Available online at www.sciencedirect.com
SCIENCE DIRECT
Molecular Brain Research 112 (2003) 8–23
MOLECULAR BRAIN RESEARCH
www.elsevier.com/locate/molbrainres

Research report

The human tyrosine hydroxylase gene promoter

Mark A. Kessler[†], Ming Yang, Kandace L. Gollomp, Hao Jin, Lorraine Iacovitti*

Department of Neurology and Farber Institute for the Neurosciences, Thomas Jefferson University, Philadelphia, PA 19107, USA

Accepted 12 December 2002

Abstract

13.329 kilobases of the single copy human tyrosine hydroxylase (hTH) gene were isolated from a genomic library. The 5′ flanking 11 kilobases fused to the reporter green fluorescent protein (GFP) drove high level expression in TH+ cells of the substantia nigra of embryonic and adult transgenic mice as determined by double label fluorescence microscopy. To provide a basis for future analysis of polymorphisms and structure–function studies, the previously unreported distal 10.5 kilobases of the hTH promoter were sequenced with an average coverage of 20-fold, the remainder with 4-fold coverage. Sequence features identified included four perfect matches to the bicoid binding element (BBE, consensus: BBTAATCYV) all of which exhibited specific binding by electrophoretic mobility shift assay (EMSA). Comparison to published sequences of mouse and rat TH promoters revealed five areas of exceptional homology shared by these species in the upstream TH promoter region −2 kb to −9 kb relative to the transcription start site. Within these conserved regions (CRs I–V), potential recognition sites for NR4A2 (Nurr1), HNF-3beta, HOXA4, and HOXA5 were shared across human, mouse, and rat TH promoters.

Theme: Cellular and molecular biology

Topic: Gene structure and function: general

Keywords: Human; Tyrosine hydroxylase; Brain; Neurons; CNS; Parkinson's disease; Stem cells; Alcoholism; Transplantation; Transgenic mice; Nurr1; NR4A2; Bicoid; Gscl; Goosecoid-like; HOXA4; HOXA5; GFP; EGFP; HNF-3beta; NRSE; Neuron restrictive silencer element; Development; Pitx3; Promoter; Transgene; Neurons; Dopaminergic neurons; Dopamine; Catecholamines; Movement disorders; Genetics

1. Introduction

The enzyme tyrosine hydroxylase (EC 1.14.16.2, TH) is encoded by a single copy gene on human chromosome 11p15 [11,67] and catalyzes the rate limiting step in the synthesis of catecholamine neurotransmitters in the central and peripheral nervous systems: the hydroxylation of tyrosine to yield dopa [59]. The regulation of TH expression has been intensively studied because of the fundamental role of catecholamines in neurophysiology and the alterations associated with psychiatric illnesses. Classically, all drugs effective for treatment of psychotic symptoms have been antagonists of dopamine receptors. Therefore, a number of studies have assessed the linkage between mutations in the TH gene and disorders such as schizophrenia, bipolar disorder, and alcoholism. There is no evidence for a role in these diseases of mutations in the protein coding regions. However, polymorphisms elsewhere in the sequence [51], and in TH-regulating transcription factors [7,30] have been described.

TH also plays a critical role in a number of neurological disorders. For example, in Parkinson's disease, it is precisely the TH+ dopaminergic cells of the substantia nigra whose degeneration parallels the symptoms of the disease. Cell replacement or transplantation has been hindered by the limited supply of cells as up to six fetal brains are harvested for treatment of a single patient. Furthermore, the TH+ cells represent only 0.5–1% of the fetal mesencephalon. This may well underlie the variability and low success rate of this treatment and has further stimulated interest in the TH gene with the view of using

*Corresponding author. Tel.: +1-215-955-8118; fax: +1-215-955-2993.
E-mail address: lorraine.iacovitti@mail.tju.edu (L. Iacovitti).
[†]Recently deceased.

0169-328X/02/$ – see front matter © 2002 Published by Elsevier Science B.V.
doi:10.1016/S0169-328X(02)00694-0

the promoter for targeting a selectable marker to permit purification of dopaminergic cells.

For Parkinson's, the goal of growing replacement tissue from stem cells is attractive because this approach may circumvent problems with tissue availability and immunological rejection. Current protocols for derivation of therapeutic cells for this purpose are quite inefficient [48,95]. A recent review has emphasized the need for further knowledge of developmental pathways to realize the potential of this renewable tissue source [16]. Understanding the molecular basis of TH gene expression will enable us to rationally intervene to direct differentiation of cells toward the desired phenotype, as well as permitting promoter-driven selection of the TH+ cells.

As there is no cell line that can recapitulate the complex anatomical and developmental regulation seen in vivo, several laboratories have studied the TH promoter in transgenic mice. Using the rat TH promoter, it was found that 4.5–9 kb were required for high level expression of the reporter in the known sites of dopamine synthesis [45,76,89]. Likewise, both 5 kb [58] or 2.5 kb of the human TH promoter [32] could direct tissue-specific expression in a transgenic mouse, although, in the latter case, the entire exon–intron structure and 0.5 kb of 3′ flanking region were also included in the construct [32].

Because of this prior work, and because the sequence of the human TH promoter upstream of −500 has not been reported, we undertook cloning and sequencing of a lambda human genomic clone including ~11 kb (10,967 bp) upstream from the TH gene transcription start. To assess the function of this sequence in vivo, transgenic mice were generated in which this sequence drives the reporter, green fluorescent protein (GFP). While lacking the sensitivity of enzymatically amplified reporters, GFP allows resolution at the cellular and subcellular level and is non-destructive. Thus, in addition to convenient visualization by fluorescent microscopy, GFP allows subsequent purification of living expressing cells by fluorescent activated cell sorting (FACS). The present work demonstrates for the first time the ability of a human TH promoter sequence to direct GFP expression to TH+ neurons of the substantia nigra.

2. Materials and methods

2.1. Molecular cloning of the human tyrosine hydroxylase promoter

A partial human tyrosine hydroxylase (hTH) cDNA was purchased from ATCC (ATCC 100604). An EcoRI-Xho I fragment of this cDNA was isolated and used to screen a commercially available lambda cDNA library purchased from Stratagene (catalog no. 936201). The longest hTH cDNA clone of 2 million total plaques screened was isolated and the 5′ 350 bp extending from an internal Xho I site to an Eco RI site within the vector was purified for use as a hybridization probe. We screened 3 million plaques of a lambda human genomic library purchased from BIO 101 [60]. Of seven positive clones recovered by cre/lox mediated excision into plasmid pPS [60], two contained identical (by restriction mapping) ~15 kb inserts encompassing exons 1 and 2 of the hTH gene and 11 kb of 5′ flanking (promoter) region as determined by Southern blot analysis.

The proximal 516 bp of the hTH promoter has been previously reported [38,61]. To sequence the more distal promoter, we isolated an overlapping Not I/SacII 11 kb restriction fragment by electroelution. For 'shotgun' sequencing this fragment was subjected to titrated ultrasonic shearing to produce an average fragment size of 1.5 kb. After treatment with mung bean nuclease and 'polishing' with T4 DNA polymerase, the size range of 1.5–2 kb was isolated by electrophoresis followed by electroelution and ligated into Sma I-cut plasmid pBCKS- (Stratagene). Isolated plasmids were sequenced using the big dye terminator polymerase chain reaction (PCR) sequencing method with ABI Prism apparatus and software. This was performed by the Thomas Jefferson University Nucleic Acid Core Facility (TJUNAF) under the direction of Dr Hansjuerg Alder [29]. The resulting data was analyzed and assembled using Lasergene software (DNAstar, Madison, Wisconsin), the RepeatMasker program available at http://ftp.genome.washington.edu/cgi-bin/RepeatMasker and the Signal Scan pattern recognition software from http://bimas.dcrt.nih.gov/sw.html [19,68,91] The assembled sequence has been submitted to Genbank (accession number: AF536811; release date of 11-20-02).

2.2. Construction of transgene with GFP reporter

Plasmid phTH-11 kb-EGFP (=pMAK 288-12) places the hTH promoter upstream of the EGFP reporter gene in a manner such that the promoter–reporter cassette can subsequently be isolated from the rest of the vector as an intact 12,007 bp NotI-AflII restriction fragment. Three DNA fragments are ligated to produce plasmid phTH-11 kb-EGFP:

1. The distal (5′) hTH promoter is isolated as a 10.794 kb NotI-SacII fragment from the original genomic clone pMAK 221-21. The sequence from the unique Not I to the 5′ Sau3A (gatc) is vector-derived [60].
2. The proximal (3′) hTH promoter, including the transcription start site, was isolated as a 1.168 kb SacII-KpnI fragment and cloned into pBSIISK- (Stratagene) to yield pMAK 237. One sequencing primer (pBSII reverse primer=MAKIL 120: GGAAACAGCTATGACCATG) and one mutagenic primer MAKIL119 GACA**GGATCC**GGGCTCCGTCTCCACA) were used to amplify the proximal promoter and add a

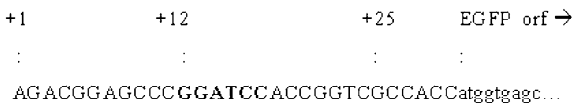

Fig. 1. Human tyrosine hydroxylase-EGFP gene fusion. The fusion of the EGFP reporter to the 5′ untranslated region of the hTH promoter is illustrated. Numbers above identify position with respect to the start of transcription of the native hTH gene. The synthetic *BamHI* site is in bold, and the open reading frame of the reporter is distinguished by lower case.

synthetic BamHI site at position +12 of the 5′ untranslated region.

3. The vector and EGFP reporter are provided as a NotI-BamHI fragment of pMAK 285, derived from pEGFP-1 (Clontech) as follows. Unique NotI and XbaI sites were removed sequentially by cleavage, fill-in, re-ligation, and transformation. A unique Not I site was inserted adjacent (3′) to the original SacII site in the multiple cloning site of the vector. The resulting construct is designated pMAK 285. The larger Not I-BamHI fragment of this plasmid is isolated and included as the vector in the ligation. The fusion of the hTH promoter 5′ untranslated region to the EGFP reporter is illustrated in Fig. 1, below, with the transcript starting at +1 in uppercase. The synthetic *BamHI* site is in bold starting at position +12. The EGFP coding sequence begins at +31 (lower case).

2.3. Generation of transgenic mice

The 12,007 bp NotI-AflII restriction fragment from plasmid phTH-11 kb-EGFP, Fig. 2, was isolated by electroelution, precipitated four times from ethanol/1 M ammonium acetate, dried, and resuspended in TE buffer. This DNA was microinjected into B6C3F2 single cell embryos that were transferred to ICR (Taconic, Germantown, NY) pseudopregnant females (Kimmel Cancer Center, Core Transgenic Mouse Facility, Thomas Jefferson University) Transgenic pups were identified by Southern blot, using the Not I-Afl II fragment as template for the random-primed probe. These founders were mated to the C57bl/6 strain for generation of F1 mice.

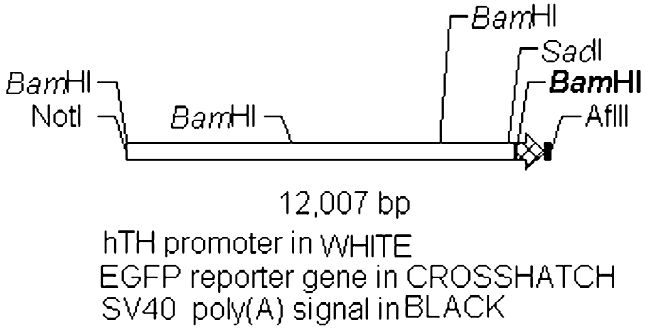

Fig. 2. Diagram of the hTH-EGFP transgene used to generate transgenic animals. This 12.007 kb NotI-AflII fragment was isolated from plasmid phTH-11 kb-EGFP, purified, and injected into single-celled embryos for generation of transgenic mice described in this application. In this diagram the hTH promoter is in white, the reporter in crosshatch, and the SV40 polyadenylation sequence in solid black.

2.4. Histology

Fixation procedures, peroxidase and fluorescence immunohistochemistry, and TH antiserum have been previously described [93]. For immunodetection of EGFP, chicken anti-GFP polyclonal antiserum (AB16901), and donkey anti-chicken IgG FITC conjugate (AP194F) were purchased from Chemicon International Inc., Temecula, CA. GFP autofluorescence may be sensitive to fixatives and mounting agents [9]. However, strong reporter signal was retained in periodate–lysine–paraformaldehyde (4%) fixed tissue and preserved for months in slides stored in the dark at 4 °C.

2.5. Graphics

Microscope images were captured with software from Scanalytics, Inc., Fairfax, VA, and composed in Adobe Photoshop 5.5. Figs. 2 and 6 were drafted using Winplas 2.7 (Goldstein Software from http://www.goldsteinsoftware.com/ Goldstein Software, 150 Green Street, Woodbridge, NJ).

2.6. Gel shift (EMSA) and supershift assays

The method for EMSA and supershift assays has been previously described [22]. For analysis of bicoid binding specificity, we employed antisera and a cDNA clone of murine goosecoid-like generously supplied by Dr J.A. Epstein, University of Pennsylvania School of Medicine, Philadelphia, PA [17]. For these experiments, nuclear extracts were prepared from COS-7 cells transiently transfected with goosecoid-like cDNA in the expression vector pcDNA3.1 (Invitrogen, Carlsbad, CA). A negative control extract was prepared from cells transfected with the same vector expressing GFP. Probes consisted of blunt-ended double-stranded 25 bp oligonucleotides representing the four bicoid binding elements (BBEs), as well as a negative control oligonucleotide representing an AP-1 response element.

BBE I:	GCCGAGCC**TGTAATCCC**AGAATTTG
BBE II:	CGAGCCCC**GGGATTAGC**AGAGGTAC
BBE III:	CTGCAACT**GCTAATCCC**TGATGCCC
BBE IV:	CTCACGCC**TGTAATCCC**AGCACTTT
AP-1:	GCTGAGGA**TGATTCA**GACAGGGC

3. Results

3.1. Expression of TH and GFP in transgenic mice

Of 93 offspring, five tested positive for the transgene by

Southern blot of tail DNA (5.4%). These were mated to non-transgenic BL6 mice. Thus, the four transgenic F1 lines we were able to obtain are necessarily heterozygous. We examined single adult individuals of each of these lines by immunocytochemistry for expression of TH and GFP (Table 1).

Robust transgene expression was detected in most traditional catecholaminergic (CA) tissues with the notable exceptions of the adrenal medulla and locus ceruleus. Conversely, ectopic expression of GFP was detected in a number of regions within the CNS. These include several areas consistently noted by authors employing shorter transgenic constructs derived from rat or human TH [45,53,76] the amygdala, interpeduncular nucleus, nucleus accumbens, and septum.

3.2. Adult basal ganglia

We used line F1-09 (Founder 7827) for examination of TH+ cells of the adult and embryonic nigro-striatal pathway. Fig. 3 compares the expression of endogenous TH and the GFP reporter within the caudate nucleus (CN, A, B) and substantia nigra (SN, C, D) of an adult transgenic mouse. While these are not adjacent sections, peroxidase immunohistochemistry nevertheless reveals close anatomical correlation between TH and GFP both in

Table 1
Summary of expression patterns of transgene in four mouse lines[a]

	Transgenic mice				Non-transgenic mice	
	GFP				GFP	TH
Founder no.	7833	7827	7829	7840		
Copy number	2	3	1	3		
Expression (cell bodies)						
Traditional CA sites						
Hypothalamus (A11–14)						
Arcuate n.	+++	+	++	+	–	++
Periventricular n.	+	++	+	+	–	++
Paraventricular n.	+	++++	++++	++	–	++
Zona incerta	++	+++	++++	++	–	++++
Olfactory bulb	ND	ND	+++	+++	–	++++
Substania nigra (A9)	++++	++++	++	+++	–	++++
VTA(A10)	++++	++++	++	+++	–	++++
Dorsal raphe n.	++	++	++	+++	–	++
Locus ceruleus (A6)	+	–	–	+	–	++++
A5	+	+	+	+	–	+++
A2	++	++	+++	++	–	+++
A1	++	+	++	+	–	++
Area postrema	+	++	++++	+	–	+
Adrenal gland	–	–	–	–	–	++++
Non-traditional CA sites						
Anterior olfactory n.	ND	ND	+	+++		
Infralimbic cortex	+	–	ND	ND		
Insular cortex	+	–	++++	++		
Somatic motor cortex	+	+	+++	++		
Somatic sensory cortex	+	+	+++	+		
N. accumbens	++	–	++	ND		
Septum	+	+	+	+		
Amygdala	++++	++	++++	+		
Bed nucleus of stria terminalis	+	+	++	+		
Endopiriform n.	++	+	++++	++		
Preoptic area	++	++	++	++		
Supraoptic n.	+	+++	++++	+		
Hippocampus	ND	–	++++	+++		
Habenular n.	ND	+	++++	+++		
Interpenduncular n.	++	++	++	++		
Inferior colliculus	–	–	+++	–		
Parabrachial n.	++	++	++++	+		
Vestibular n.	+++	+	++++	++		
Paragigantocellular n.	++	+	++	+		

ND = not determined.
[a] Two individuals of each line were examined for transgene copy number; one individual of each line for expression.

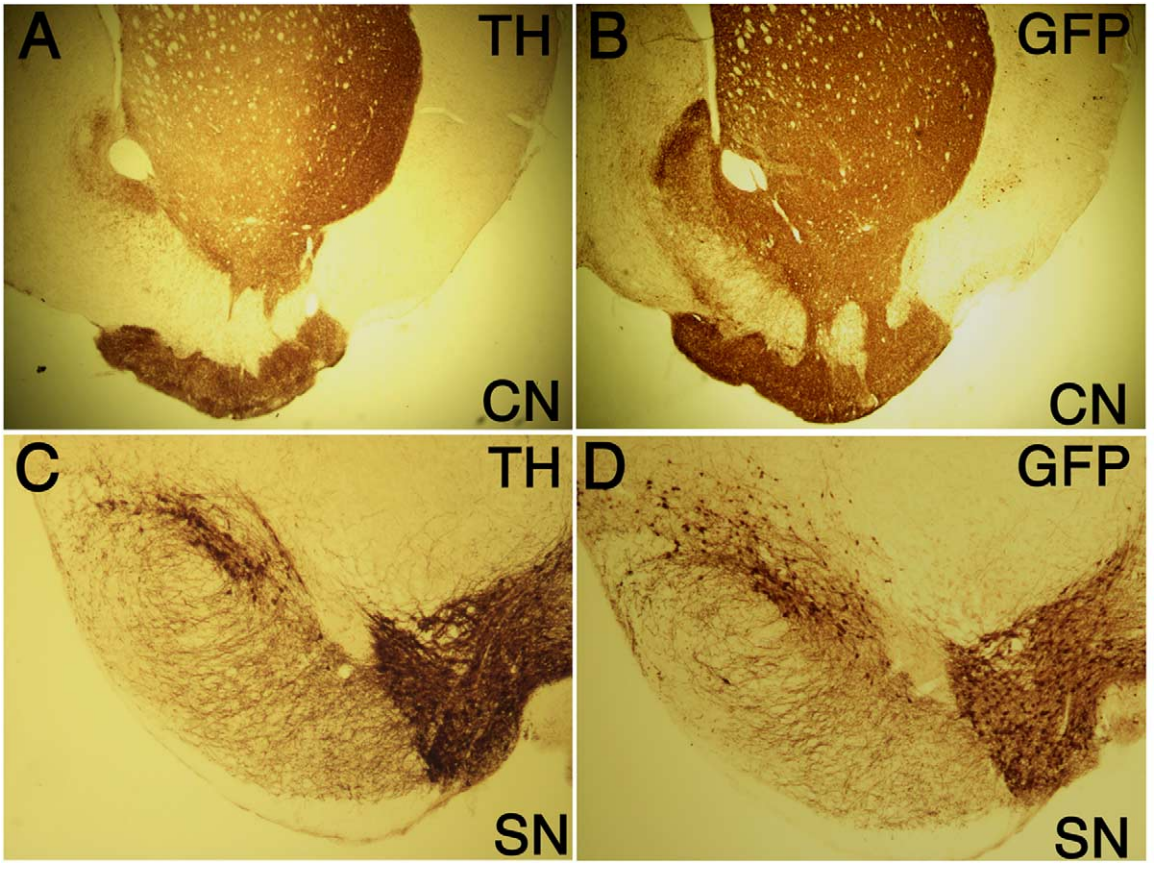

Fig. 3. Immunoperoxidase localization of endogenous TH and the GFP reporter in the basal ganglia of an adult transgenic mouse (founder no. 7827 line). Expression of TH (A, C) and GFP (B, D) is highly correlated in the caudate nucleus (CN; A, B) and substantia nigra (SN; C, D).

the cell bodies of the substantia nigra and their axonal projections to the caudate nucleus.

In Fig. 4 is illustrated a further study of TH-GFP coexpression in the ventral tegmental area (A–C) and SN (D–F) using dual-label immunofluorescence. The GFP protein is detected with a FITC conjugate (Fig. 4A, D) yielding a green signal, while a rhodamine signal marks the expression of TH in red (Fig. 4B, E). Coexpression of TH and GFP both in neuronal cell bodies and efferent fibers is indicated by the yellow signal in the merged images (Fig. 4C, F). In this field, a correlation of ~70% was determined for TH and GFP with a slight excess of singly-labeled GFP+ cells.

3.3. Embryonic midbrain

An ultimate goal of this work is to genetically tag living TH+ neurons for cell replacement treatment of Parkinson's disease. Therefore, we next examined GFP expression in a dissected E13 transgenic embryonic mouse mesencephalon (Fig. 5A). The neural tube has been opened longitudinally along the dorsal surface and is here oriented with the anterior mesencephalon at the bottom of the frame. The same field, imaged by GFP fluorescence, is shown in Fig. 5B where the strong signal within the ventral midline is easily distinguished from the surrounding tissue. This has practical significance as it is possible to choose transgenic midbrains by eye, using an inverted fluorescent microscope, and the GFP signal can guide further dissection. Thus far, the yield of E13 transgenics per pregnancy has been indistinguishable from 50% implying that the transgene is not toxic, at least in heterozygotes. Pooled GFP+ ventral midbrains (Fig. 5B) were dissociated with trypsin to a single-cell suspension and plated on poly-ornithine-coated 8-well chamber slides. We examined the cultures for morphology and retention of the GFP signal by inverted bright-field and fluorescence microscopy after incubation in culture overnight. To indicate the number of GFP+ cells among the total dissociated population, the image in Fig. 5C was captured as a single frame with both UV and bright-field illumination. It can be appreciated that the GFP+ cells represent a small minority of total cells, consistent with the presence or 1–5% TH+ cells that is normally present in midbrain cultures. Within 18 h of plating, labeled cells have begun to elaborate processes, whose network continued to grow over 4 days (experiment terminated). One set of cultures was fixed after 2 days and examined for TH/GFP co-expression by dual label immunofluorescence microscopy. A single field is shown labeled for TH (Fig. 5D, rhodamine), or GFP (Fig. 5E,

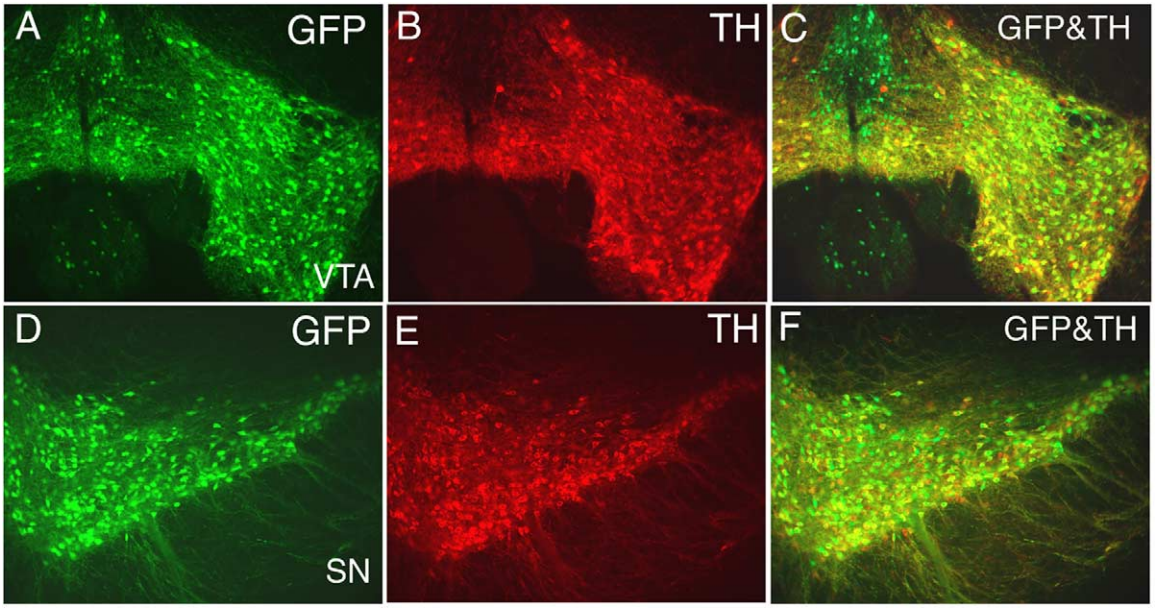

Fig. 4. Immunofluorescence localization of endogenous TH and the GFP reporter in the ventral tegmental area (VTA; A–C) and substantia nigra (SN; D–F) of an adult transgenic mouse (founder no.7827 line). The GFP protein was localized with a green FITC conjugate (A, D), while a rhodamine signal was used to detect TH in red (B, E). The merged images revealed extensive co-localization as indicated by the yellow signal (C, F). Note the exclusive expression of GFP (no overlap with TH) in the midline region of the VTA (C).

FITC). The yellow signal in the merged image (Fig. 5F) indicates near complete correlation between the two, although, a few individual GFP+ singly labeled cells can be detected.

3.4. Sequence analysis of the human TH promoter

The proximal hTH promoter and the complex RNA splicing that generates isoforms of the enzyme have been previously reported [21,38,61]. We therefore concentrated our sequencing effort on the ~11 kb of promoter distal to the SacII site at −215. This region was sequenced with 20-fold average coverage while the 3′ remainder of the genomic clone (~2.3 kb) received 4-fold minimum coverage. The complete sequence is submitted to Genbank, accession number AF536811. We present here only selected findings.

A map of the 13,329 bp determined hTH promoter sequence is given in Fig. 6.

For clarity, and to avoid overlap of the symbols, features are divided into several panels. The symbols are centered on the correct position but are larger than scale. The positions of all sites depicted in Fig. 6 are given in the online supplementary data. Fig. 6E–G are discussed below in Section 3.5. The top line (Fig. 6A) indicates position (in kb) with respect to the transcription start site. Fig. 6B shows the four identified bicoid binding elements (BBE I–IV) in crosshatch, while the initial exons are illustrated as solid black bars. The unlabeled exon is included in a minor population of TH transcripts [31,61]. Because the BBE site specifically binds the transcription factor Pitx3, which is thought to be important in specification of the dopaminergic phenotype [83], we designed double-stranded oligonucleotides from BBE I–IV to test bicoid binding specificity by electrophoretic mobility shift assay (EMSA) and supershift (see Fig. 7). Fig. 6C shows the location of repeat elements [84]. The TTCA tetranucleotide repeat within the first intron has been previously noted to be polymorphic, and a role in regulation of gene expression has been demonstrated in vitro [51]. The positions of response elements for several transcription factors are depicted in Fig. 6D. The quality of match to the published sequence motifs was 97% for the AP-3 site, 100% for EGR, AP-1, 2, and 4, GLI, and CRE. No exact match to the 21 bp consensus of the neuron restrictive silencer element (NRSE) could be located [46,78]. However two sites with 70% homology to the NRSE are depicted as smaller open boxes (see Fig. 8).

3.5. Regions of homology among rat, mouse, and human TH promoters

We compared the sequence of the human TH promoter to available corresponding sequence from mouse (locus: AP003184; [gi16303287]) and rat (locus: AF069036; [gi:5724776]) in an effort to detect evolutionarily conserved features outside the proximal promoter which may play a significant role in developmental and short-term regulation of TH expression. Five areas of homology (conserved regions, CRs I–V) upstream of −2 kb are

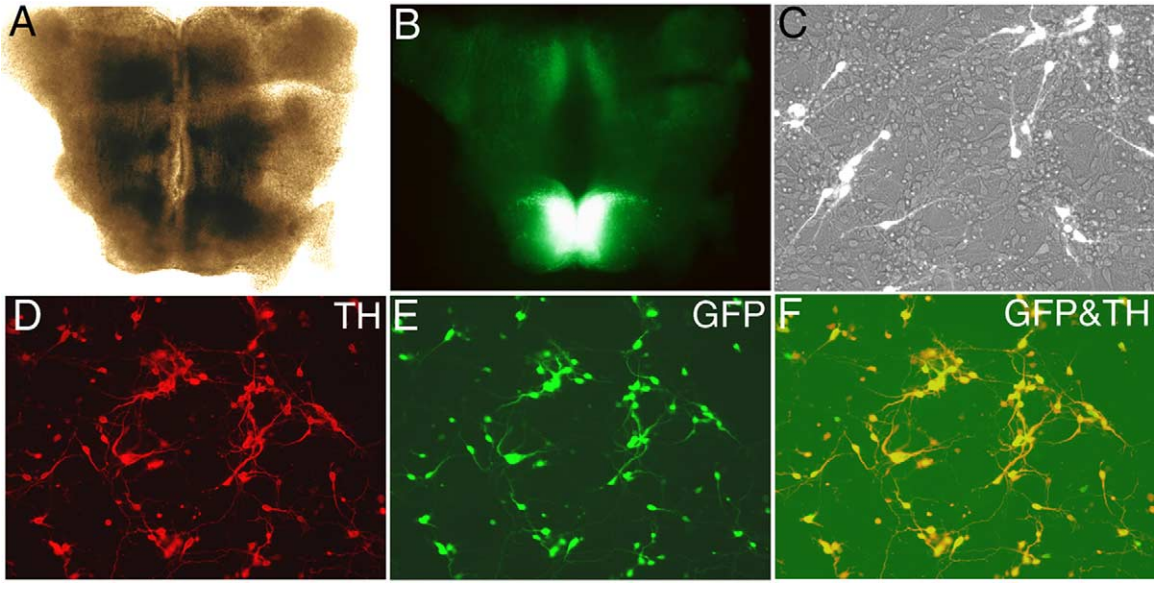

Fig. 5. (A–C) Demonstration of function of hTH transgene in living tissue and cells. (D–F) Demonstration of the specificity of reporter (EGFP) targeting to TH+ cells. (A) ×40 bright-field image of a freshly dissected E13 transgenic mouse mesencephalon is shown. The neural tube has been cut longitudinally along the dorsal surface to reveal the underlying dorsal aspect of the tegmentum. The rostral portion, including the mesencephalic flexure, is at the bottom of the frame. (B) This is the same field as A, but imaged for EGFP fluorescence. The brightness of the signal within the mesencephalic flexure approximates that seen by eye. For this image, the background signal has been enhanced so that the shape of the entire dissected specimen can be directly compared with the bright-field image in A. (C) The tissue sample of A,B was dissociated with trypsin into a single-cell suspension, plated, and incubated overnight. This ×200 image was captured as a single frame by turning up the white light in conjunction with the UV. It is a remarkable demonstration of the high signal/noise ratio of this transgenic. Most fluorescent microscopy is performed with the white light off and the room darkened to avoid overwhelming the fluorescent image. The bright-field was turned up to clearly visualize the EGFP- 'background' cells. TH+ neurons are only a small percentage of mesencephalic cells. This image illustrates a corresponding low ratio of EGFP+ to EGFP− (non-fluorescent) cells. The cultures illustrated in C were maintained until 48 h after plating, fixed, and processed for double-label fluorescent immunocytochemistry (D–F). (D) ×200 image of cultures stained for TH with rhodamine (red) conjugate. Note that the network of processes is more developed at 48 h than 12 h after plating (compare to C). (E) This is the same field as D imaged for the green signal of the anti-EGFP FITC conjugate. It can be appreciated that the same individual TH+ cells and processes (D) also express GFP (E). (F) This image is made by merging D (red), and E (green). The yellow color of the bright cells confirms the co-expression of TH and the reporter, EGFP.

shown in Fig. 6, panels E (human), F (mouse), and G (rat) using solid black rectangles. These were notable for several reasons. First, it can be appreciated from the figure that the overall topology and linear order of these regions has been preserved. Second, the degree of homology is near 100% over sites with a length of 18 to 60 bp. Third, across these three species the orientation of individual sites is the same (same strand). Finally, a BLAST search [3] of the Genbank database using these sequences returned exclusively tyrosine hydroxylase genes. A closer examination of CRs I–V is shown in Fig. 9. Bases identical to the human sequence are indicated by an asterisk (*). The response elements identified by the TFD [19], Transfac [91], or IMD [68] signal scan programs are located by name and a line above or below the corresponding sequence. Two sites correspond to unidentified factors. The site marked 'a' in CR I corresponds to an element in the whey acidic protein promoter which binds nuclear proteins in an extract of lactating mammary glands [47]. In region V, 'b' matches an octamer motif within the HLA-DQ beta promoter [54]. The two sites labeled 'NurRE 1/2 site' are portrayed in Fig. 10 and discussed below in Section 4.2.1.

3.6. LIM family members

Below, Fig. 11, is the alignment of the FLAT element of the rat insulin I gene enhancer [18] to human, mouse, and rat TH promoters. Also shown, because of homology, is the Emx2 site from the Wnt-1 3′ enhancer. Emx2 restricts the expression of Wnt-1 in the developing brain [28] and, like Otx1 and -2 is expressed in embryonic mesencephalon [27,80].

4. Discussion

4.1. Expression of TH and GFP in transgenic mice

We have demonstrated that 11 kb of the human tyrosine hydroxylase promoter is capable of driving reporter gene expression in TH+ neurons of the embryonic and adult CNS. In particular, the strong GFP fluorescence in the embryonic substantia nigra permits ready identification of TH+ cells suitable for transplant. This implies that this construct may be clinically useful for deriving material for

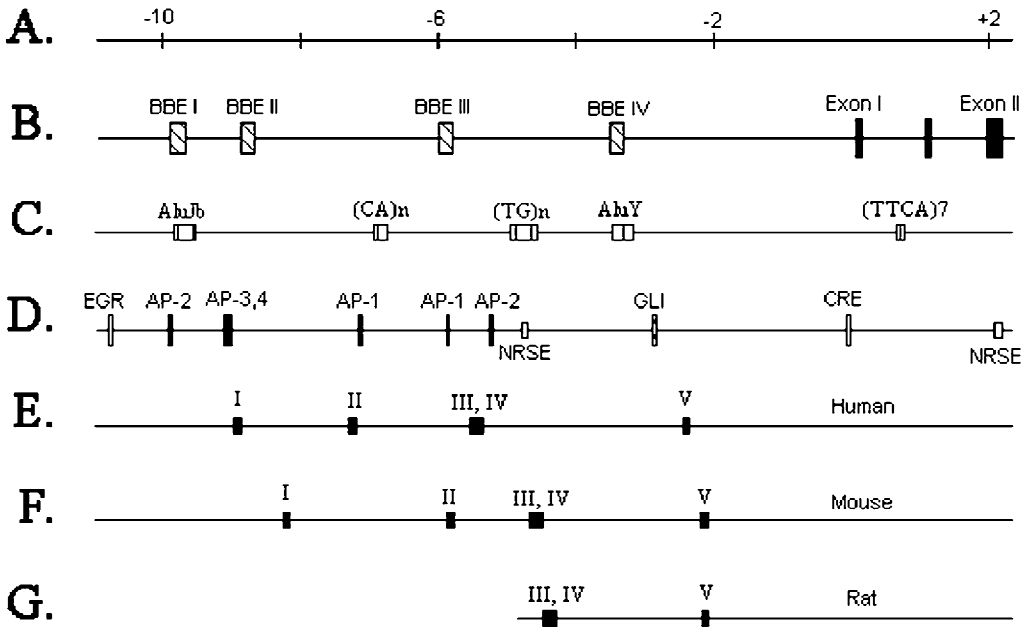

Fig. 6. Overview of TH promoter structure. (A) This scale, in kilobases, marks the nucleotide positions with respect to the start of transcription (Exon I in B). (B) The 5′ end of the human TH gene is represented, including 11 kb of promoter and the first Exons. BBEs I–IV are recognition sites for the transcription factor bicoid, associated with regulation of the TH gene by Pitx3 [40]. (C) Repeat sequences within the human TH promoter including a previously described [1,50] tetranucleotide (TTCA) repeat in the first intron. (D) The locations of select potential response elements are shown. These include the early response factors EGR and AP-1–4; a single cyclic-AMP response element (CRE), a Gli site which may mediate Shh action, and two 70% matches to the neuron restrictive silencer element, NRSE [78]. (E) Human TH promoter. Solid black squares denote conserved regions I–V. (F) Mouse TH promoter (Genbank accession locus: AP003184; [gi:16303287]). Solid black squares denote conserved regions I–V. (G) Rat TH promoter (Genbank accession locus: AF069036; [gi:5724776]). Solid black squares denote conserved regions III–V.

cell replacement therapy of Parkinson's disease. Protocols have been reported for the derivation of human dopaminergic cells from embryonic stem (ES) cells and expandable human neural precursors [12,35,36,42,73,95]. However, the final yield of TH+ cells represents only a fraction of total cells. The hTH-GFP reporter construct reported here may thus be of use for final purification of therapeutic cells from such renewable sources. Indeed, using a similar rat TH-GFP transgenic mouse, Sawamoto and colleagues [76] have reported that dopamine neurons can be greatly enriched (60% GFP+) for transplantation by FAC sorting.

Although the majority of reporter expression in our transgenic mice was region appropriate, there were areas of ectopic expression in the CNS. Many of these showed parallels to those described in other models using shorter fragments of rat or human TH promoter. In those studies, authors have consistently noted expression in the amygdala, interpeduncular nucleus, nucleus accumbens, and septum [45,53,58,89]. Some of this ectopic expression may be ascribed to positional effects dependent on the site of integration of the transgene. Long range chromosome conformation may affect the more than 25 kb native TH locus in a manner not replicated by shorter constructs. However, the consistency of these observations suggest alternative explanations. One possibility is that the 11 kb hTH promoter lacks repressor elements which may lie outside this region. Another possibility is that gene dosage could titrate out repressors that define the anatomical limits of normal expression. Finally, it is unlikely that the simple, intron-less reporter gene is subject to the same regulation as the native 12 kilobase TH transcript (see below).

The surprising failure of TH promoter-driven expression within the adrenal medulla and locus ceruleus observed here has been noted before in experiments utilizing a 5.3 kb fragment of the rat TH promoter [89]. In that case it was possible to ascribe this anomaly to instability of the reporter in those tissues. The *lacZ* reporter construct utilized contained a polyadenylation signal from SV40, but not a splice site, similar to our GFP reporter (EGFP-1, Clontech). Thus, one explanation for these findings may be that there is relatively rapid turnover of reporter mRNA in the adrenal and locus ceruleus [2].

4.2. Potential response elements within the human TH promoter

With the exception of the bicoid element, we have not yet determined the functional significance (EMSA assays) of response elements in the upstream portion of the TH promoter. Nonetheless, provided below is a summary of *cis*-elements which bind transcription factors of known importance in TH expression (NR4A2, BBE, HNF-3,

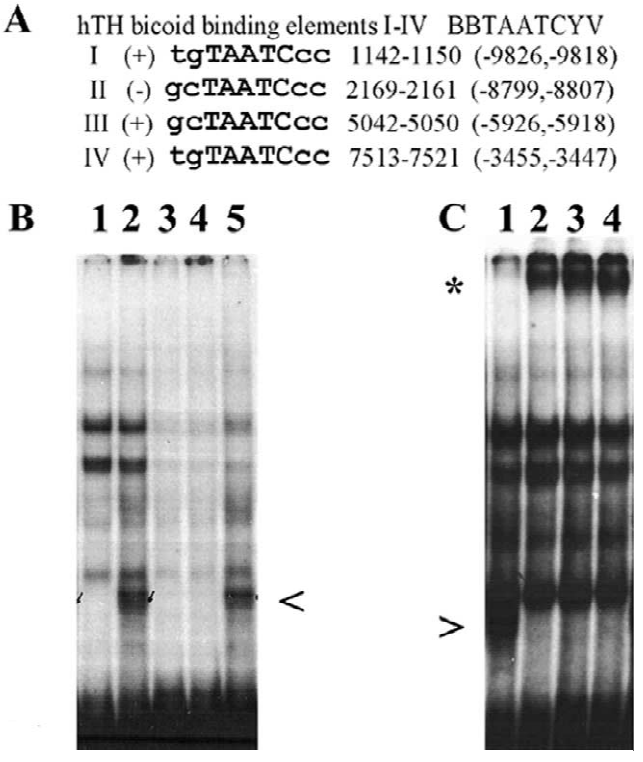

Fig. 7. Human tyrosine hydroxylase promoter bicoid binding elements (BBEs). (A) BBE I–IV sequences all match 9 bp consensus BBTAATCYV (B=C, G or T; Y=C or T; V=A, C or G). The orientation of BBE II has been reversed '(−)' in this figure to emphasize alignment with other sites. Numbers indicate positions of sites within 13,329 bp gene sequence, or, in parentheses, with respect to transcript start. (B) EMSA employing 25 bp ^{32}P-labeled double-stranded oligonucleotide corresponding to BBE II (see Materials and methods, results were comparable with BBE I, III, and IV). Additions, Lane 1: Nuclear protein extract from COS-7 cells transiently transfected with GFP-expressing plasmid. Lane 2: Nuclear protein extract from COS-7 cells transiently transfected with plasmid expressing murine goosecoid-like (gscl) homeodomain protein. Arrowhead (<) indicates bicoid/gscl-specific complex. Lanes 3 and 4: As in 2, but with competition by 25 (lane 3) and 12.5-fold (lane 4) molar excess of unlabeled BBE II oligonucleotide. Lane 5: As in 2 but with competition by 50-fold molar excess of AP-1 oligonucleotide. (C) Supershift assay, performed as in B, lane 2 with the following additions: Lane 1: 3 μl of rabbit preimmune serum; Lanes 2, 3 and 4: 1, 2 and 3 μl of affinity purified anti-gscl rabbit antiserum. Arrowhead (>) indicates bicoid/gscl-specific complex, while the supershifted band is indicated by an asterisk (*).

LIM) and those that are potentially new with an, as yet, undefined role (HoxA4/5) in TH regulation.

4.2.1. NR4A2 (Nurr1)

The NGFI-B family of orphan nuclear receptors derives its name from the founding member which was isolated from PC12 cells as a transcript induced by nerve growth factor [52]. The NGFI-Bbeta subgroup includes the human gene formerly known as NOT or NURR1, and murine Nurr1, now officially designated by the cross-species nomenclature NR4A2 [4,20]. Although in tissue culture, both TH and NR4A2 are induced by cAMP analogs or forskolin [57], recently, the dependence of TH induction on NR4A2 has been seriously questioned [88]. Nevertheless, NR4A2 has been implicated in the expression of TH by genetic knockout studies in mice which showed loss of TH expression in the substantiae nigrae of NR4A2 null animals [6,75]. The canonical NR4A2 9 bp monomer binding site (NBRE) AAAAGGTCA was defined by genetic assay in yeast [90]. However, in cultured mammalian cells the sequences implicated in NR4A2 action may be quite different as the protein can also bind DNA as a homodimer, or heterodimer complexed with the retinoid X receptor RXR [63]. In the POMC promoter the protein binds as a dimer to a palindromic site (NurRE) only loosely related to NBRE [66]. A model has been proposed for NR4A2 regulation of the human dopamine transporter (DAT) gene in which it acts as a transcriptional coactivator, without direct DNA contact [72]. No NBRE or complete NurRE is found within the human TH gene from −11 to +2 kb. However, 1/2 sites of the NurRE palindrome are present at two sites, which are exactly conserved between rodent and human. In Figs. 9 and 10, these are each labeled 'NurRE 1/2 site' where they occur in CRs I and IV.

4.2.2. Bicoid binding elements (BBE)

Disruption of the *Drosophila melanogaster bicoid* gene prevents organization of anterior structures leaving animals with two caudal ends [13]. Its unique DNA binding sequence specificity is shared by the mammalian homeoprotein families of the paired-like class: Otx and Pitx. The important role of these transcription factors in determining cell type-specific gene expression in the pituitary gland has been recently reviewed [14]. That Pitx3 may play a role in specification of the dopaminergic phenotype was suggested originally by its coexpression in TH+ cells of the mesencephalon [83]. Subsequently, Pitx3 regulation of the TH gene has been ascribed to a single high-affinity binding site (bicoid binding element, BBE) within the rat TH promoter [40]. That four sites were identified in the human TH promoter (Figs. 6 and 7 above)

```
I       G*A*TG*C*********_*A**   6220-6240    (-4748,-4728)
NRSE    TTCAGCA_CCACGGACAGCGCC
II      ****_**G*TCG**********   13104-13125  (+2137,+2158)
```

Fig. 8. Homology of neuron restrictive silencer element (NRSE) to sites within the human tyrosine hydroxylase gene. Identical bases are designated with an asterisk (*). A single base gap has been introduced into each sequence (_) to maximize alignment. Numbers indicate the position within the 13,329 bp sequence, or (in parentheses) with respect to the transcription start site.

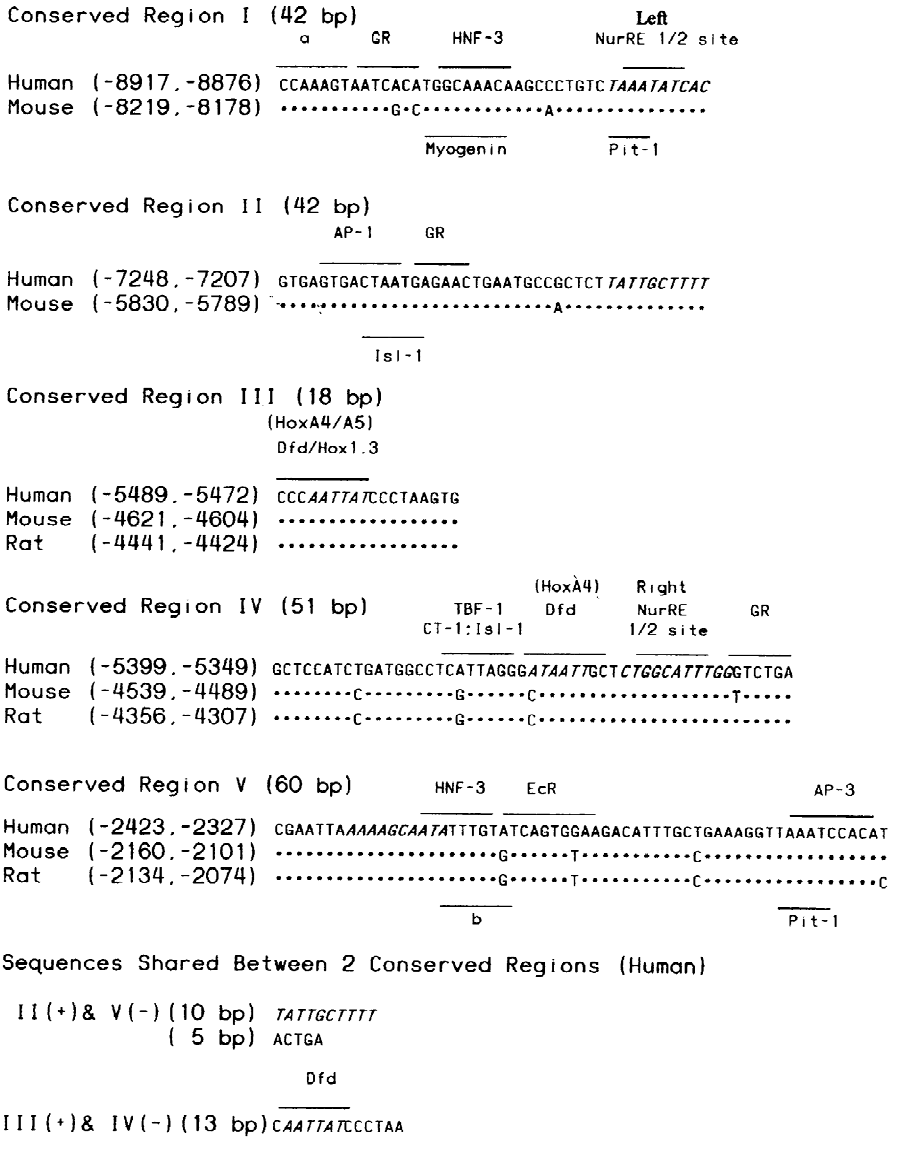

Fig. 9. Analysis of conserved regions I through V (CR I–V). Shown are the sequences of the five conserved regions. Available sequence for the rat includes CRs III–V, but not I and II. Homologous regions are aligned with positions in parentheses indicating their location with respect to the start of transcription of the respective TH genes. Exact matches to the human sequence are indicated by '*', and only the discordant bases are shown for mouse and rat. Above or below the sequence are labeled with the corresponding transcription factors identified in a search of the available databases [19,69,92]. Sites marked 'a' (region I) and 'b' (region V) correspond to unidentified factors described in Refs. [47] and [54], respectively. CR I: 'Left NurRE half-site' is overlined and and is homologous to the NuRE site of the pro-opiomelanocortin (POMC) promoter [67]. The entire 10 bp match to the POMC NurRE is in bold italics (see Fig. 10). CR II: The 3′ 10 bases are italicized, as is their reverse complement in CR V. The available databases do not contain a transcription factor which recognizes this motif. At the bottom of the Fig. ('sequences shared') this homology and that between regions III and IV are emphasized. CR III: The HoxA4/Dfd site is shared with CR IV. Within the Dfd element the italicized six base sequence AATTAT is homologous to the described Lmx1b recognition site [56]. CR IV: The 'Right NurRE half site' [67] is overlined, and the 11 base match to the POMC gene NurRE is in italics (see Fig. 10). CR V: The italicized 10 bases are the exact reverse complement of those at the 3′ end of CR II.

may, in part, explain the high levels of expression observed with the hTH-11 kb-EGFP construct.

4.2.3. LIM family of homeodomain proteins

The LIM homeodomain proteins derive their name from three founding family members—*C. elegans* lin-11, rat Isl-1, and *C. elegans* mec-3. They are characterized by a unique homeodomain as well as two zinc-finger structures which mediate binding to protein co-factors (recently reviewed in [24]). The presence of two Isl-1 sites within conserved regions II and IV of the TH promoter (Fig. 9, above) may underlie the expression of TH in embryonic pancreatic stem cells [87] and in adult beta and delta cells [64]. Many members of the LIM family have been implicated in the determination of neuronal cell fate, and loss of a population of mesencephalic dopaminergic cells has been demonstrated in Lmx1b null mice [82]. Also, this sequence has previously been shown on gel shift to bind

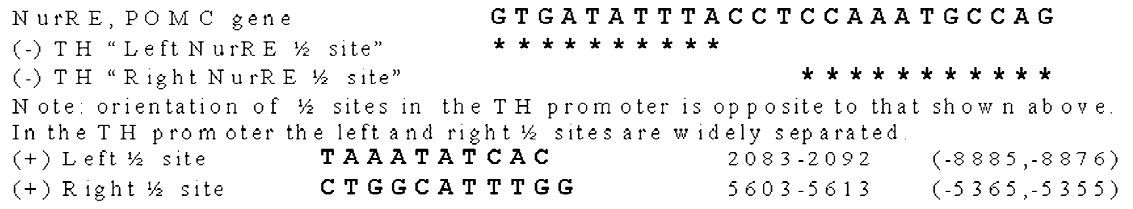

Fig. 10. Alignment of TH promoter sequences within conserved regions I and IV to the NurRE site originally described within the pro-opiomelanocortin (POMC) gene [67]. Matching bases are indicated by an asterisk (*). Note that the 'left' and 'right' designations refer to the positions of imperfectly symmetrical 'NurRE 1/2 sites' within the complete POMC NurRE. Numbers indicate base positions within the 13,329 bp of the hTH gene, or, in parentheses, with respect to the hTH transcription start site.

Arix/Phox2a, a paired-like homeodomain protein which plays a key role in noradrenergic specification [94]. Within the human TH promoter there is no exact match to the TAATTA hexamer Lmx1b binding site [56]. Partial matches (5/6) are included in the AATTAT motif found at three locations. One of these is at position −6847, the other two overlap the Dfd (*D. melanogaster* deformed) sites in conserved CRs III and IV (Fig. 9). As this short sequence might be expected to occur by chance, we sought alternative potential Limx1b binding sites. Several LIM family members, including Limx1a [18] and Isl-1 [33] have been isolated as transcription factors which bind to the FLAT element of the insulin enhancer ('Isl' stands for islets of Langerhans). Alignment of the rat insulin gene FLAT element to TH promoters and the Emx2 site of the Wnt 3′ enhancer (Fig. 11) demonstrates perfect homology over 7–11 bases.

4.2.4. Sonic hedgehog, Gli, and BMP

Sonic hedgehog, the vertebrate homolog of *Drosophila hedgehog* gene [69], mediates the induction of midbrain DA neurons by the floor plate [26]. Human Gli (homolog of *Drosophila melanogaster* cubitus interruptus) was originally discovered as a gene amplified in gliomas, and subsequently shown to be a zinc-finger transcription factor and oncogene capable of transforming primary cells in conjunction with adenovirus E1A protein [37,71]. Gli1 transcription is stimulated by Sonic hedgehog (Shh), and exogenous Gli1 mimics the floor-plate inducing activity of Shh [41]. In turn, Gli proteins bind to the sequence TGGGTGGTC to modulate transcription of target genes. Multiple copies of this sequence are found in the BMP-4 and BMP-7 gene promoters, which, in cultured cells, are induced by transfected Gli1, or Gli3 [34]. As Gli3 and Shh are antagonistic [8,44], an implied signal pathway is Shh> Gli1/Gli3>BMP. We have recently reported the induction of TH by BMP-2, -4,or -6 in ex vivo cultures of embryonic striatum [85]. A single perfect copy of the nonamer Gli response element is located within the human TH promoter at position −2857, −2849. The available sequence for the rat and mouse TH promoters reveals imperfect matches (7/9) at positions −2179, −2171; and −9833, −9825; respectively.

4.2.5. Conserved regions I–V

Previous studies of the rat and human TH promoters in transgenic mice have established that sequences distal to −2000 are important for proper levels and anatomical distribution of reporter expression [45,58,77,92] but their identity is unknown. The evolutionarily conserved regions between −2000 and −9000 are prime candidates for those regulatory domains. Isl-1 and NR4A2 sites within these regions have been discussed above.

4.2.6. HNF-3beta

HNF-3beta is a winged helix transcription factor related to *Drosophila* forkhead. The potential importance of this class of transcription factors has been recently emphasized by studies of *C. elegans* in which an HNF-3beta homolog, pha-4, was found responsible for determining pharyngeal organ identity [25]. In addition, it has been considered a marker of early floor plate [70], and is induced by Shh [26]. Since Shh is thought to be one of the two necessary in vivo signals needed to induce TH in developing

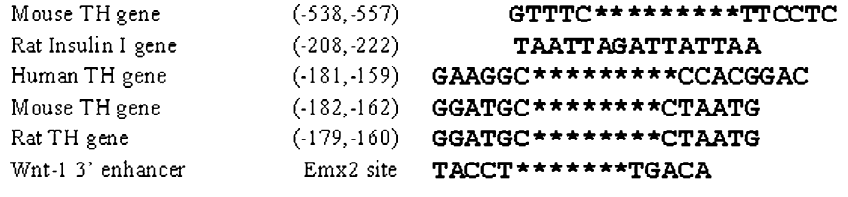

Fig. 11. Alignment of FLAT element of rat insulin promoter (second line), and Emx2 site of Wnt-1 3′ enhancer to sites within the mouse, human and rat tyrosine hydroxylase promoters. Numbers in parentheses indicate the positions of these sites with respect to the transcription start of the corresponding gene. Identical residues are designated by an asterisk (*). No gaps have been introduced. Human sequence is from Genbank AF536811 (this manuscript), mouse from AP003184, rat from AF069036, and Emx2 site from Ref. [28].

midbrain neurons [26,27], the two copies of HNF-3β found in CRs I and V (Fig. 9), may be critically involved in mediating this expression.

4.2.7. Deformed (Dfd) and Sex combs reduced (Scr)

The *Drosophila* homeobox genes Dfd and Scr determine head segment identity and have persistent expression within the adult CNS [43,49,50]. Vertebrate homologs include HoxA-D4 (Dfd) and HoxA-C5 (Scr). Most of these are expressed in embryonic murine CNS, and all are inducible by retinoic acid (RA) in neurogenic teratocarcinoma or neuroblastoma cell lines [65,81]. An important feature shared by Dfd, HoxA4, and HoxA5 is the ability to positively autoregulate their expression [5,39,62]. Such a positive feedback loop is an ideal mechanism for the establishment and maintenance of a particular phenotype. Potential response elements for HoxA4 are also designated 'Dfd' in Fig. 9, conserved regions III and IV. A HoxA5 response element in conserved region III is also denoted by 'Hox1.3', an older name for murine HoxA5 [79].

Known transcription factors within CRs I–V, those present in public databases, are indicated in Fig. 9 and have been discussed. Unassigned conserved sequences may simply represent evolutionary relics. However, the degree of conservation across three mammalian species suggests a functional role for sequences for which a known factor cannot be assigned. Possibilities include: constraints on DNA conformation, correct spacing of interacting transcription factors, or binding of a previously unidentified factor. The last possibility may be most sensitively addressed by use of electrophoretic mobility shift assay with a nuclear protein extract derived from FACS-purified TH+ neurons.

These data inform future studies of non-structural mutations of the human tyrosine hydroxylase locus which may be of clinical import. A major caveat regarding the present work is that all these in vitro data await confirmation by biological experiments. The availability of our murine transgenic model enables the design of future genetic, biochemical, and functional studies to examine this issue.

Acknowledgements

For assistance with olfactory bulb anatomy, we are grateful for the expertise of Professor Harriet Baker, Department of Neurology and Neuroscience, Weill Medical College of Cornell University at the W. M. Burke Medical Research Institute, 785 Mamaroneck Avenue, White Plains, NY 10605. For the goosecoid-like cDNA and antiserum we thank Professor J.A. Epstein, University of Pennsylvania School of Medicine, Philadelphia, PA. This work was supported by NS24204 and NS32519 to LI.

Appendix A. Online Supplementary Material

This refers to 13,329 bp assembled sequence from lambda genomic clone pMAK 221-21 (AF536811). It is the unmodified hTH gene sequence and includes 5′ Sal I g×tcgac (153–158)

5′ end is gatc of Sau3A. Note that vector sequences upstream of Sau3A cloning site have been eliminated.

3′ end is gatc of Sau3A. Note that vector sequences downstream of this site have been eliminated

Extends −10,967 to +2362 with respect to transcription start site.

Transcription start (**agacggagcccgg**) 10,968

Translation start (**ATGCCCACCC**) 10,997

For standard numeration with respect to transcription start:

(+): n−10967 (for positions 3′ to transcription start; $n=10{,}968$)

(−): n−10968 (for positions 5′ to transcription start; $n=10{,}967$)

Exons [21,38,61]

I1 10968– 11086 (transcription starts 10,968)

I2 10968–11098

I3 12013–12093

II 12913–13134

EGR-1 and -2 [10], are zinc finger transcription factors involved in early growth response and have been alternatively designated Zif268, Krox-24, and Krox-20. Only the 5′-most EGR response element achieved a high matrix score and is illustrated in Fig. 6D. A space is added to the EGR-1 sequence below to accommodate the 3′-most site which differs from the others by addition of a 'T.'

Human TH (+)	**CT****_****	207–216
EGR-1	**GCGTGG_GGCG**	
Human TH (+)	***A****_****	2363–2372
Human TH (+)	******_***C**	2788–2797
Human TH (−)	******_***A**	6893–6884
Human TH (+)	***AT***_****	8384–8393
Human TH (−)	******T****	11,021–11,011

hTH bicoid binding elements BBEs I–IV BBTAATCYV [23]

I (+)	**tgTAATCcc**	1142–1150	(−9826, −9818)
II (−)	**gcTAATCcc**	2169–2161	(−8799, −8807)
III (+)	**gcTAATCcc**	5042–5050	(−5926, −5918)
IV (+)	**tgTAATCcc**	7513–7521	(−3455, −3447)

Human TH gene (AF536811) conserved regions, CRs I–V

CR I	2051–2092	(−8917, −8876)	42 bp
CR II	3720–3761	(−7248, −7207)	42 bp
CR III	5479–5496	(−5489, −5472)	18 bp
CR IV	5569–5619	(−5399, −5349)	51 bp
CR V	8545–8641	(−2423, −2327)	60 bp

Mouse TH gene (AP003184) conserved regions, CRs I–V

CR I	65,892–65,933	(−8219, − 8178)	42 bp
CR II	68,281–68,322	(−5830, − 5789)	42 bp
CR III	69,490–69,507	(−4621, − 4604)	18 bp
CR IV	69,572–69,622	(−4539, − 4489)	51 bp
CR V	71,951–72,010	(−2160, − 2101)	60 bp

Rat TH gene (AF069036) conserved regions, CRs III–V

CR III	78–95	(−4441, − 4424)	18 bp
CR IV	163–212	(−4356, − 4307)	51 bp
CR V	2385–2445	(−2134, − 2074)	60 bp

Gli (−) **GACCACCCA** 8119–8111 (−2849,−2857) [41,74,86]

AP-1 [15]

(+)	**TGAGTCA**	3834–3840	(−7134, − 7128)
(+)	**TGAGTCA**	5102–5108	(−5866, − 5860)
(+)	**TGATTCAG**	10,764–10,771	(−204, − 197)

The most 3′ AP-1 site (−204,−197), just upstream of the CRE site (−45,−38), is not included in Fig. 6D because of the low matrix score.

AP-2 [15]

(+)	**CCCCAGGC**	1076–1083	(−9892, − 9885)
(+)	**CCCCAGGC**	5720–5727	(−5248, − 5241)
(+)	**CCCCAGGC**	11,179–11,186	(+ 212, + 219)

AP-3 [15]

(+)	**TGTGGAAA**	1869–1876	(−9099, − 9092)

AP-4 [15]

(−)	**CAGCTGTGG**	1957–1949	(−9011,−9019)
(+)	**CAGCTGTGG**	12,846–12,854	(+1879,+1887)

CRE [54,55]

(+)	**TGACGTCA**	10,923–10,930	(−45, − 38)

NRSE [77]

I	**G*A*TG*C********_*A****	6220–6240	(−4748,−4728)
NRSE	**TTCAGCA_CCACGGACAGCGCC**		
II	**** _ **G*TCG**********	13,104–13,125	(+2137,+2158)

NurRE [65]

NurRE, POMC gene	**GTGATATTTA CCT CCAAATGCCAG**
(−) TH ‘Left NurRE 1/2 site’	**********
(−) TH ‘Right NurRE 1/2 site’	**********

Note: orientation of 1/2 sites in the TH promoter is opposite to that shown above. In the TH promoter the left and right 1/2 sites are widely separated.

(+) Left 1/2 site	**TAAATATCAC**	2083–2092	(−8885,−8876)
(+) Right 1/2 site	**CTGGCATTTGG**	5603–5613	(−5365,−5355)

Homologies among TH promoter, Insulin ‘FLAT’ element [18], and EMX2 response element [28]:

Mouse TH gene	(−538,−557)	**GTTTC*********TTCC**
Rat insulin I gene	(−208,−222)	**TAATTAGATTATTAA**
Human TH gene	(−181,−159)	**GAAGGC*********CCACGGAC**
Mouse TH gene	(−182,−162)	**GGATGC********CTAATGG**
Rat TH gene	(−179,−160)	**GGATGC********CTAATG**
Wnt-1 3′ enhancer	EMX2 site	**TACCT*******TGACA**

REPEATS (http://ftp.genome.washington.edu/cgi-bin/RepeatMasker)

begin	end (left)	repeat	class/family	begin	end
1120	1424 (11,905)	+AluJb	SINE/Alu	1	305
4026	4083 (9246)	+(CA)n	Simple repeat	2	58
4185	4215 (9114)	+(CA)n	Simple repeat	1	31
6003	6185 (7144)	+(TG)n	Simple repeat	1	178
6278	6451 (6878)	+(TG)n	Simple repeat	1	172
7490	7795 (5534)	+AluY	SINE/Alu	1	306
11,667	11,697 (1632)	+(TTCA)n	Simple repeat	1	31
11,942	11,979 (1350)	+C-rich	Low complexity	1	38

References

[1] V. Albanese, N.F. Biguet, H. Kiefer, E. Bayard, J. Mallet, R. Meloni, Quantitative effects on gene silencing by allelic variation at a tetranucleotide microsatellite, Hum. Mol. Genet. 10 (2001) 1785–1792.

[2] J. Alterio, J. Mallet, N.F. Biguet, Multiple complexes involved in tyrosine hydroxylase mRNA stability in rat adrenal medulla, after reserpine stimulation, Mol. Cell. Neurosci. 17 (1) (2001) 179–189.

[3] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, Basic local alignment search tool, J. Mol. Biol. 215 (1990) 403–410.

[4] Anonymous, A unified nomenclature system for the nuclear receptor superfamily, Cell 97 (1999) 161–163.

[5] J. Aubin, M. Lemieux, M. Tremblay, R.R. Behringer, L. Jeannotte, Transcriptional interferences at the Hoxa4/Hoxa5 locus: importance of correct Hoxa5 expression for the proper specification of the axial skeleton, Dev. Dyn. 212 (1998) 141–156.

[6] J.S. Baffi, M. Palkovits, S.O. Castillo, E. Mezey, V.M. Nikodem, Differential expression of tyrosine hydroxylase in catecholaminergic neurons of neonatal wild-type and Nurr1-deficient mice, Neuroscience 93 (1999) 631–642.

[7] S. Buervenich, A. Carmine, M. Arvidsson, F. Xiang, Z. Zhang, O. Sydow, E.G. Jonsson, G.C. Sedvall, S. Leonard, R.G. Ross, R. Freedman, K.V. Chowdari, V.L. Nimgaonkar, T. Perlmann, M. Anvret, L. Olson, NURR1 mutations in cases of schizophrenia and manic-depressive disorder, Am. J. Med. Genet. 96 (2000) 808–813.

[8] D. Buscher, B. Bosse, J. Heymer, U. Ruther, Evidence for genetic control of Sonic hedgehog by Gli3 in mouse limb development, Mech Dev. 62 (1997) 175–182.

[9] M. Chalfie, Green fluorescent protein. [Review] [42 refs], Photochem. Photobiol. 62 (1995) 651–656.

[10] B. Christy, D. Nathans, DNA binding site of the growth factor-inducible protein Zif268, Proc. Natl. Acad. Sci. USA 86 (1989) 8737–8741.

[11] S.P. Craig, V.J. Buckle, A. Lamouroux, J. Mallet, I. Craig, Localization of the human tyrosine hydroxylase gene to 11p15: gene

duplication and evolution of metabolic pathways, Cytogenet. Cell Genet. 42 (1986) 29–32.
[12] M.M. Daadi, S. Weiss, Generation of tyrosine hydroxylase-producing neurons from precursors of the embryonic and adult forebrain, J. Neurosci. 19 (1999) 4484–4497.
[13] W. Driever, C. Nusslein-Volhard, The bicoid protein determines position in the Drosophila embryo in a concentration-dependent manner, Cell 54 (1988) 95–104.
[14] J. Drouin, B. Lamolet, T. Lamonerie, C. Lanctot, J.J. Tremblay, The PTX family of homeodomain transcription factors during pituitary developments. [Review] [16 refs], Mol. Cell. Endocrinol. 140 (1998) 31–36.
[15] S. Faisst, S. Meyer, Compilation of vertebrate-encoded transcription factors, Nucleic Acids Res. 20 (1992) 3–26.
[16] F.H. Gage, Mammalian neural stem cells. [Review] [55 refs], Science 287 (2000) 1433–1438.
[17] N. Galili, S. Nayak, J.A. Epstein, C.A. Buck, Rnf4, a RING protein expressed in the developing nervous and reproductive systems, interacts with gscl, a gene within the DiGeorge critical region, Dev. Dyn. 218 (1) (2000) 102–111.
[18] M.S. German, J. Wang, R.B. Chadwick, W.J. Rutter, Synergistic activation of the insulin gene by a LIM-homeo domain protein and a basic helix–loop–helix protein: building a functional insulin minienhancer complex, Genes Dev. 6 (1992) 2165–2176.
[19] D. Ghosh, TFD: the transcription factors database, Nucleic Acids Res. 20 (Suppl.) (1992) 2091–2093.
[20] V. Giguere, Orphan nuclear receptors: from gene to function. [Review] [639 refs], Endocr. Rev. 20 (1999) 689–725.
[21] B. Grima, A. Lamouroux, C. Boni, J.F. Julien, F. Javoy-Agid, J. Mallet, A single human gene encoding multiple tyrosine hydroxylases with different predicted functional characteristics, Nature 326 (1987) 707–711.
[22] Z. Guo, X. Du, L. Iacovitti, Regulation of tyrosine hydroxylase gene expression during transdifferentiation of striatal neurons: changes in transcription factors binding the AP-1 site, J. Neurosci. 18 (1998) 8163–8174.
[23] S.D. Hanes, R. Brent, A genetic model for interaction of the homeodomain recognition helix with DNA, Science 251 (1991) 426–430.
[24] O. Hobert, H. Westphal, Functions of LIM-homeobox genes. [Review] [78 refs], Trends Genet. 16 (2000) 75–83.
[25] M.A. Horner, S. Quintin, M.E. Domeier, J. Kimble, M. Labouesse, S.E. Mango, pha-4, an HNF-3 homolog, specifies pharyngeal organ identity in Caenorhabditis elegans, Genes Dev. 12 (1998) 1947–1952.
[26] M. Hynes, J.A. Porter, C. Chiang, D. Chang, M. Tessier-Lavigne, P.A. Beachy, A. Rosenthal, Induction of midbrain dopaminergic neurons by Sonic hedgehog, Neuron 15 (1995) 35–44.
[27] M. Hynes, A. Rosenthal, Specification of dopaminergic and serotonergic neurons in the vertebrate CNS. [Review] [85 refs], Curr. Opin. Neurobiol. 9 (1999) 26–36.
[28] N. Iler, D.H. Rowitch, Y. Echelard, A.P. McMahon, C. Abate-Shen, A single homeodomain binding site restricts spatial expression of Wnt-1 in the developing brain, Mech. Dev. 53 (1995) 87–96.
[29] H. Inoue, H. Ishii, H. Alder, E. Snyder, T. Druck, K. Huebner, C.M. Croce, Sequence of the FRA3B common fragile region: implications for the mechanism of FHIT deletion, Proc. Natl. Acad. Sci. USA 94 (1997) 14584–14589.
[30] H. Ishiguro, Y. Okubo, T. Ohtsuki, K. Yamakawa-Kobayashi, T. Arinami, Mutation analysis of the retinoid X receptor beta, nuclear-related receptor 1, and peroxisome proliferator-activated receptor alpha genes in schizophrenia and alcohol dependence: possible haplotype association of nuclear-related receptor 1 gene to alcohol dependence, Am. J. Med. Genet. 114 (2002) 15–23.
[31] N. Kaneda, K. Kobayashi, H. Ichinose, F. Kishi, A. Nakazawa, Y. Kurosawa, K. Fujita, T. Nagatsu, Isolation of a novel cDNA clone for human tyrosine hydroxylase: alternative RNA splicing produces four kinds of mRNA from a single gene, Biochem. Biophys. Res. Commun. 146 (1987) 971–975.
[32] N. Kaneda, T. Sasaoka, K. Kobayashi, K. Kiuchi, I. Nagatsu, Y. Kurosawa, K. Fujita, M. Yojoyama, T. Nomura, M. Katsuki, T. Nagatsu, Tissue-specific and high-level expression of the human tyrosine hydroxylase gene in transgenic mice, Neuron 6 (1991) 583–594.
[33] O. Karlsson, S. Thor, T. Norberg, H. Ohlsson, T. Edlund, Insulin gene enhancer binding protein Isl-1 is a member of a novel class of proteins containing both a homeo- and a Cys–His domain, Nature 344 (1990) 879–882.
[34] S. Kawai, T. Sugiura, Characterization of human bone morphogenetic protein (BMP)-4 and -7 gene promoters: activation of BMP promoters by Gli, a sonic hedgehog mediator, Bone 29 (1) (2001) 54–61.
[35] H. Kawasaki, K. Mizuseki, S. Nishikawa, S. Kaneko, Y. Kuwana, S. Nakanishi, S.I. Nishikawa, Y. Sasai, Induction of midbrain dopaminergic neurons from ES cells by stromal cell-derived inducing activity, Neuron 28 (1) (2000) 31–40.
[36] H. Kawasaki, H. Suernori, K. Mizuseki, K. Watanabe, F. Urano, H. Ichinose, M. Haruta, M. Takahashi, K. Yoshikawa, S.I. Nishikawa, N. Nakatsuji, Y. Sasai, Generation of dopaminergic neurons and pigmented epithelia from primate ES cells by stromal cell-derived inducing activity, Proc. Natl. Acad. Sci. USA 99 (2002) 1580–1585.
[37] K.W. Kinzler, S.H. Bigner, D.D. Bigner, J.M. Trent, M.L. Law, S.J. O'Brien, A.J. Wong, B. Vogelstein, Identification of an amplified, highly expressed gene in a human glioma, Science 236 (1987) 70–73.
[38] K. Kobayashi, N. Kaneda, H. Ichinose, F. Kishi, A. Nakazawa, Y. Kurosawa, K. Fujita, T. Nagatsu, Structure of the human tyrosine hydroxylase gene: alternative splicing from a single gene accounts for generation of four mRNA types, J. Biochem. 103 (1988) 907–912.
[39] M.A. Kuziora, W. McGinnis, Autoregulation of a Drosophila homeotic selector gene, Cell 55 (1988) 477–485.
[40] M. Lebel, Y. Gauthier, A. Moreau, J. Drouin, Pitx3 activates mouse tyrosine hydroxylase promoter via a high-affinity binding site, J. Neurochem. 77 (2001) 558–567.
[41] J. Lee, K.A. Platt, P. Censullo, A. Ruiz i Altaba, Gli1 is a target of Sonic hedgehog that induces ventral neural tube development, Development 124 (1997) 2537–2552.
[42] S.H. Lee, N. Lumelsky, L. Studer, J.M. Auerbach, R.D. McKay, Efficient generation of midbrain and hindbrain neurons from mouse embryonic stem cells, Nat. Biotechnol. 18 (6) (2000) 675–679.
[43] P.K. LeMotte, A. Kuroiwa, L.I. Fessler, W.J. Gehring, The homeotic gene sex combs reduced of Drosophila: gene structure and embryonic expression, EMBO J. 8 (1989) 219–227.
[44] Y. Litingtung, C. Chiang, Specification of ventral neuron types is mediated by an antagonistic interaction between Shh and Gli3, Nat. Neurosci. 3 (2000) 979–985.
[45] J. Liu, J.P. Merlie, R.D. Todd, K.L. O'Malley, Identification of cell type-specific promoter elements associated with the rat tyrosine hydroxylase gene using transgenic founder analysis, Brain Res. Mol. Brain Res.. 50 (1997) 33–42.
[46] P. Lonnerberg, C.J. Schoenherr, D.J. Anderson, C.F. Ibanez, Cell type-specific regulation of choline acetyltransferase gene expression. Role of the neuron-restrictive silencer element and cholinergic-specific enhancer sequences, J. Biol. Chem. 271 (1996) 33358–33365.
[47] H. Lubon, L. Hennighausen, Nuclear proteins from lactating mammary glands bind to the promoter of a milk protein gene, Nucleic Acids Res. 15 (1987) 2103–2121.
[48] N. Lumelsky, O. Blondel, P. Laeng, I. Velasco, R. Ravin, R. McKay, Differentiation of embryonic stem cells to insulin-secreting structures similar to pancreatic islets. [erratum appears in Science 2001 293 (5529):428], Science 292 (2001) 1389–1394.
[49] J.W. Mahaffey, T.C. Kaufman, Distribution of the sex combs

reduced gene products in *Drosophila melanogaster*, Genetics 117 (1987) 51–60.
[50] W. McGinnis, T. Jack, R. Chadwick, M. Regulski, C. Bergson, N. McGinnis, M.A. Kuziora, Establishment and maintenance of position-specific expression of the *Drosophila* homeotic selector gene deformed. [Review] [71 refs], Adv. Genet. 27 (1990) 363–402.
[51] R. Meloni, V. Albanese, P. Ravassard, F. Treilhou, J. Mallet, A tetranucleotide polymorphic microsatellite, located in the first intron of the tyrosine hydroxylase gene, acts as a transcription regulatory element in vitro, Hum. Mol. Genet. 7 (1998) 423–428.
[52] J. Milbrandt, Nerve growth factor induces a gene homologous to the glucocorticoid receptor gene, Neuron 1 (1988) 183–188.
[53] N. Min, T.H. Joh, K.S. Kim, C. Peng, J.H. Son, 5′ upstream DNA sequence of the rat tyrosine hydroxylase gene directs high-level and tissue-specific expression to catecholaminergic neurons in the central nervous system of transgenic mice, Brain Res. Mol. Brain Res. 27 (1994) 281–289.
[54] K. Miwa, J.L. Strominger, The HLA-DQ beta gene upstream region contains an immunoglobulin-like octamer motif that binds cell-type specific nuclear factors, Nucleic Acids Res. 15 (1987) 8057–8067.
[55] M.R. Montminy, L.M. Bilezikjian, Binding of a nuclear protein to the cyclic-AMP response element of the somatostatin gene, Nature 328 (1987) 175–178.
[56] R. Morello, G. Zhou, S.D. Dreyer, S.J. Harvey, Y. Ninomiya, P.S. Thorner, J.H. Miner, W. Cole, A. Winterpacht, B. Zabel, K.C. Oberg, B. Lee, Regulation of glomerular basement membrane collagen expression by LMX1B contributes to renal disease in nail patella syndrome, Nat. Genet. 27 (2001) 205–208.
[57] E.P. Murphy, O.M. Conneely, Neuroendocrine regulation of the hypothalamic pituitary adrenal axis by the nurr1/nur77 subfamily of nuclear receptors, Mol. Endocrinol. 11 (1997) 39–47.
[58] I. Nagatsu, N. Karasawa, K. Yamada, M. Sakai, T. Fujii, T. Takeuchi, R. Arai, K. Kobayashi, T. Nagatsu, Expression of human tyrosine hydroxylase–chloramphenicol acetyltransferase (CAT) fusion gene in the brains of transgenic mice as examined by CAT immunocytochemistry, J. Neural Transm. Gen. Sect. 96 (1994) 85–104.
[59] T. Nagatsu, M. Levitt, S. Udenfriend, Tyrosine hydroxylase, the initial step in norepinephrine biosynthesis, J. Biol. Chem. 239 (1964) 2910–2917.
[60] M. Nehls, M. Messerle, A. Sirulnik, A.J. Smith, T. Boehm, Two large insert vectors, lambda PS and lambda KO, facilitate rapid mapping and targeted disruption of mammalian genes, Biotechniques 17 (1994) 770–775.
[61] K.L. O'Malley, M.J. Anhalt, B.M. Martin, J.R. Kelsoe, S.L. Winfield, E.I. Ginns, Isolation and characterization of the human tyrosine hydroxylase gene: identification of 5′ alternative splice sites responsible for multiple mRNAs, Biochemistry 26 (1987) 2910–2914.
[62] A.I. Packer, D.A. Crotty, V.A. Elwell, D.J. Wolgemuth, Expression of the murine Hoxa4 gene requires both autoregulation and a conserved retinoic acid response element, Development 125 (1998) 1991–1998.
[63] T. Perlmann, L. Jansson, A novel pathway for vitamin A signaling mediated by RXR heterodimerization with NGFI-B and NURR1, Genes Dev. 9 (1995) 769–782.
[64] S. Persson-Sjogren, S. Forsgren, I.B. Taljedal, Tyrosine hydroxylase in mouse pancreatic islet cells, in situ and after syngeneic transplantation to kidney, Histol. Histopathol. 17 (2002) 113–121.
[65] F.A. Peverali, M. D'Esposito, D. Acampora, G. Bunone, M. Negri, A. Faiella, A. Stornaiuolo, M. Pannese, E. Migliaccio, A. Simeone, Expression of HOX homeogenes in human neuroblastoma cell culture lines, Differentiation 45 (1990) 61–69.
[66] A. Philips, S. Lesage, R. Gingras, M.H. Maira, Y. Gauthier, P. Hugo, J. Drouin, Novel dimeric Nur77 signaling mechanism in endocrine and lymphoid cells, Mol. Cell. Biol. 17 (1997) 5946–5951.
[67] J.F. Powell, C. Boni, A. Lamouroux, I.W. Craig, J. Mallet, Assignment of the human tyrosine hydroxylase gene to chromosome 11, FEBS Lett. 175 (1984) 37–40.
[68] D.S. Prestridge, SIGNAL SCAN: a computer program that scans DNA sequences for eukaryotic transcriptional elements, Comput. Appl. Biosci. 7 (1991) 203–206.
[69] R.D. Riddle, R.L. Johnson, E. Laufer, C. Tabin, Sonic hedgehog mediates the polarizing activity of the ZPA, Cell 75 (1993) 1401–1416.
[70] H. Roelink, J.A. Porter, C. Chiang, Y. Tanabe, D.T. Chang, P.A. Beachy, T.M. Jessell, Floor plate and motor neuron induction by different concentrations of the amino-terminal cleavage product of sonic hedgehog autoproteolysis, Cell 81 (1995) 445–455.
[71] J.M. Ruppert, B. Vogelstein, K.W. Kinzler, The zinc finger protein GLI transforms primary cells in cooperation with adenovirus E1A, Mol. Cell. Biol. 11 (1991) 1724–1728.
[72] P. Sacchetti, T.R. Mitchell, J.G. Granneman, M.J. Bannon, Nurr1 enhances transcription of the human dopamine transporter gene through a novel mechanism, J. Neurochem. 76 (2001) 1565–1572.
[73] R. Sanchez-Pernaute, L. Studer, K.S. Bankiewicz, E.O. Major, R.D. McKay, In vitro generation and transplantation of precursor-derived human dopamine neurons, J. Neurosci. Res. 65 (4) (2000) 284–288.
[74] H. Sasaki, C. Hui, M. Nakafuku, H. Kondoh, A binding site for Gli proteins is essential for HNF-3beta floor plate enhancer activity in transgenics and can respond to Shh in vitro, Development 124 (1997) 1313–1322.
[75] O. Saucedo-Cardenas, J.D. Quintana-Hau, W.D. Le, M.P. Smidt, J.J. Cox, F. De Mayo, J.P. Burbach, O.M. Conneely, Nurr1 is essential for the induction of the dopaminergic phenotype and the survival of ventral mesencephalic late dopaminergic precursor neurons, Proc. Natl. Acad. Sci. USA 95 (1998) 4013–4018.
[76] K. Sawamoto, N. Nakao, K. Kobayashi, N. Matsushita, H. Takahashi, K. Kakishita, A. Yamamoto, T. Yoshizaki, T. Terashima, F. Murakami, T. Itakura, H. Okano, Visualization, direct isolation, and transplantation of midbrain dopaminergic neurons, Proc. Natl. Acad. Sci. USA 98 (2001) 6423–6428.
[77] J.J. Schimmel, L. Crews, S. Roffler-Tarlov, D.M. Chikaraishi, 4.5 kb of the rat tyrosine hydroxylase 5 ′ flanking sequence directs tissue specific expression during development and contains consensus sites for multiple transcription factors, Brain Res. Mol. Brain Res. 74 (1–2) (1999) 1–14.
[78] C.J. Schoenherr, A.J. Paquette, D.J. Anderson, Identification of potential target genes for the neuron-restrictive silencer factor, Proc. Natl. Acad. Sci. USA 93 (1996) 9881–9886.
[79] M.P. Scott, Vertebrate homeobox gene nomenclature, Cell 71 (1992) 551–553.
[80] A. Simeone, D. Acampora, M. Gulisano, A. Stornaiuolo, E. Boncinelli, Nested expression domains of four homeobox genes in developing rostral brain. [see comments.], Nature 358 (1992) 687–690.
[81] A. Simeone, D. Acampora, V. Nigro, A. Faiella, M. D'Esposito, A. Stornaiuolo, F. Mavilio, E. Boncinelli, Differential regulation by retinoic acid of the homeobox genes of the four HOX loci in human embryonal carcinoma cells, Mech. Dev. 33 (1991) 215–227.
[82] M.P. Smidt, C.H. Asbreuk, J.J. Cox, H. Chen, R.L. Johnson, J.P. Burbach, A second independent pathway for development of mesencephalic dopaminergic neurons requires Lmx1b, Nat. Neurosci. 3 (2000) 337–341.
[83] M.P. Smidt, H.S. van Schaick, C. Lanctot, J.J. Tremblay, J.J. Cox, A.A. van der Kleij, G. Wolterink, J. Drouin, J.P. Burbach, A homeodomain gene Ptx3 has highly restricted brain expression in mesencephalic dopaminergic neurons, Proc. Natl. Acad. Sci. USA 94 (1997) 13305–13310.
[84] A.F.A. Smit, P. Green. (2002) (Unpublished observations).
[85] N.D. Stull, J.W. Jung, L. Iacovitti, Induction of a dopaminergic phenotype in cultured striatal neurons by bone morphogenetic proteins, Brain Res. Dev. Brain Res. 130 (2001) 91–98.
[86] A. Tanimura, H. Teshima, J. Fujisawa, M. Yoshida, A new regula-

tory element that augments the Tax-dependent enhancer of human T-cell leukemia virus type 1 and cloning of cDNAs encoding its binding proteins, J. Virol. 67 (1993) 5375–5382.
[87] G. Teitelman, S. Alpert, J.M. Polak, A. Martinez, D. Hanahan, Precursor cells of mouse endocrine pancreas coexpress insulin, glucagon and the neuronal proteins tyrosine hydroxylase and neuropeptide Y, but not pancreatic polypeptide, Development 118 (1993) 1031–1039.
[88] N. Tornqvist, E. Hermanson, T. Perlmann, I. Stromberg, Generation of tyrosine hydroxylase-immunoreactive neurons in ventral mesencephalic tissue of Nurr1 deficient mice, Brain Res. Dev. Brain Res. 133 (2002) 37–47.
[89] C. Trocme, C. Sarkis, J.M. Hermel, R. Duchateau, S. Harrison, M. Simonneau, R. Al-Shawi, J. Mallet, CRE and TRE sequences of the rat tyrosine hydroxylase promoter are required for TH basal expression in adult mice but not in the embryo, Eur. J. Neurosci. 10 (1998) 508–521.
[90] T.E. Wilson, T.J. Fahrner, M. Johnston, J. Milbrandt, Identification of the DNA binding site for NGFI-B by genetic selection in yeast, Science 252 (1991) 1296–1300.
[91] E. Wingender, X. Chen, E. Fricke, R. Geffers, R. Hehl, I. Liebich, M. Krull, V. Matys, H. Michael, R. Ohnhauser, M. Pruss, F. Schacherer, S. Thiele, S. Urbach, The TRANSFAC system on gene expression regulation, Nucleic Acids Res. 29 (2001) 281–283.
[92] S.C. Wong, M.A. Moffat, K.L. O'Malley, Sequences distal to the AP1/E box motif are involved in the cell type-specific expression of the rat tyrosine hydroxylase gene, J. Neurochem. 62 (1994) 1691–1697.
[93] M. Yang, N.D. Stull, E.Y. Snyder, M.A. Berk, L. Iacovitti, Neuronal stem cells spontaneously differentiate into dopaminergic neurons after transplantation into the intact or 6-hydroxydopamine lesioned rat, Exp. Neurol. 177 (2002) 50–60.
[94] E. Zelmer, Z. Zhang, D. Greco, J. Rhodes, S. Cassel, E.J. Lewis, A homeodomain protein selectively expressed in noradrenergic tissue regulates transcription of neurotransmitter biosynthetic genes, J. Neurosci. 15 (1995) 8109–8120.
[95] S.C. Zhang, M. Wernig, I.D. Duncan, O. Brustle, J.A. Thomson, In vitro differentiation of transplantable neural precursors from human embryonic stem cells [see comments.], Nat. Biotechnol. 19 (2001) 1129–1133.