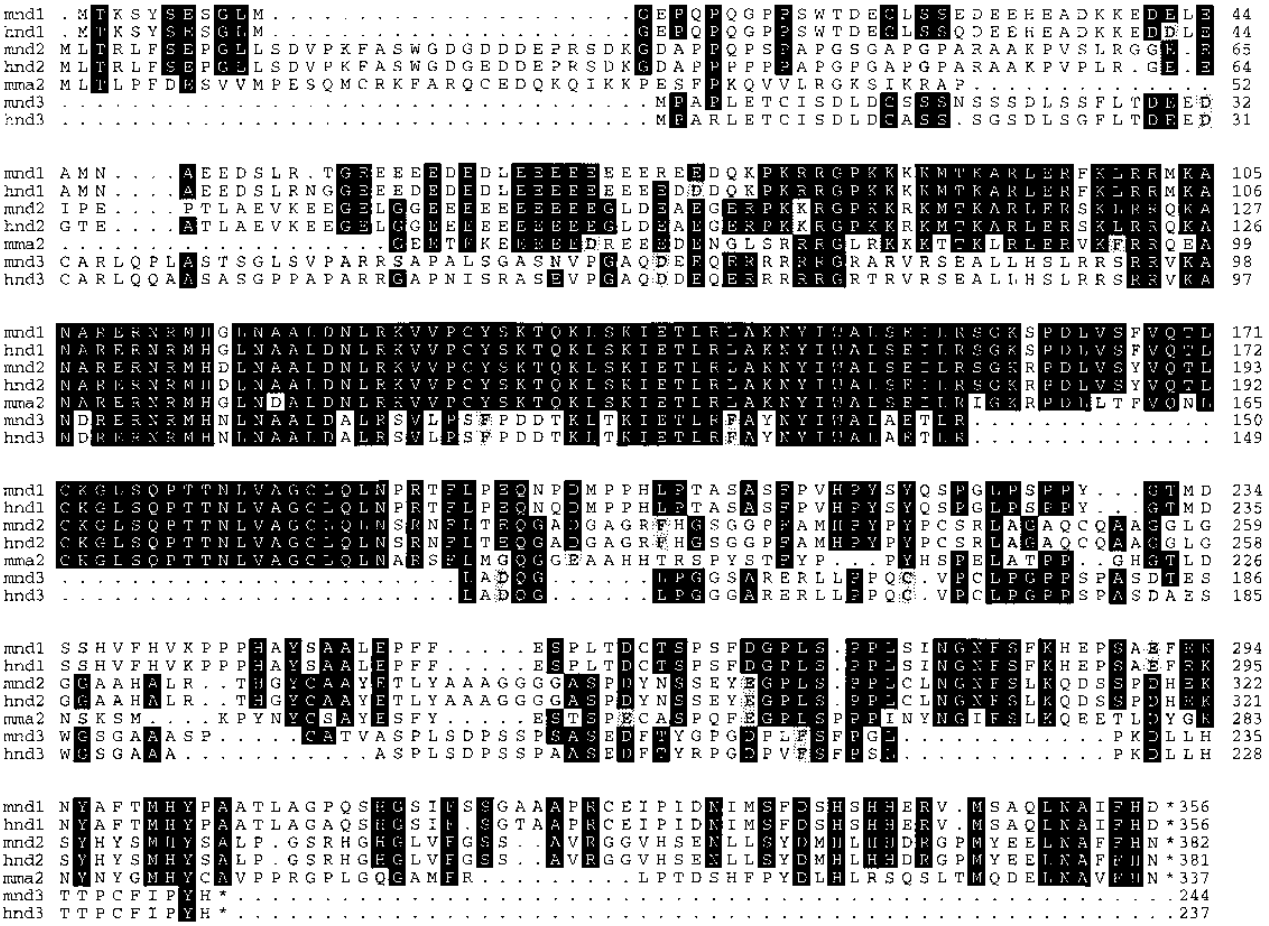

FIG. 2. Alignment of mouse and human neuroD1, neuroD2, MATH2, and neuroD3 proteins. The neuroD3 sequence is based on the major ORF in the mouse and human genomic clones.

were detected in the brains of both newborn and adult mice, whereas, *neuroD3* was not detected in any of the tissues tested. RNA extracted from dissected regions of the adult mouse nervous system demonstrated that *neuroD1* was more abundant in the cerebellum than the cortex, whereas *neuroD2* was expressed at relatively equivalent levels in both cerebellum and cortex (Fig. 3B).

To determine when during mouse embryonic development *neuroD2* and *neuroD3* were expressed in comparison with *neuroD1*, RNA was prepared from whole embryos at various developmental stages. As previously reported (21), *neuroD1* mRNA was first detected at low levels at embryonic day 9.5 and at increasing levels through embryonic day 12.5, the latest embryonic stage tested (Fig. 3C). *neuroD2* mRNA was first detected at embryonic day 11 (Fig. 3D) and also increased in abundance through embryonic day 12.5 (Fig. 3C). Although we did not detect *neuroD3* in the adult tissues, the embryonic expression pattern showed transient expression between embryonic days 10 and 12 and then declined to undetectable levels by embryonic day 16 (Fig. 3D). Collectively, these data demonstrate that *neuroD3* is expressed transiently during embryogenesis, similar to the expression pattern of *MATH1* (1), and that the temporal expression of *neuroD1* and *neuroD2* partly overlaps that of *neuroD3* but their expression persists in the adult nervous system.

To address the question of whether *neuroD1* and *neuroD2* were expressed in neurons in the adult mouse brain and whether they were expressed in the same cells, in situ hybridizations were performed with ^{35}S-UTP-labeled RNA probes. In the cerebellum, *neuroD1* was easily detected in the granule layer (Fig. 4A), whereas the *neuroD2* signal was less intense in this region and was largely restricted to the region of the Purkinje cells (Fig. 4C). In contrast, the *neuroD1* and *neuroD2* signal in the pyramidal cells and dentate gyrus of the hippocampus was easily detected (Fig. 4B and D). These results demonstrate that *neuroD1* and *neuroD2* are expressed in neuronal populations in the mature nervous system and that their relative level of expression varies among neuronal populations.

Mouse neuroD2 and human neuroD2 induce ectopic neuronal differentiation in *X. laevis*. The injection of *Xenopus neuroD1* RNA into *X. laevis* embryos resulted in the formation of ectopic neurons as well as the premature conversion of neuronal precursor cells into neurons (21). Ectopic neurons were detected with a number of neuron-specific markers, including NCAM, a neuron-specific β-tubulin, and tanabin (21). To determine if *neuroD2* was capable of inducing ectopic neuronal development in the frog, we injected mouse *neuroD2* RNA, made from pCS2-MTmND2 (see Materials and Methods), into one side of a two-cell embryo, with the uninjected side serving as a control. As with *Xenopus neuroD1*, mouse *neuroD2* was able to induce ectopic neuronal development, as determined by immunohistochemistry with an anti-NCAM antibody (Fig.

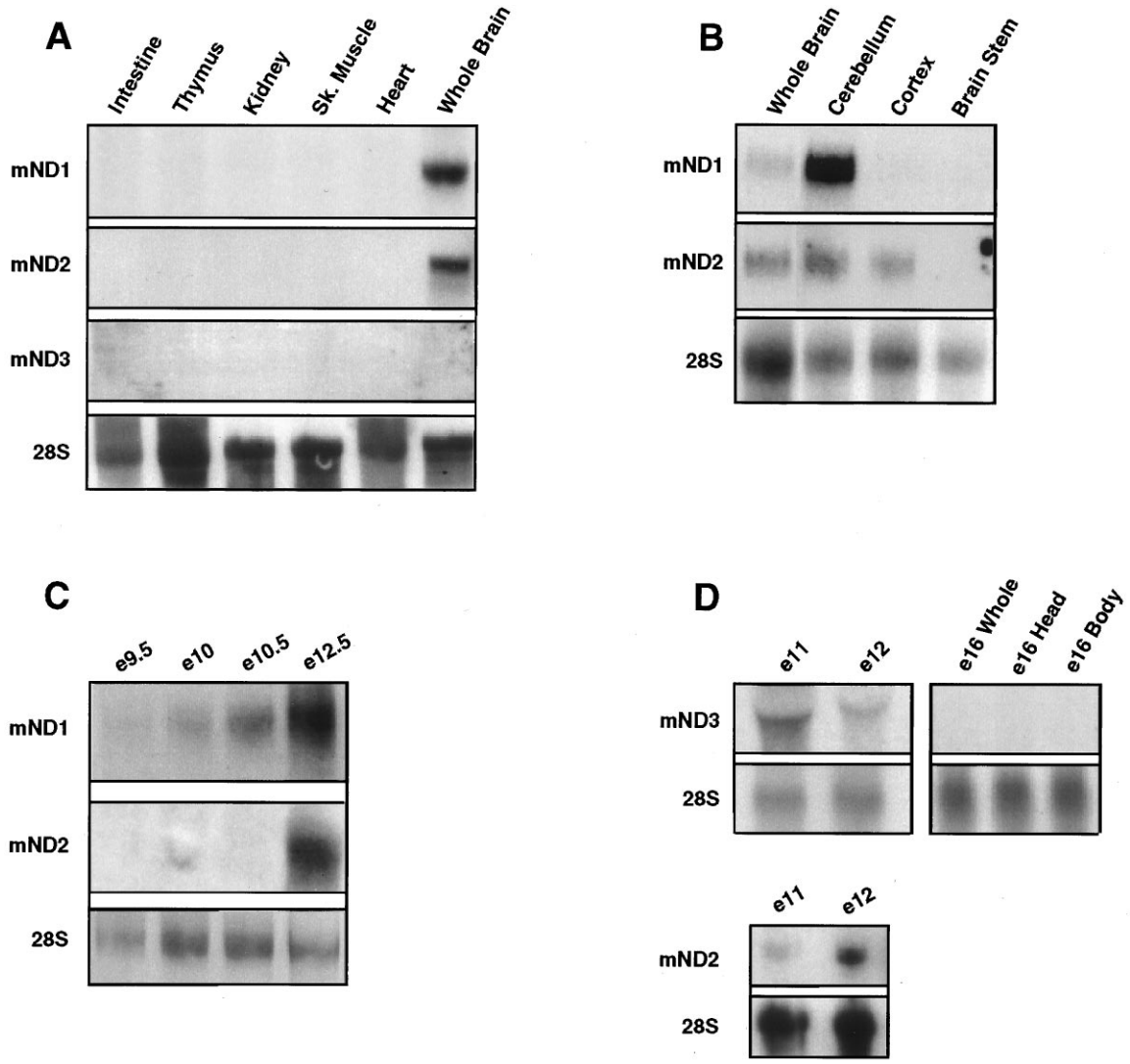

FIG. 3. Northern analysis of *neuroD* gene expression in mouse development. (A) Total RNA from the indicated adult mouse tissue was hybridized to probes from the 3′ regions of the indicated *neuroD* gene. (B) RNA from isolated subregions of adult mouse brain was prepared and hybridized to the indicated probes. (C and D) RNAs from whole mouse embryos or dissected embryonic tissue of the indicated stages were hybridized to *neuroD2* or *neuroD3* probe. Sk, skeletal.

5A). An anti-myc tag antibody, 9E10, was used to confirm that most ectodermal cells on the injected side of the frog expressed the myc-tagged mouse neuroD2 (data not shown), and approximately 80 to 90% of injected embryos stained positively with either the anti-myc or anti-NCAM antibodies. Injection of RNA encoding the human *neuroD2* gene resulted in an ectopic neuronal phenotype similar to that seen with *Xenopus neuroD1* and murine *neuroD2* (Fig. 5B). This demonstrates that both *neuroD1* and *neuroD2* can regulate the formation of neurons and that the human and mouse neuroD2 proteins are capable of functioning in the developing *Xenopus* embryo.

***neuroD* expression patterns and activity in P19 embryonic carcinoma cells.** P19 cells are a well-characterized mouse embryonic carcinoma cell line with the ability to differentiate into numerous cell types, including skeletal and cardiac muscle, or neurons and glia after treatment with dimethyl sulfoxide or retinoic acid (RA) (18), respectively. To determine whether P19 cells expressed endogenous *neuroD* genes during neuronal differentiation, we analyzed RNA for *neuroD1*, *neuroD2*, and *neuroD3* in both uninduced and induced P19 cells. To induce the formation of neurons, P19 cells were cultured as aggregates in petri dishes in the presence of RA for 4 days. The aggregates were then plated into tissue culture dishes in the absence of RA, and neuronal differentiation occurred during a 5-day period, as evidenced by the formation of neurofilament-positive process-bearing cells (data not shown). *neuroD1* mRNA was most abundant after the cells were aggregated and treated with RA for 4 days (Fig. 6, lane 6) and continued to be expressed at decreased levels during the period of neuronal differentiation. *neuroD2* was not detected during the period of RA induction, but became abundant during the period of neuronal differentiation (Fig. 6, lane 8). Both *neuroD1* and *neuroD2* signals were modestly enhanced when the differentiated P19 cultures were grown in the presence of AraC, which eliminates some of the nonneuronal dividing cells (Fig. 6, lane 9), suggesting that the *neuroD1* and *neuroD2* genes are preferentially expressed in the postmitotic cell population, but further experiments will be necessary to prove this point. *neuroD3* was first detected after 2 days of induction (Fig. 6, lane 4) and was most abundant after 4 days of induction (Fig. 6, lane 6); however, unlike *neuroD1*,

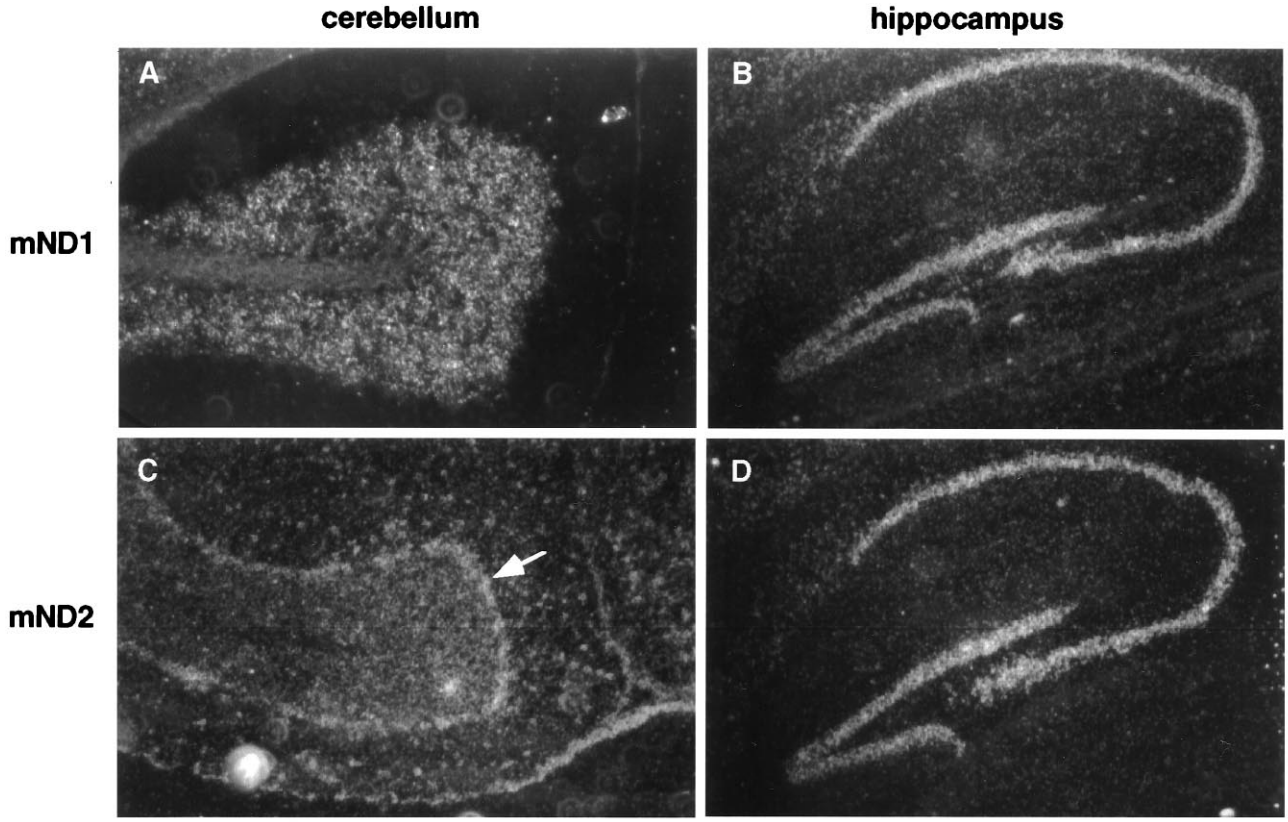

FIG. 4. In situ localization of *neuroD1* and *neuroD2* RNA in sections of adult mouse brain. Dark-field optics illuminates the silver grains as white spots. Sections were hybridized to ^{35}S-UTP-labeled RNA antisense probes to the 3′ region of *neuroD1* (A and B) or *neuroD2* (C and D). (A and C) Region of the cerebellum (magnification, ×320) with silver grains clustered over the region containing the granule cell layer in the section hybridized to the *neuroD1* probe (A). The *neuroD2* probe shows preferential hybridization to the region of the Purkinje cell layer (C [arrow]). (B and D) Hippocampus and dentate gyrus (magnification, ×160). Both groups of neurons are labeled with probe to *neuroD1* (B) and *neuroD2* (D).

neuroD3 mRNA was not detected at the later, more differentiated time points. Therefore, the temporal expression pattern of *neuroD1*, *neuroD2*, and *neuroD3* in differentiating P19 cells was similar to that seen during embryonic development: a peak of *neuroD3* expression at the time of neuronal commitment and early neurogenesis, early and persistent expression of *neuroD1*, and slightly later and persistent expression of *neuroD2*. At this time, we have not determined whether *neuroD1* and *neuroD2* are expressed in the same differentiated P19 neurons.

neuroD1 and *neuroD2* are both expressed in neurons, and both can induce neurogenesis when expressed in frog embryos. To determine if they have the ability to activate similar target genes, we constructed expression vectors driving the human *neuroD1* or *neuroD2* coding regions from a simian cytomegalovirus promoter, pCS2-hND1 and pCS2-hND2, and assayed the activity of neuroD1 and neuroD2 on reporter constructs cotransfected into P19 cells. Other members of the neuroD family have been shown to bind consensus E-box sequences in vitro. Gel shift assays have demonstrated that MATH-1 and NEX-1/MATH-2 bind the consensus E-box site CAGGTG as a heterodimer with the E47 protein and activate the transcription of reporter constructs (1, 6, 30). In vitro gel shift assays demonstrated that neuroD1 and neuroD2 can bind to an oligonucleotide containing the core E-box CACCTG as a heterodimer with an E-protein (27a). Therefore, we tested the ability of neuroD1 and neuroD2 to activate transcription of a simple reporter construct composed of a multimerized E-box with the same core sequence and the minimal promoter from the thymidine kinase gene driving the luciferase gene, p4RTK-luc. Transfection of either pCS2-hND1 or pCS2-hND2 modestly increased the level of activity from p4RTK-luc in P19 cells, increasing the activity between two- and fourfold (Fig. 7B).

Additional reporter constructs were tested in P19 cells to determine whether neuroD1 and neuroD2 had different transcriptional activation potentials. We tested the ability of pCS2-hND1 and pCS2-hND2 to transactivate a luciferase reporter construct driven by a 760-bp core promoter of the *GAP-43* gene (pGAP43-luciferase), a neuron-specific promoter construct which is upregulated in vivo in postmitotic, terminally differentiating neurons (26). In contrast to the simple E-box-driven reporter, pCS2-hND1 did not show significant transactivation of the pGAP43-luciferase, while pCS2-hND2 induced expression from this construct by approximately fourfold over the basal activity (Fig. 7C).

The myogenic bHLH proteins show auto- and cross-regulation, and expression of NEX-1/MATH-2 has been shown to activate a reporter driven by the *NEX-1/MATH-2* promoter (6). To determine if neuroD1 or neuroD2 could activate a construct containing the *neuroD2* promoter, we made a construct that contained a 1-kb fragment upstream of the mouse *neuroD2* gene, terminating in the first exon, driving the luciferase reporter gene. P19 cells were cotransfected with this pND2-luc reporter construct and the neuroD expression vectors. Both pCS2-hND1 and pCS2-hND2 transactivated this reporter construct (Fig. 7D), suggesting that *neuroD2* may be

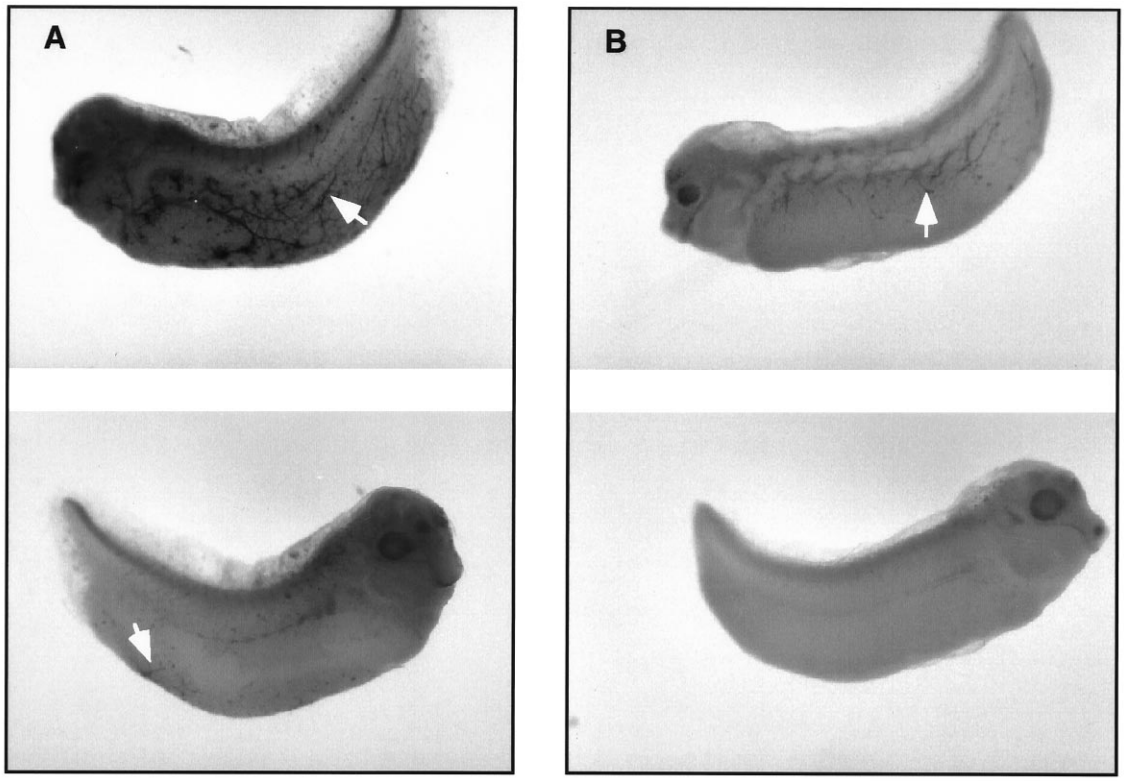

FIG. 5. Expression of *neuroD2* in *X. laevis* embryos induces ectopic neurogenesis. Embryos were injected at the two-cell stage with mouse (A) or human (B) *neuroD2* RNA. At stage 32, the embryos were stained with anti-NCAM antibody localized with alkaline phosphatase-conjugated secondary antibody. (A) Embryos injected with myc-tagged mouse *neuroD2* RNA (top, injected side of embryo, arrow indicates an NCAM-positive neuronal process; bottom, uninjected side of same embryo, arrow indicates a neuronal process that originated on the other side of the embryo). (B) Embryos injected with human *neuroD2* RNA (top, injected side of embryo, arrow indicates an NCAM-positive neuronal process; bottom, uninjected side of same embryo).

autoregulated and cross-regulated by other members of the neuroD family, in a manner analogous to the regulation of the myogenic bHLH genes.

Together these transfection experiments demonstrate that neuroD1 and neuroD2 can both activate some target genes, such as a multimerized E-box reporter and the *neuroD2* promoter, whereas the reporter construct driven by the *GAP43* promoter seems to be preferentially activated by neuroD2. At this time, we cannot quantitate the amount of protein made from each vector after transfection, and our interpretations rely on the relative activity of the reporter constructs. Further analysis of the specificity of neuroD1 and neuroD2 will require identifying specific *cis*-acting sequences in these reporters that mediate activity.

DISCUSSION

We have isolated two novel genes, *neuroD2* and *neuroD3*, that share a high degree of amino acid sequence similarity to *neuroD*/*Beta2* (21, 25) and *NEX-1*/*MATH2* (6, 30). In this report we have referred to the original *neuroD*/*Beta2* clone as *neuroD1*. The amino acid sequence similarity within the bHLH domain between *neuroD1*, *neuroD2*, and *NEX-1*/*MATH-2* is greater than 90%, suggesting the possibility of a conserved function for this subfamily of bHLH genes. In addition, like *NEX-1*/*MATH-2* and *neuroD1*, we have determined that the coding sequence of *neuroD2* is contained in a single second exon. This is similar to the *Drosophila atonal* gene structure (17), indicating that the gene structure is also conserved within this subfamily. Outside of the bHLH region of *neuroD1*, *neu-*

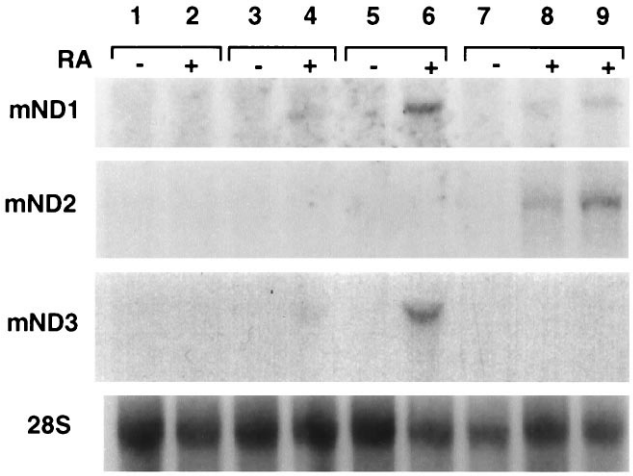

FIG. 6. Northern blot analysis of *neuroD* expression in P19 cells. P19 cells were either grown on tissue culture dishes (lanes 1 and 2) or aggregated in petri dishes for 2 days (lanes 3 and 4) or 4 days (lanes 5 and 6) either in the presence or in the absence of 0.1 μM RA as indicated. Lanes 7 to 9 show RNA from cells plated in tissue culture dishes after 4 days of aggregation and allowed to differentiate for 5 days. Both RA treatment and aggregation are necessary to induce neurogenesis. Cytosine arabinoside (0.5 mM) was added to the culture in lane 9 for the last 3 days to enrich for the nondividing neuronal population.

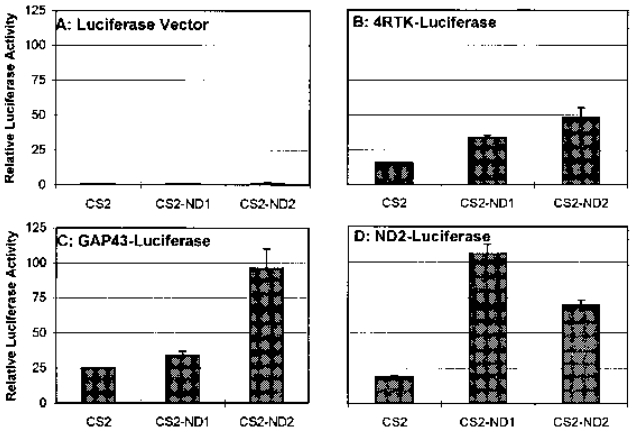

FIG. 7. Transient transfection of *neuroD* expression vectors into P19 cells. P19 cells were cotransfected with the indicated *neuroD* expression vector and reporter construct together with a cytomegalovirus-*lacZ* reporter construct. Levels of luciferase activity were normalized to the level of *lacZ* activity in each plate, and transfections were performed in triplicate. (A) The promoterless luciferase vector is not activated by any of the constructs. (B) The p4RTK-luciferase vector shows approximately twofold activation by pCS2-hND1 and fourfold activation by pCS2-hND2. (C) The pGAP43-luciferase vector is not significantly activated by pCS2-hND1, whereas there is an approximately fourfold activation by pCS2-hND2. (D) Both pCS2-hND1 and pCS2-hND2 activate a reporter driven by the *neuroD2* promoter region by four- to fivefold. Error bars represent standard errors of the mean.

roD2, and *NEX-1/MATH-2*, there are numerous domains with significant sequence similarity. All three genes contain an amino-terminal domain rich in glutamate residues, which may represent an acidic region indicative of a transcriptional activation domain. Immediately carboxy terminal to the HLH region lies a stretch of approximately 30 amino acids which is highly conserved (over 80% sequence similarity among these three family members) that has a leucine repeat periodically interrupted by proline residues. Although the overall sequence similarity in the remainder of the carboxy-terminal region of these proteins is only 60%, there are scattered regions showing sequence conservation, and potentially these conserved motifs represent important regulatory regions of the proteins.

The predicted amino acid sequence of mouse and human *neuroD3* genes is based on the major ORF in the genomic DNA, since cDNAs have not been cloned for these genes. The difficulty in obtaining *neuroD3* cDNA may be secondary to instability of the construct in the library, since deletions in the genomic DNA were common during amplification. The entire coding region of other *neuroD* family members is contained within a single exon, and therefore it is possible that the ORF in the *neuroD3* genomic DNA represents the entire coding region, a notion supported by the conservation between mouse and human sequences that extends to the stop codon. The major ORF predicts a smaller protein than related neuroD family members, and lacks the glutamate-rich amino-terminal region. The bHLH region has some elements of the loop that are similar to *MATH1*, but the overall level of homology in the bHLH region is closer to the *neuroD*-related genes.

Developmental expression patterns suggest two distinct subfamilies of neurogenic bHLH genes. *MATH1* and *neuroD3* share similarity in the bHLH region and have similar temporal expression patterns, with RNA expression detected around embryonic day 10, but not persisting in the mature nervous system. *MATH-1* RNA was localized to the dorsal neural tube in 10.5- to 11.5-day embryos, but by birth was present only in the external granule cell layer of the cerebellum, the progenitors of the cerebellar granule cell layer (1). In contrast, the *neuroD1*, *neuroD2*, and *MATH2/NEX-1* genes are expressed in both differentiating and mature neurons. Our Northern analysis demonstrates that *neuroD2* expression begins around embryonic day 11 and continues through day 16, the latest embryonic time point tested. *neuroD2* is detected in the brain of neonates as well as adult mice, with relatively equal abundance in both the cerebellum and cortex. Similar to *neuroD2*, the central nervous system expression of *neuroD1* persists postnatally, as well as showing expression in the beta cells of the pancreas (25). Northern blot analysis results presented in this paper indicate that *neuroD1* expression in the adult mouse brain is most abundant in the cerebellum, with lower levels in the cerebral cortex and brain stem. *NEX-1/MATH-2* gene expression is reported to occur by embryonic day 11.5, and at embryonic day 15.5, its expression is limited to the intermediate zone adjacent to the mitotically active ventricular zone, suggesting that *NEX-1/MATH2* is expressed primarily in the newly differentiating neurons at this stage (6, 30). In mature brain tissue, *NEX-1/MATH-2* is expressed in neurons composing the hippocampus, subsets of cortical neurons, and postmigratory cerebellar granule cells, but the reports disagree on whether this gene is expressed in the dentate gyrus of the hippocampus. It is interesting to note that Northern analysis of *MATH2* expression reported by Shimizu et al. (30) shows high levels in the cerebral cortex and low levels in the cerebellum, the opposite of the expression pattern that we see for *neuroD1*, suggesting that these genes may also have significant differences in relative abundance in specific regions of the nervous system. Therefore, it appears that *MATH-1* and *neuroD3* are expressed early in nervous system development and may have a role in either determining or expanding a population of neuronal precursors, whereas the persistent expression of *neuroD1*, *neuroD2*, and *NEX-1/MATH-2* suggests a role in initiating and maintaining expression of genes related to neuronal differentiation. While this paper was in preparation, Kume et al. (20) reported the cloning of an HLH gene from rat brain tissue by a strategy designed to identify genes that are expressed during tetanic stimulation of hippocampal neurons in a model of long-term potentiation. The gene they describe, *KW8*, is the rat homolog of the mouse and human *neuroD2* gene that we describe in this report. They also describe expression in the adult brain, including the hippocampus. Subsequently, while this paper was in review, Yasunami et al. (41) reported the mouse NDRF gene, which is nearly identical to what we have cloned as *neuroD2* and demonstrated a similar expression pattern in adult brain tissue by in situ hybridization.

While expression of either *neuroD1* or *neuroD2* in *X. laevis* embryos resulted in ectopic neuronal development, it is interesting to note that neither neuroD1 nor neuroD2 was capable of converting all cell types in which it was present into neurons. As in the case of neuroD1, the ectopic neurons induced by neuroD2 were confined to a subpopulation of ectodermal cells, as shown by the spotty NCAM-positive staining pattern. The apparent restricted activity of the neuroD proteins to a subset of cells derived from the ectoderm suggests that other factors may regulate their activity, such as the notch pathway that mediates lateral inhibition during *Drosophila* neurogenesis.

While the induction of ectopic neurogenesis by both neuroD1 and neuroD2 in *Xenopus* embryos suggests a similar function, the developmental expression patterns and in vitro transfection experiments indicate that the family members may serve both overlapping and distinct functions. Previous studies have demonstrated that neuroD1/Beta2 and NEX-1/MATH2 can bind the core CANNTG sequence of an E-box as a heterodimer with an E-protein and activate transcription. In the

work presented here, we have shown that both neuroD1 and neuroD2 can activate a construct containing multimerized E-boxes. They also activate a construct driven by a genomic fragment from the *neuroD2* gene that presumably contains regulatory regions for *neuroD2*, and the temporal expression pattern of *neuroD1* and *neuroD2* in embryogenesis and P19 differentiation suggests a model in which neuroD1 may activate *neuroD2* expression during development. Most important, however, is the demonstration that neuroD1 and neuroD2 have different capacities to activate a construct driven by the core regulatory sequences of the *GAP-43* gene, demonstrating that the highly related neuroD1 and neuroD2 are capable of regulating specific subsets of genes. This promoter contains several E-boxes, and it remains to be determined if neuroD2 directly binds to these sites.

In the bHLH region, neuroD1 and neuroD2 differ by only two amino acids, and it would be anticipated that they recognize the same core binding sequences. Therefore, the differential regulation of transcriptional activity may be determined independently of DNA binding. The amino acid following the histidine in the junction region (at position "o" in Fig. 1) of the basic region is a glycine in neuroD1, NEX-1/MATH2, and MATH1; an aspartate in neuroD2; and an asparagine in neuroD3. This residue is positioned at the same site as the lysine residue in the myogenic bHLH proteins that has been shown to be one of the sites critical for myogenic activity (12, 13, 40). In this case, it has been postulated to be a site of potential interaction with coactivator factors that regulate transcriptional activity. If the *neuroD* genes have a similar mechanism of regulation, it is possible that amino acid variability in this amino acid mediates different target specificities. Alternatively, the more divergent amino- and carboxy-terminal regions could confer regulation by interaction with other activators or repressors.

The different expression patterns in the mature nervous system and the subtle differences in target genes are similar to those of myogenic bHLH proteins. In mature muscle, *myoD* is expressed in fast muscle fibers and the myogenin gene is expressed in slow fibers (4, 16), and transfection studies demonstrate that sequences adjacent to the core E-box sequence can differentially regulate the ability of MyoD and myogenin to function as transcriptional activators (3), presumably by interaction of other regulatory factors with the non-bHLH regions of MyoD and myogenin. For the *neuroD*-related genes, the partially overlapping expression patterns and partially overlapping target genes suggest that they may act in a combinatorial fashion to directly regulate overlapping subsets of genes and thereby confer specific neuronal phenotypes. In this model, it is possible that a small family of neuroD-related transcription factors act to establish the identity of a limited number of neuronal subtypes and that local inductive events influence the generation of a higher complexity. Alternatively, it is possible that many additional members of this subfamily are yet to be identified, and they may act to directly determine specific neuronal attributes.

ACKNOWLEDGMENTS

We are grateful to Josh Weintraub, Rachel Hernandez, Jacqueline Lee, Anthony Gerber, Frank Gertler, Brian Zambrowicz, Glen Friederich, Phil Soriano, and David Turner for assistance with important aspects of this work.

S.J.T. is supported by grants from NIH, the American Cancer Society, and the Muscular Dystrophy Association. M.B.M. and A.A. are supported by a fellowship from the Muscular Dystrophy Association.

REFERENCES

1. **Akazawa, C., M. Ishibashi, C. Shimizu, S. Nakanishi, and R. Kageyama.** 1995. A mammalian helix-loop-helix factor structurally related to the product of *Drosophila* proneural gene *atonal* is a positive transcriptional regulator expressed in the developing nervous system. J. Biol. Chem. **270:**8730–8738.
2. **Anderson, D. J.** 1995. Spinning skin into neurons. Curr. Biol. **5:**1235–1238.
3. **Asakura, A., A. Fujisawa-Sehara, T. Komiya, Y. Nabeshima, and Y.-I. Nabeshima.** 1993. MyoD and myogenin act on the chicken myosin light-chain 1 gene as distinct transcriptional factors. Mol. Cell. Biol. **13:**7153–7162.
4. **Asakura, A., G. E. Lyons, and S. J. Tapscott.** 1995. The regulation of MyoD gene expression: conserved elements mediate expression in embryonic axial muscle. Dev. Biol. **171:**386–398.
5. **Balak, K., M. Jacobson, J. Sunshine, and U. Rutishauser.** 1987. Neural cell adhesion molecule expression in *Xenopus* embryos. Dev. Biol. **119:**540–550.
6. **Bartholoma, A., and K. A. Nave.** 1994. NEX-1: a novel brain-specific helix-loop-helix protein with autoregulation and sustained expression in mature cortical neurons. Mech. Dev. **48:**217–228.
7. **Brasier, A. R.** 1989. Current protocols in molecular biology. John Wiley & Sons, New York.
8. **Burgess, R., P. Cserjesi, K. L. Ligon, and E. N. Olson.** 1995. Paraxis: a basic helix-loop-helix protein expressed in paraxial mesoderm and developing somites. Dev. Biol. **168:**296–306.
9. **Campuzano, S., and J. Modolell.** 1992. Patterning of the *Drosophila* nervous system: the achaete-scute gene complex. Trends Genet. **8:**202–208.
10. **Cross, J. C., M. L. Flannery, M. A. Blanar, E. Steingrimsson, N. A. Jenkins, N. G. Copeland, W. J. Rutter, and Z. Werb.** 1995. Hxt encodes a basic helix-loop-helix transcription factor that regulates trophoblast cell development. Development **121:**2513–2523.
11. **Cserjesi, P., D. Brown, K. L. Ligon, G. E. Lyons, N. G. Copeland, D. J. Gilbert, N. A. Jenkins, and E. N. Olson.** 1995. Scleraxis: a basic helix-loop-helix protein that prefigures skeletal formation during mouse embryogenesis. Development **121:**1099–1110.
12. **Davis, R. L., P. F. Cheng, A. B. Lassar, and H. Weintraub.** 1990. The MyoD DNA binding domain contains a recognition code for muscle-specific gene activation. Cell **60:**733–746.
13. **Davis, R. L., and H. Weintraub.** 1992. Acquisition of myogenic specificity by replacement of three amino acid residues from MyoD into E12. Science **256:**1027–1030.
14. **Hasty, P., A. Bradley, J. H. Morris, D. G. Edmondson, J. M. Venuti, E. N. Olson, and W. H. Klein.** 1993. Muscle deficiency and neonatal death in mice with a targeted mutation in the myogenin gene. Nature (London) **364:**501–506.
15. **Hollenberg, S. M., R. Sternglanz, P. F. Cheng, and H. Weintraub.** 1995. Identification of a new family of tissue-specific basic helix-loop-helix proteins with a two-hybrid system. Mol. Cell. Biol. **15:**3813–3822.
16. **Hughes, S. M., J. M. Taylor, S. J. Tapscott, C. M. Gurley, W. J. Carter, and C. A. Peterson.** 1993. Selective accumulation of MyoD and myogenin mRNAs in fast and slow adult skeletal muscle is controlled by innervation and hormones. Development **118:**1137–1147.
17. **Jaman, A. P., Y. Grau, L. Y. Jan, and Y. N. Jan.** 1993. *atonal* is a proneural gene that directs chordotonal organ formation in the *Drosophila* peripheral nervous system. Cell **73:**1307–1321.
18. **Jones-Villeneuve, E. M. V., M. A. Rudnicki, J. F. Harris, and M. W. McBurney.** 1983. Retinoic acid-induced neural differentiation of embryonal carcinoma cells. Mol. Cell. Biol. **3:**2271–2279.
19. **Kreig, P. A., and D. A. Melton.** 1987. In vitro RNA synthesis with SP6 RNA polymerase. Methods Enzymol. **155:**397–415.
20. **Kume, H., K. Maruyama, T. Tomita, T. Iwatsubo, T. C. Saido, and K. Obata.** 1996. Molecular cloning of a novel basic helix-loop-helix protein from the rat brain. Biochem. Biophys. Res. Commun. **219:**526–530.
21. **Lee, J. E., S. M. Hollenberg, L. Snider, D. L. Turner, N. Lipnick, and H. Weintraub.** 1995. Conversion of *Xenopus* ectoderm into neurons by *NeuroD*, a basic helix-loop-helix protein. Science **268:**836–844.
22. **Lo, L.-C., J. E. Johnson, C. W. Wuenschell, T. Saito, and D. J. Anderson.** 1991. Mammalian *achaete-scute* homolog 1 is transiently expressed by spatially restricted subsets of early neuroepithelial and neural crest cells. Genes Dev. **5:**1524–1537.
23. **Masters, B. A., C. J. Quaife, J. C. Erickson, E. J. Kelly, G. J. Froelick, B. P. Zambrowicz, R. L. Brinster, and R. D. Palmiter.** 1994. Metallothionein III is expressed in neurons that sequester zinc in synaptic vesicles. J. Neurosci. **14:**5844–5857.
24. **Nabeshima, Y., K. Hanaoka, M. Hayasaka, E. Esumi, S. Li, and I. Nonaka.** 1993. Myogenin gene disruption results in perinatal lethality because of severe muscle defect. Nature (London) **364:**532–535.
25. **Naya, F. J., M. M. Stellrecht, and M. J. Tsai.** 1995. Tissue-specific regulation of the insulin gene by a novel basic helix-loop-helix transcription factor. Genes Dev. **9:**1009–1019.
26. **Nedivi, E., G. S. Basi, I. V. Akey, and J. H. P. Skene.** 1992. A neural-specific GAP-43 core promoter located between unusual DNA elements that interact to regulate its activity. J. Neurosci. **12:**691–704.
27. **Nieuwkoop, P. D., and J. Faber.** 1967. Normal table of *Xenopus laevis*. North-Holland Publishing Co., Amsterdam.

27a. **Olson, J. M., and S. J. Tapscott.** Unpublished data.
28. **Rudnicki, M. A., P. N. Schnegelsberg, R. H. Stead, T. Braun, H. H. Arnold, and R. Jaenisch.** 1993. MyoD or Myf-5 is required for the formation of skeletal muscle. Cell **75:**1351–1359.
29. **Shen, C. P., and T. Kadesch.** 1995. B-cell-specific DNA binding by an E47 homodimer. Mol. Cell. Biol. **15:**4518–4524.
30. **Shimizu, C., C. Akazawa, S. Nakanishi, and R. Kageyama.** 1995. MATH-2, a mammalian helix-loop-helix factor structurally related to the product of *Drosophila* proneural gene atonal, is specifically expressed in the nervous system. Eur. J. Biochem. **229:**239–248.
31. **Steingrimsson, E., K. J. Moore, M. L. Lamoreux, A. R. Ferre-D'Amare, S. K. Burley, D. C. Zimring, L. C. Skow, C. A. Hodgkinson, H. Arnheiter, N. G. Copeland et al.** 1994. Molecular basis of mouse microphthalmia (mi) mutations helps explain their developmental and phenotypic consequences. Nat. Genet. **8:**251–255.
32. **Tamimi, R., E. Steingrimsson, N. G. Copeland, K. Dyer-Montgomery, J. E. Lee, R. Hernandez, N. A. Jenkins, and S. J. Tapscott.** 1996. The NEUROD gene maps to human chromosome 2q32 and mouse chromosome 2. Genomics **34:**418–421.
33. **Tamura, M., and M. Noda.** 1994. Identification of a DNA sequence involved in osteoblast-specific gene expression via interaction with helix-loop-helix (HLH)-type transcription factors. J. Cell Biol. **126:**773–782.
34. **Tapscott, S. J., M. J. Thayer, and H. Weintraub.** 1993. Rhabdomyosarcomas lack a factor necessary for both transcriptional activation by MyoD and myogenesis. Science **259:**1450–1453.
35. **Turner, D. L., and H. Weintraub.** 1994. Expression of achaete-scute homolog 3 in *Xenopus* embryos converts ectodermal cells to a neural fate. Genes Dev. **8:**1434–1447.
36. **Venuti, J. M., J. H. Morris, J. L. Vivian, E. N. Olson, and W. H. Klein.** 1995. Myogenin is required for late but not early aspects of myogenesis during mouse development. J. Cell Biol. **128:**563–576.
37. **Weintraub, H.** 1993. The MyoD family and myogenesis: redundancy, networks, and thresholds. Cell **75:**1241–1244.
38. **Weintraub, H., R. Davis, D. Lockshon, and A. Lassar.** 1990. MyoD binds cooperatively to two sites in a target enhancer sequence: occupancy of two sites is required for activation. Proc. Natl. Acad. Sci. USA **87:**5623–5627.
39. **Weintraub, H., R. Davis, S. Tapscott, M. Thayer, M. Krause, R. Benezra, T. K. Blackwell, D. Turner, R. Rupp, S. Hollenberg et al.** 1991. The myoD gene family: nodal point during specification of the muscle lineage. Science **251:**761–766.
40. **Weintraub, H., V. J. Dwarki, I. Verma, R. Davis, S. Hollenberg, L. Snider, A. Lassar, and S. J. Tapscott.** 1991. Muscle-specific transcriptional activation by MyoD. Genes Dev. **5:**1377–1386.
41. **Yasunami, M., K. Suzuki, H. Maruyama, H. Kawakami, Y. Nagai, M. Hagiwara, and H. Ohkubo.** 1996. Molecular cloning and characterization of a cDNA encoding a novel basic helix-loop-helix protein structurally related to neuroD/BHF1. Biochem. Biophys. Res. Commun. **220:**754–758.
42. **Zhuang, Y., P. Soriano, and H. Weintraub.** 1994. The helix-loop-helix gene E2A is required for B cell formation. Cell **79:**875–884.