



Conserved Upstream Regulatory Regions in Mammalian *Tyrosine Hydroxylase*

Meng Wang¹ · Lilah Fones¹ · John W. Cave^{1,2}

Received: 5 November 2017 / Accepted: 25 January 2018
© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract

Tyrosine hydroxylase (Th) encodes the rate-limiting enzyme in catecholamine biosynthesis, and the regulation of its transcription is critical for the specification and maintenance of catecholaminergic neuron phenotypes. For many genes, regulatory genomic DNA sequences that are upstream of the proximal promoter control expression levels as well as region-specific expression patterns. The regulatory architecture of the genomic DNA upstream of the *Th* proximal promoter, however, is poorly understood. In this study, we examined the 11 kb upstream nucleotide sequence of *Th* from nine mammalian species and identified five highly conserved regions. Using cultured human cells and mouse olfactory bulb tissue, chromatin immunoprecipitation (ChIP) assays show that these conserved regions recruit transcription factors that are established regulators of *Th* transcription (such as NURR1, PITX3, FOXA2, MEIS2, and PAX6). This analysis also identified a conserved binding site for CTCF, and functional studies in cultured human cells and ChIP assays with mouse tissue show that CTCF is a novel regulator of *Th* transcription in the forebrain. Together, the findings in this study provide key insights into the upstream regulatory genomic architecture and regulatory mechanisms controlling mammalian *Th* gene transcription.

Keywords Transcription · Evolution · Genomic · Dopamine · Catecholamine

Introduction

Tyrosine hydroxylase (Th) encodes the rate-limiting enzyme for catecholamine neurotransmitter biosynthesis. Homozygous loss of *Th* in mice is embryonic lethal due to disruption of cardiac and/or cardiovascular development [1, 2]. Heterozygous *Th* mutant mice are viable and fertile with a normal physical appearance, but they have reduced noradrenaline levels in multiple brain regions that are linked to impaired associative and latent learning [3]. In humans, individuals with mutations in the *Th* coding region on both *Th* alleles can develop *Th* deficiency, which encompasses a spectrum of movement disorders that typically first manifests in infants [4].

The spatial organization of catecholaminergic neurons is conserved in the mammalian brain [5], which indicates that

there are evolutionarily conserved regulatory mechanisms controlling the expression of genes required for catecholaminergic phenotypes. Studies examining molecular mechanisms regulating *Th* transcription have concentrated on the human and rodent *Th* proximal promoter region (< 1 kb upstream), and these studies have identified several promoter *cis*-regulatory elements that modulate gene expression [6–10]. Although the proximal promoter is necessary for *Th* expression, it is not sufficient to drive reporter gene expression *in vivo* [11–13] and regulatory regions outside the proximal promoter are required to activate *Th* expression. Studies with the human and rat *Th* locus have shown that the 11 and 9 kb upstream regions, respectively, can drive reporter gene expression in adult catecholaminergic regions with minimal ectopic expression [6, 12, 14, 15]. Despite their importance, a systematic and thorough examination of evolutionary conservation within these upstream sequences is lacking. Previous studies have compared human and rodent upstream sequences, but these studies concentrated on consensus transcription factor binding sites and short conserved motifs, and did not define potential upstream enhancer or repressor territories [6, 8]. Given the conservation in the spatial organization of mammalian catecholaminergic neurons as well as the extensive use of rodents and other mammals to understand the function of

✉ John W. Cave
joc2042@med.cornell.edu

¹ Burke Medical Research Institute, 785 Mamaroneck Ave, White Plains, NY 10605, USA

² The Feil Family Brain and Mind Research Institute of Weill Cornell Medicine, New York, NY 10021, USA

these neurons in humans, there is a high priority on identifying conserved upstream regulatory regions that control *Th* expression in the mammalian nervous system. Moreover, identifying conserved upstream regions will also provide novel insight into the molecular mechanisms controlling *Th* expression necessary for the specification and maintenance of catecholaminergic neuronal phenotypes.

To establish whether there are conserved territories upstream of the mammalian *Th* promoter, this study aligned 11 kb genomic sequences upstream of *Th* from nine mammalian species. The conserved regions identified by this alignment were tested for their ability to recruit established transcription factors that regulate *Th* expression in the midbrain and olfactory bulb (OB), which contain the two largest groups of dopaminergic neurons in the brain.

Materials and Methods

Nucleotide Sequence Alignment

Upstream genomic DNA sequences for *Th* were downloaded from Ensembl (<http://www.ensembl.org>). The species used for the alignments were baboon (*Papio anubis*), cow (*Bos taurus*), dog (*Canis lupus familiaris*), dolphin (*Tursiops truncatus*), human (*Homo sapiens*), mouse (*Mus musculus*), panda (*Ailuropoda melanoleuca*), rat (*Rattus norvegicus*), and vervet (*Chlorocebus sabaesus*). Sequence alignments and visualization were performed using Multi-LAGAN and mVista web-based services (<http://genome.lbl.gov/vista/mvista/submit.shtml>) [16, 17]. All sequence alignments and comparisons were made relative to the human *Th* 11 kb upstream region.

ChIP Assays

For ChIP experiments to detect NURR1 occupancy, human SH-SY5Y cells were grown in 60 mm culture dishes to ~80% confluence and then transfected with p3XFlag-CMV-m*Nurr1* (a gift from Dr. Kaoru Saijo, UC Berkeley) using Lipofectamine LTX (Life Technologies). Twenty-four hours after transfection, cells were washed with PBS and then cross-linked with 1% formaldehyde in PBS at room temperature for 9 min. Fixation was terminated by addition of 125 mM glycine, and cells were washed twice with PBS before being pelleted by centrifugation at 4 °C for 5 min and resuspended with SDS Lysis Buffer (Millipore). The chromatin was sheared with a Bioruptor sonicator (Diagenode) and immunoprecipitation of cross-linked protein-DNA complexes used the Magna ChIP Protein-A/G kit (Millipore) following the manufacturer's instructions. Immunoprecipitation reactions used 4 µg of either mouse anti-FLAG M2 antibody (Sigma) or normal mouse IgG (Santa Cruz Biotechnology). Reverse

cross-linking was done overnight in the presence of proteinase K at 62 °C with shaking.

For ChIP assays to detect binding by PITX3 and FOXA2, lysates from non-transfected SH-SY5Y cells were prepared as described above. Immunoprecipitation reactions used either 2.5 µg of mouse anti-PITX3 antibody (Thermo, 38-2850), 1.5 µg of rabbit anti-FOXA2 antibody (Abcam, ab108422) or an equivalent amount of normal IgG (Santa Cruz Biotech).

For ChIP experiments to detect CTCF, MEIS2, PBX1/2/3, and PAX6 occupancy in vivo, adult C57BL6 mice (aged 2–5 months) of both sexes were used for tissue from the OB, cortex, ventral midbrain, and liver. The tissue was washed with ice cold PBS and then cross-linked with 1% formaldehyde in PBS on ice for 15 min. Fixation was terminated with 125 mM glycine, and the tissue was washed twice with PBS before addition of SDS Lysis Buffer (Millipore). Immunoprecipitation of protein/chromatin complexes was performed as described above with SH-SY5Y cells and used either 6 µg of goat anti-CTCF antibody (Santa Cruz), 10 µg of goat anti-MEIS2 antibody (Santa Cruz Biotech, sc-10600), 2 µg of mouse anti-PBX1/2/3 antibody (Santa Cruz Biotech, sc-28313), 5 µg of rabbit anti-PAX6 antibody (Abcam, ab5790-100), or an equivalent amount of normal IgG (Santa Cruz Biotech).

For ChIP experiments with neurospheres, cultures were generated from dissociated anterior subventricular zone tissue of adult mice (aged 3 months) grown in DMEM/F12 media supplemented with B27 (Life Technologies), bFGF (20 ng/mL; BD Bioscience), and EGF (20 ng/mL; BD Bioscience). Initial cultures were expanded for 7–8 days and then dissociated and used to reseed new cultures. The passaged cultured were expanded for 7 days, after which the neurospheres were pelleted and washed with PBS before being cross-linked with 1% formaldehyde in PBS at room temperature for 9 min. Fixation reactions were terminated by the addition of 125 mM glycine, and the tissue was washed twice with PBS before addition of SDS Lysis Buffer (Millipore). Immunoprecipitation of protein/chromatin complexes was performed as described above with SH-SY5Y cells and used 6 µg of either goat anti-CTCF antibody (Santa Cruz) or normal goat IgG (Santa Cruz).

All immunoprecipitated *Th* promoter genomic DNA fragments were amplified and quantified using a 7600 Fast Real-time PCR System (Applied Biosystems) using SYBR Green PCR master mix (Applied Biosystems). Primer sequences used for amplifying were:

mouse *Th* E1 forward strand 5'-GGGATTTGCAGGAGCTTGCTCA-3'.

mouse *Th* E1 reverse strand 5'-CTTGGACTCTCAGGAGCCAACT-3'.

mouse *Th* E2 forward strand 5'-TTCCATGAAAGCACAACTGGC-3'.

mouse *Th* E2 reverse strand 5'-CAGGGTCGGCTGCT GAGGAT-3'.

mouse *Th* E3 forward strand 5'-TGGTCTGACTTTCA GCTGCCCAAT-3'.

mouse *Th* E3 reverse strand 5'-CAATACCACTCACT GACCTCACTG-3'.

mouse *Th* E4 forward strand 5'-GTGACCACCACTCA CGGGCT-3'.

mouse *Th* E4 reverse strand 5'-CCTGTGCACCACTG AGTCACATAA-3'.

mouse *Th* E5 forward strand 5'-TCCAGGAGAACAGA CGCCAGC-3'.

mouse *Th* E5 reverse strand 5'-GCCAGGCTGAAGGC AAGCACA-3'.

mouse *Th* negative control forward strand 5'-TGCCTCAG CAGAGCCTGAGT-3'.

mouse *Th* negative control reverse strand 5'-AAGCTCCC CGTGACTGTGTG-3'.

human *Th* E1 forward strand 5'-CCAAATCCTTCTGG GCCAGGA-3'.

human *Th* E1 reverse strand 5'-CCGTTCTCTTCA ACAATAGCC-3'.

human *Th* E2 forward strand 5'-TTCCATGAAAGCAC AACTGGC-3'.

human *Th* E2 reverse strand 5'-CAGGGTCGGCTGCT GAGGAT-3'.

human *Th* E3 forward strand 5'-TCGCTCTGGGCTG ACTTCC-3'.

human *Th* E3 reverse strand 5'-AACACAGGACAGAA TCCGCCGT-3'.

human *Th* E5 forward strand 5'-TTGGAGCAAAGCGG ACAAGCTCA-3'.

human *Th* E5 reverse strand 5'-GCGCATTCACTTCA GGTACCTC-3'.

human *Th* negative control forward strand 5'-AGGCTGAGGCTCTCCTTCCA-3'.

human *Th* negative control reverse strand 5'-GAACTCCA CCGTGAACCAGTACA-3'.

All ChIP experiments were conducted as three independent assays and the mean relative enrichment of the target region is reported with error bars representing the standard error of the mean. Statistical significance was assessed using either two-tailed Student's *t* test or ANOVA with appropriate post-hoc tests.

CTCF Over-Expression and Knock-down

To measure *Th* promoter activity when *Ctcf* was over-expressed, SH-SY5Y cultures were seeded in Primaria-coated 6-well plates (Corning) at 3×10^5 cells/well and incubated at 37 °C for 24 h before using Lipofectamine LTX reagent (Thermo Fisher Scientific) to co-transfect 4 µg of pCMV-Myc-CTCF (a gift from Dr. Mary Donohoe, Burke

Medical Research Institute), 2 µg of pGL4.20 reporter plasmid containing Firefly luciferase under control of the rat *Th* 4.5 kb promoter (generated in our laboratory), and 1 µg of pRL-CMV (Promega), which constitutively expresses Renilla luciferase in order to control for variations in transfection efficiency. After 24 h, cells were harvested and Firefly and Renilla luciferase activity levels were measured using the Dual-Glo Luciferase Assay System (Promega) with a LMaxII luminometer (molecular devices). Luciferase activities are reported as the mean of at least three independent measurements with error bars representing the standard error of the mean.

To measure *Th* promoter activity when *Ctcf* was knocked-down, SH-SY5Y cultures were seeded in Primaria-coated 6-well plates (Corning) at 3×10^5 cells/well and incubated at 37 °C for 24 h before being transfected with siRNA (Dharmacon Accell siRNA) according to the manufacture's instruction. Control cultures were treated with Accell Delivery Media (Dharmacon). All cultures were maintained at 37 °C for 48 h following transfection before were divided to conduct either Western blot or qRT-PCR analyses.

Knock-down of CTCF in cultures was confirmed by Western blots with goat anti-CTCF (Santa Cruz Biotech., sc-15914X at 1:1000 dilution) and mouse anti-beta-Actin (Sigma, A5316, at 1:6000 dilution). Blots were imaged with an Odyssey Imaging System (Li-Cor Biosciences) and intensities were quantified using ImageJ software (National Institutes of Health). Protein band intensities are reported as the mean of three individual trials with error bars representing the standard error of mean.

Th expression levels in transfected cells were measured by qRT-PCR. Total RNA from cells was collected using the GenCatch Total RNA Extraction System (Epoch Life Science). First-strand cDNA syntheses were generated with SuperScript III reverse transcriptase (Invitrogen). Quantitative PCR reactions were performed with TaqMan assays (Applied Biosystems) for human *Th* (Hs00165941_m1) and *beta-Actin* (Hs3044422). All reactions were carried out on a 7500 Fast Real-Time PCR System (Applied Biosystems). All samples were run in triplicate and expression levels for *Th* were normalized to *beta-Actin*. The mean relative expression levels are reported with error bars representing the standard error of the mean.

Results

Genomic Sequence Alignments and Identification of Conserved Upstream Regions

To identify regions of evolutionary conservation, 11 kb genomic DNA sequences upstream of the *Th* proximal promoter from nine mammalian species were aligned. This analysis

identified five regions of high homology based on the criteria of being a minimum length of 75 nucleotides length, having more than 50% homology to human, and being conserved in all nine species. These conserved regions were named E1–E5, going from the proximal to distal end of the *Th* upstream region (Figs 1, 2, 3). All five of the regions were found in only placental mammals and were absent in alignments with *Th* upstream regions from avians, reptiles, amphibians, and fish (data not shown).

Recruitment of Established Regulators of *Th* Transcription

Several individual transcription factor binding sites upstream of the *Th* proximal promoter have been reported, and many of these sites overlap with the conserved regions identified by our alignment (Fig. 4a). NURR1 (NR4A2) is a transcription factor necessary for *Th* expression in the midbrain [18, 19]. NURR1 can also drive dopaminergic differentiation in the OB [20]. The E2 region identified in this study contains a previously reported NURR1 binding site [21, 22] and our analysis shows that this site is highly conserved in mammals (Fig. 2b, Fig. 4a). Our analysis also identified potential novel NURR1 binding sites in the E3 and E5 regions (Fig. 2c, Fig. 3b, Fig. 4a). To test whether these unreported binding sites recruit NURR1, chromatin immunoprecipitation (ChIP) assays were performed in SH-SY5Y cells transfected with a FLAG-tagged NURR1 expression plasmid. These ChIP assays showed that NURR1 is recruited to the E2, E3, and E5 regions, but not to a negative control region in intron 10–11 (Fig. 4b). Together, the ChIP assays indicate that NURR1 regulates *Th* transcription by targeting multiple upstream regulatory regions.

The specification and maintenance of the midbrain dopaminergic phenotype is regulated by the interaction of NURR1

with other transcription factors, such as paired-like homeodomain 3 (PITX3) and forkhead box A2 (FOXA2). Both PITX3 and FOXA2 physically interact with NURR1 and function as co-activators [22, 23]. In the E5 region, our analysis found that a previously reported PITX3 binding site [24] is conserved in mammals (Fig. 3b, Fig. 4a) and also identified a previously unreported FOXA2 core binding motif (‘5-AAAYA-3’) [25, 26] that is conserved in rodents and primates (Fig. 2b). In addition, the E2 region contains three conserved FOXA2 binding sites (Fig. 2b) that were previously identified in embryonic midbrain progenitors [22]. A previous ChIP-seq study also indicated PITX3 targets a region that overlaps E2 [23], but a consensus PITX3 binding site motif (5'-TAATCC-3') was not identified within this region. The analysis of the E2 region, however, did find a core homeodomain sequence (5'-TAAT-3') 15 bp upstream from the NURR1 sites (Fig. 2b, Fig. 4a). Inspection of the E3 region also found a core homeodomain motif 5 bp upstream from a Nurr1 site (Fig. 3a, Fig. 4a).

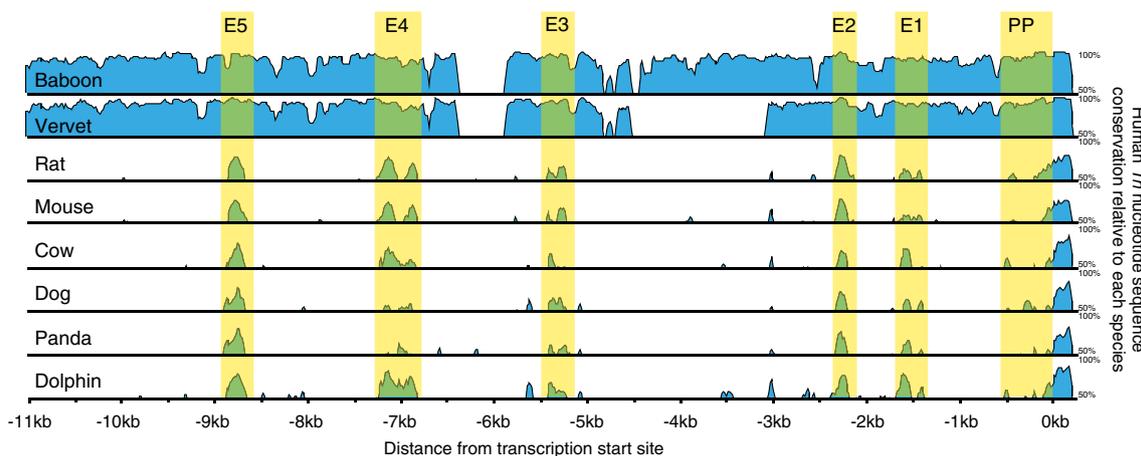


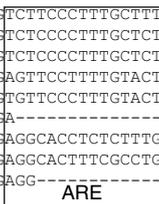
Fig. 1 Conservation of 11 kb genomic sequences upstream of mammalian *Th*. Graphic representation of human nucleotide sequence conservation as a percentage relative to each of the other mammalian

species examined. The five upstream regions of high conservation (E1–E5) and proximal promoter (PP) are highlighted in yellow

A

E1 region

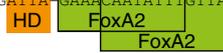
Human	-1741	GAGT	GAGCAGCAGCGGA	AACTGGGTTTTGTGGGATGCATAGGAGTTCACCCGGATAAGAGGTGGGTGAGGAATGACACTGCAAAACCGGGGATCAC	-1648
Baboon	-1742	GAGT	GAGCAGCAGCAGCGAGGTGGGTTTTGTGGGATGCGCAGGAGCTCACCCTGGATAAGAGGTGGCTGAGGAATGACACTGTGAACCTGGGGATCAC	-1649	
Vervet	-1745	GAGT	GAGCTGCGACCCAGATGGGTTTTGTGGGATGCACAGGAGTTCACCTGGAT-AGAGGTGGCTGAGGAATGACACTGTGAACCTGGGGATCAC	-1653	
Rat	-1665	GAGT	AAAATAGTCACTGAGCCGGGTTTTATGGGTTTTAGGAGCTTACTCAGA--GAAAGTAGATGAGAGATGCCATGCCAGTCTGAGTATCAC	-1574	
Mouse	-1661	GAGT	AAAATAGTCCCTGAGCTGGGGTTTTATGGGTTTTGCAGGAGCTTGTCTAGA--GAAGGCAGAGGAGAGATGCTGCGCCAAAGCTGGGTATCAC	-1570	
Cow	-1721	GAGT	GAATCGTAGCTGAGCT-GGCTTTGTGGGATGCATAGGAGTTCGCCAGGAGAAGAAACGGGTGAGGAACAGCATTTCAAATGGGGCTCAC	-1629	
Dog	-1499	GAAT	GAATCGTAGCTGAGCTGGGTTTTGTGGGATGCATAGGAGTTCACCCTGACAGCAAGAGGGGGTG-----AATCTTGCACACTAGGTATCGT	-1412	
Panda	-1495	GAAT	GAACCGTGTCTAAGTGGGTTTTGTGGGATGCATAGGAACTCAGCAGGAGGAGAAGGGGGTG-----AACCTTGCACAACTGGGTGTCTAT	-1408	
Dolphin	-1389	GAGT	GAATCGTAGCTGAGCT-GGTTTTGTGGGATGCATAGGAGTTCGCCAGGAGAAGAAAGGGTGGGGGAGTATTTCAAACAGGGTGTCTAT	-1297	
Human	-1647	GGAG	CCCCAAATCTTCTGG-GCCAGGAAGTGGGAAG-----GGTTGGGG-----GCTCTCCCTTTGCTTTGACT	-1583	
Baboon	-1648	GGAG	CCCCAAATCTTCTGG-GCCAGGAAGTGGGAAG-----GGTCGGGGG-----GCTCTCCCTTTGCTCTGACT	-1583	
Vervet	-1652	GGAG	CCCCAAATCTTCTGG-GCCAGGAAGTGGGAAG-----GGTCGGGG-----GCTCTCCCTTTGCTCTGACT	-1589	
Rat	-1573	AGAG	CCCCAGGCTCTCTGG-GAACGGAACGTGTAGGGCCAGAAGTTCAGCAAGGGA---GGTTAGGGA---GAGTCTCTTTGTACTGACT	-1489	
Mouse	-1569	AGAG	CCTCAGA--CTCTCG-AACAGGAACGTGGGGTCT--AGGTCAGCAGGGGA---GGTTAGGGA---GTGTCCCTTTGTACTGACT	-1490	
Cow	-1628	GGAG	CCCCAAATCTTCTGG--AGCCGAAACGGGGATTTTAGGAGTGGGACCAACAGC--GGGAGGTGAGGAGA-----AATCTTGCACACTAGGTATCGT	-1557	
Dog	-1411	GGCT	TCCCAAGCACCAGGAGCGGGAAGTGAGAACAGCCAGGGTGAGAGACCCGGCAGGGGAGGCATAGGAGGCACCTCTCTTTGACT	-1318	
Panda	-1407	GGAC	CCCCAGTCCCTCTGG-AGCAGGAAGTGGGAGAACCAAGGATGGGAGGGTTCAGCAGGGCAAG--GGAGAGGCACCTTTCGCTGACT	-1317	
Dolphin	-1296	GAAG	TCCAAACCTCTTT--AGCCGAAATGGGATGTTGGGGTGAGATCAATAG--GGTGGTGGGGAGG-----G	-1223	
Human	-1582	GAG	CACTCAGCCTCCCTGCAGAGGGCAGCGAGGAGCCACGGAGGGGTGTGGGACA--	-1528	
Baboon	-1582	GAG	CACTCA----CCCTGCAGAGGACAGCGAGGAGCCCGAGGGGTGTGGAACA--	-1532	
Vervet	-1588	GAG	CACTCA----CCCTGCAGAGGACAGCGAGGAGCCCGAGGGGTGTGGAACA--	-1538	
Rat	-1488	CAG	CACTTATCCTGCTCCAGGGGGCAA-TGGGGCCAGTGAGGATGCAGAGCA--	-1435	
Mouse	-1489	TAG	CACTTATCCTGCTTCTAGGGGGAA-GGGGGCCAGTGGGGATGCACAGCA--	-1436	
Cow	-1556	-GTC	ACTCGGGC-ACTTCCAAAGGGCAA-GGGGGCCACAGAAGATTGTGGAGCAAG	-1503	
Dog	-1317	GACT	TCTTGGCTGCTTCCAGAGGGCAATGGGAGCCATCCAGGGTTGTGAAGGA--	-1262	
Panda	-1316	GACT	GAGCTCTCTGCTTCCAGAGGGCAATGGGAGCCACAGAAGTGTGAGGTAAG	-1260	
Dolphin	-1222	GGT	CACTCAGGCTCTTCCAGAGGCAATGGGAGCCATAGAGGATTGTGGAGCAAG	-1166	



B

E2 region

Human	-2443	TGAA	AGCAC-AACTGGCCC-GGCAGGAAACCGAATTA-AAAAGCAATATTTGTAT-CAGTGGAAAGACATTTGCTGAAAGGTTAATCCACATCGG	-2353
Baboon	-2424	TGAA	AGCAC-AACTGGTCCAGGCAGGAAACCGAATTA-AAAAGCAATATTTGTAT-CAGTGGAAAGACATTTGCTGAAAGGTTAATCCACATCGG	-2333
Vervet	-2427	TGAA	AGCAC-AACTGGCCAGGCAGGAAACCGAATTA-AAAAGCAATATTTGTAT-CAGTGGAAAGACATTTGCTGAAAGGTTAATCCACATCGG	-2336
Rat	-2199	TGCA	AGCAC-CTCCAGCCGAGACAAGAAACGAATTA-AAAAGCAATATTTGTAT-CAGCGTAAGACATTTGCCGAAAGGTTAATCCACACTG-	-2109
Mouse	-2218	TGCA	AGCAC-CTCCAGCCGAGACAAGAAACGAATTA-AAAAGCAATATTTGTAT-CAGTGTAAAGACATTTGCCGAAAGGTTAATCCACATTC-	-2128
Cow	-2447	CGCA	AAG----CGCTGGCTCAGGGCCGCAAGGGGATTA-GAAAAGCAATATTTGTAT-CAGCAG-AGACATTTGCCGAAAGGTCAAGTCCACACCCG	-2361
Dog	-2111	CGCA	AGCCAGCAGCTGCTCAGGTAGGCAAGGGAAATTA-AAAAGCAATATTTGTAT-CAGCAGAAAGACATTTGCCGAAAGGTTAATCCACACCGA	-2019
Panda	-2121	TGCA	AGCCAGCCAGCTGGCTCAGGTGGGCAAGGAATTA-AAAAGCAATATTTGTAT-CAGCAGAAAGACATTTGCTGAAAGGTTAATCCACACTGA	-2029
Dolphin	-2048	CGC	AGG----CACCGGCTCAGGGAAGCCAAAGGGATTA-GAAAAGCAATATTTGTATTCGGTGG-AGATATTTGCCGAAAGGTTAATCCACAGCCG	-1961



C

E3 region

Human	-5541	CGCT	TGGGCTGACTTCCAAACACCCAATTATCCCTAAGTGCATCCGATCGATGCGAGGGGGGCT-GTTCGGGGCCACCTCGTCCATG-----	-5451	
Baboon	-5634	CGCT	TGGGCTGACTTCCAAACACCCAATTATCCCTAAGTGCATCCGATCGACCCGAGGGGGCT-GCTCCGGGGCC-GCTTCGTCATG-----	-5545	
Vervet	-4332	CGCT	TGGGCTGACTTCCAAACACCCAATTATCCCTAAGTGCATCCGATCGACCCGAGGGGGCT-GCTCCGGGGCCGCTTCGTCATG-----	-4242	
Rat	-4500	TGCC	TCTGGTCTGACTTCAACTGCCCAATTATCCCTAAGTGTCTCCATCGAC-----CGACC-GCAGCTGGTGTCTAGTACATA-----	-4418	
Mouse	-4627	TGCC	TCTGGTCTGACTTTCAGCTGCCCAATTATCCCTAAGTGTCTCCAT-----CGACC-GCAGCTGGTGTCTAGTACATA-----	-4596	
Cow	-4874	CTC	-----CCGACTTCCAAACGCACAAATTATCCCTGGGCGCTCGCATCGACCCGCGGGGGCTAGCCCCAGGCCCGGCTCTGCC-----	-4794	
Dog	-3856	TCCT	GCAAGTCCGCTTCCAAATGCCCAATTAT-CCTGAGAGCATCGATCGACCCGAGGGGCTGATGCGCC--GGCTCCCGCCGACTGCCACACT-----	-3766	
Panda	-3896	TCCT	GCAAGTCTCACTTCCAAATGACCAATTAT-CCTGAGTGCCTCGATCGACCCAGTGGGAGC--GGCTCTGGTCCGACCCACCATGCGCTGCCCC	-3800	
Dolphin	-4257	CCC	-----CTGACTTCCAAATGACCAATTAT-CCTGAGCGCATCGATCGACCCAGTGGGGCGGCGGCTCCTGGCCGCTCTGCTCG-----	-4178	
Human	-5450	-CGT	CCGCCCGCCTGCTGTGG---GGTCCATCTGATGGCTCATTAGGGCTAATTGCTCTGGCATT--GGGCTGACAGGGAC-----GGCGGA	-5364	
Baboon	-5544	-CGT	CCGCCCGCCTGCTGTGG---GGTCCATCTGATGGCTCATTAGGGCTAATTGCTCTGGCATT--GGGCTGACAGGGAC-----GGCGGA	-5458	
Vervet	-4241	-CGT	CCGCCCGCCTGCTGTGG---GGTCCATCTGATGGCTCATTAGGGCTAATTGCTCTGGCATT--GGGCTGACAGGGAC-----GGCGGA	-4155	
Rat	-4417	-TTCT	GAGTCTACAGCCCCGAGGCTGTCTCCATCCGATGGCTCGTTAGGGCTAATTGCTCTGGCATT--GGGCTGATGAGGACAAGAAATGGCTGGA	-4320	
Mouse	-4595	-TTCT	GAGTCTACAGCCCCGAGGCTGTCTCCATCCGATGGCTCGTTAGGGCTAATTGCTCTGGCATT--GGGCTGATGAGGACAAGAAATGGCTGGA	-4498	
Cow	-4793	-ATC	CCCTC-----GCTCCCTGCGAGC--CTCATCAGAGCTAATTGCTCTGACATTTCAGCGCTGACA-GGC-----	-4731	
Dog	-3765	-CCCC	CGCTGCCCGCCAGCTGCTCGGTTCCATCCGATGGTCTCGTTAGAGCTAATTGCTCTGACATTTCAGCTCTGACAGGGC-----	-3681	
Panda	-3799	CAC	CCCCCGCGCTCGCCAGCTGCTCGGTTCCATCCGATGGTCTCGTTAGAGCTAATTGCTCTGACATTTCAGCTCTGACAGGGC-----	-3714	
Dolphin	-4177	-ACC	TCTCCACCCGCTTCTGCTCTGCTCCCT-CCAGTCTCTCTTAGGGCTAATTGCTCTGACATTTCAGCTCGGACA-GGC-----	-4094	



ChIP assays in SH-SY5Y cells showed that endogenous PITX3 and FOXA2 are recruited to the E5, but not E2 or E3 regions (Fig. 4c, d). These findings indicate that NURR1, PITX3, and FOXA2 collectively target the E5 region to regulate *Th* expression. By contrast, the binding site motifs

identified in the E2 and E3 regions were not sufficient to recruit either PITX3 or FOXA2 under these conditions used in these studies. Previous work with SH-SY5Y cells also detected PITX3 binding to the E5, but not E2, region [24]. By contrast, studies with mouse embryonic midbrain cells and

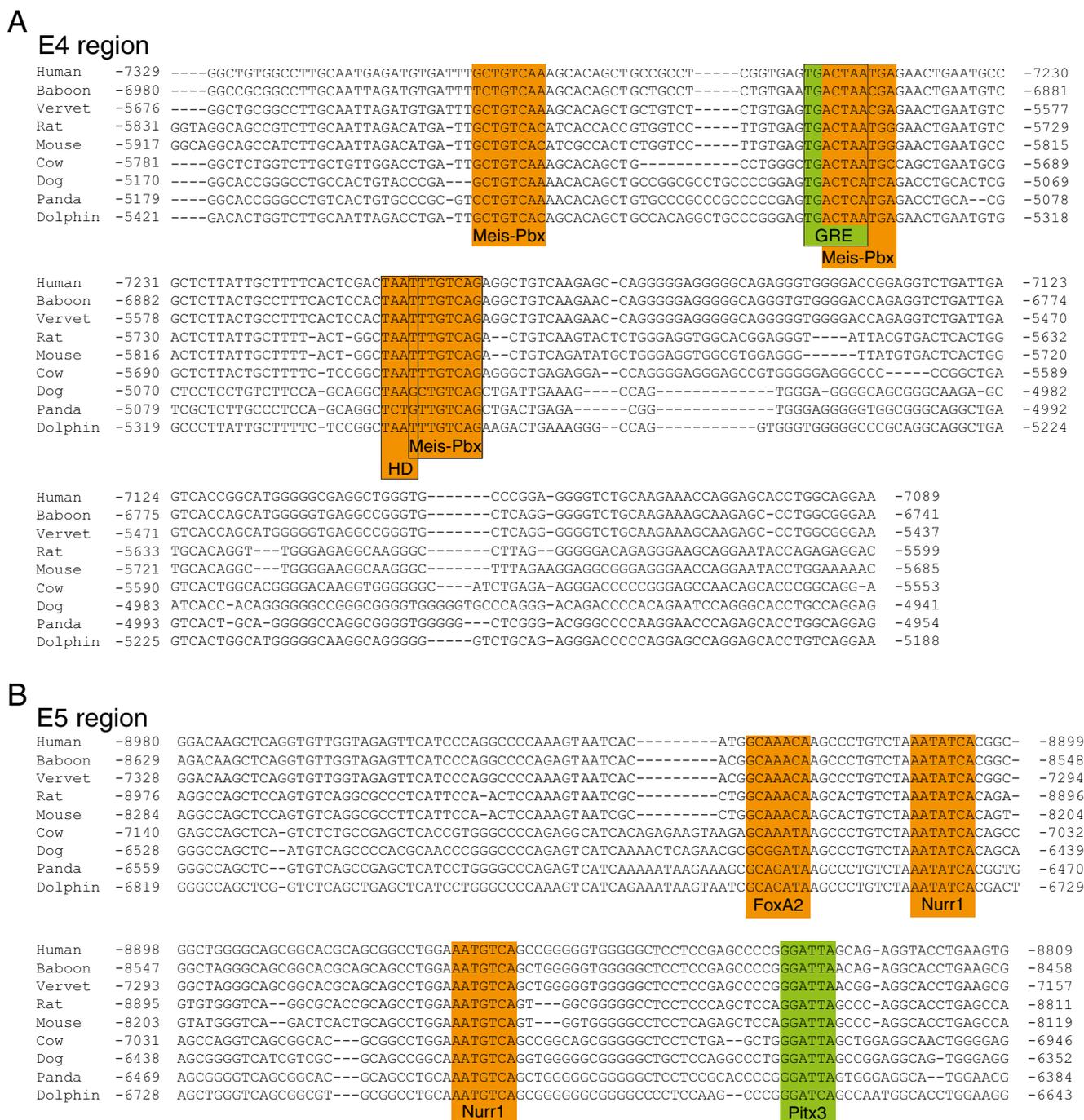


Fig. 3 Genomic sequence alignment and transcription factor binding sites for the E4 and E5 upstream regions. **a** Alignment of the E4 region reveals three novel and conserved binding site motifs and MEIS-PBX heterodimers [61] (MEIS-PBX; highlighted in orange). One of the MEIS-PBX sites overlaps a previously identified AP-1 binding site reported to mediate the induction of *Th* expression in response to glucocorticoid receptor activation [53] (GRE; highlighted in green).

Another MEIS-PBX site overlaps with partially conserved core homeodomain binding site motif (HD; highlighted in orange). **b** Alignment of the E5 region shows that a previously identified PITX3 binding site (highlighted in green) is conserved in mammals [24]. Analysis of this region also identified previously unreported and conserved binding site motifs for FOXA2 and NURR1 (highlighted in orange)

cultured MN9D cells reported PITX3 and FOXA2 binding to the E2 region [22, 23]. Given the importance of combinatorial interactions between transcription factors in regulating DNA-binding site specificity and the formation of transcription

regulatory protein complexes [27, 28], the differential recruitment of PITX3 and FOXA2 to the E2 region may reflect cell type differences in either post-translational modifications or the co-expression of other proteins.

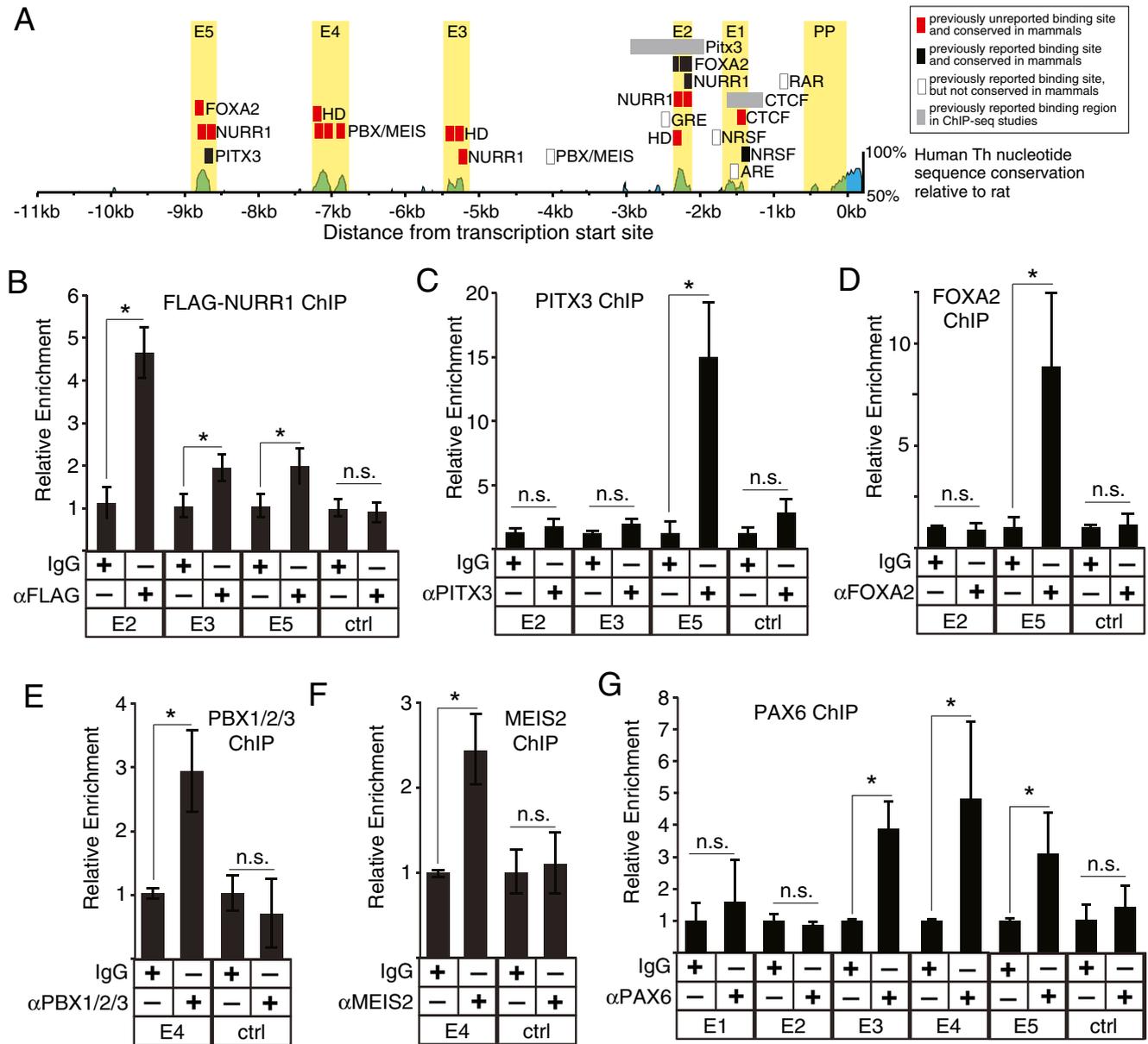


Fig. 4 Recruitment of established regulators of *Th* transcription to conserved upstream regions. **a** Graphic representation of transcription factor binding sites previously reported and identified in the current study. **b** ChIP assays testing NURR1 occupancy on the E2, E3, and E5 regions within SH-SY5Y cells over-expressing FLAG-NURR1 ($n = 3$). **c–d** ChIP assays testing PITX3 and FOXA2 occupancy, respectively, on the E2, E3, and E5 regions within SH-SY5Y cells ($n = 3$ for each). **e–f**

ChIP assays testing PBX1/2/3 and MEIS2 occupancy, respectively, on the E4 region within OB tissue ($n = 6$ and $n = 5$ for PBX and MEIS, respectively). **g** ChIP assays testing PAX6 occupancy on the E2, E3, and E4 regions within OB tissue ($n = 4$). For all ChIP assays, negative control regions are indicated by “ctrl,” asterisks indicate a significant difference ($p \leq 0.01$) and “n.s.” indicates non-significant enrichment

In OB dopaminergic neurons, the heterodimeric transcription factors pre B-cell leukemia homeodomain 1 (PBX1) and myeloid ecotropic viral integration site 2 (MEIS2) promote specification of the OB dopaminergic phenotype by directly binding to a site upstream of the mouse *Th* proximal promoter [29, 30]. Our analysis, however, found that this target site, which is between the E2 and E3 regions, was not conserved throughout mammals (Fig. 4a). By contrast, inspection of the E4 region identified three MEIS-PBX consensus binding

motifs that are well conserved (Fig. 3a, Fig. 4a), and ChIP assays with mouse OB tissue showed that both MEIS2 and PBX1/2/3 associate with the E4 region (Fig. 4e, f). The high conservation of the MEIS-PBX binding motifs within E4 suggests that this region mediates the conserved regulation of *Th* expression by MEIS and PBX proteins in the mammalian OB, whereas the previously identified MEIS-PBX site may be an important co-regulator of *Th* expression specifically in rodents.

MEIS2 also regulates *Th* transcription in the OB by physically interacting with the paired box 6 (PAX6) transcription factor [29]. PAX6 is necessary for OB dopaminergic neuron development and survival [31–33], but whether it targets a specific *Th* cis-regulatory region has not been established. PAX6 contains two functional DNA-binding domains, a paired domain and homeodomain, that each bind to distinct recognition sequences [34]. None of the upstream regions identified in this study contain consensus binding sites for the paired domain, but the E2, E3, and E4 regions contain core homeodomain (5'-TAAT-3') binding site motifs (Fig. 2b, c, Fig. 3a, Fig. 4a). ChIP assays with mouse OB tissue showed that PAX6 associates with the E3, E4, and E5 regions, but not with either E1, E2, or a negative control region (Fig. 4g). Together, these studies indicate that MEIS2, PBX1, and PAX6 regulate *Th* transcription in the OB, in part, by collectively targeting the E4 upstream region.

The neuron-restrictive silencer factor/RE1-silencing transcription factor (NRSF/REST) mediates epigenetic regulation of *Th* transcription in cultured cell lines and binds to a site found within the E1 region (Fig. 2a and Fig. 4a) [35–37]. This NRSF binding site shows conservation, but the *in vivo* relevance of this site is unclear. NRSF expression levels in the adult brain under physiological conditions are very low and there are several alternatively-spliced NRSF isoforms that have different DNA-binding affinities [38, 39]. Furthermore, NRSF is expressed at high levels during embryonic development, but the ENCODE database shows that NRSF does not occupy either the E1 region or the *Th* promoter in human embryonic stem cells [40]. Thus, NRSF may not regulate *Th* expression either during development or in the normal adult brain. Further studies are required, however, to establish whether NRSF regulates *Th* expression by targeting the E1 region under pathological conditions associated with elevated NRSF expression levels, such as MPTP-mediated neurodegeneration [36].

CTCF is a Novel, Direct Regulator of *Th* Promoter Activity in the Forebrain

Analysis of the E1 region identified an unreported and conserved CCCTC-binding factor (CTCF) consensus site adjacent to the previously reported NRSF site (Fig. 2a, Fig. 4a). CTCF is broadly expressed in many tissues and is a bifunctional regulator of transcription [41]. CTCF has not been previously reported to regulate *Th* expression, but ChIP-seq data from over 45 different cell lines in the ENCODE database show that CTCF associates with a region that overlaps with E1 (Fig. 4a). To establish whether CTCF targets the E1 region *in vivo*, ChIP assays were performed with tissue from several mouse brain regions. These assays showed that CTCF occupied the E1 region in the OB, but not in the midbrain (Fig. 5a,

b), suggesting that CTCF is a region-specific regulator of *Th* expression. In a small subset of cortical neurons, *Th* is transcribed, but not translated into protein [42]. ChIP assays showed that CTCF also occupied the E1 region in cortical tissue (Fig. 5c), suggesting that CTCF is a forebrain-specific regulator of *Th*. ChIP assays were also conducted with forebrain progenitors in neurospheres derived from the adult subventricular zone. These assays also found CTCF occupancy in the E1 region, revealing that CTCF regulation of *Th* is not limited to only mature neurons (Fig. 5d). To establish whether CTCF targeted the *Th* promoter outside of the nervous system, ChIP assays were conducted with liver tissue. These assays showed no significant occupancy of E1 (Fig. 5e) and indicate that CTCF is not required for maintaining *Th* repression in non-neural tissues.

To address whether modifying CTCF expression levels could alter *Th* promoter activity, luciferase transcription assays were conducted with SH-SY5Y cells since we had shown that CTCF is recruited to the E1 region in this cell line (Fig. 5f). These assays showed that co-transfection of a *Ctcf* expression plasmid reduced *Th* promoter activity (Fig. 6b), whereas siRNA-mediated partial knock-down of CTCF in these cells increased *Th* promoter activity (Fig. 6c–e). These findings indicate that CTCF functions as a repressor of *Th* transcription in these cells. Moreover, together with the ChIP assays, these findings show that CTCF is a forebrain-specific and direct regulator of *Th* transcription.

Discussion

Our analysis identified five highly conserved regions in the genomic DNA upstream of the *Th* proximal promoter in placental mammals. The restricted conservation of these regions to placental mammals likely reflects the presence of a *Th* gene paralogue (*Th2*) in non-placental mammals and other vertebrates (avians, reptiles, amphibians, and fish) that may have arisen from whole-genome duplication in early vertebrate evolution [43]. *In situ* hybridization studies in both developing and adult fish showed that the two *Th* paralogues have largely complementary expression patterns [43–45], suggesting that the *cis*-regulatory mechanisms controlling the single placental mammalian *Th* gene are distributed between the two paralogues in other vertebrates.

Our identification of highly conserved territories within the mammalian *Th* upstream region ranging in length between ~100 to ~200 bp significantly advances our understanding of genomic control of *Th* expression. These findings contrast with a previous analysis of humans, rats, and mice that only identified five short conserved sequences (18–60 bp in length) [6]. Four of these five previously reported sequences overlap with the regions identified in our alignments, but the limited

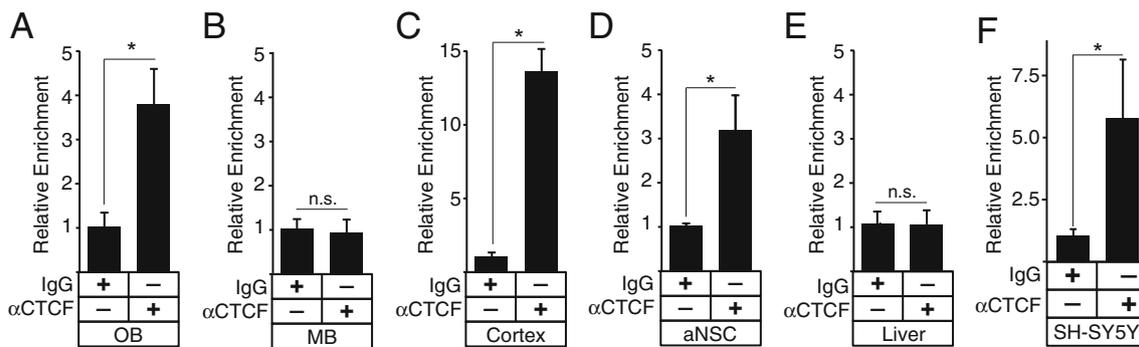


Fig. 5 Tissue-specific recruitment of CTCF to the E1 upstream region. **a–f** Relative enrichment of the E1 region in ChIP assays for CTCF with mouse olfactory bulb (OB), midbrain (MB), cortex, adult neurosphere

cultures (aNSC), liver tissue, and SH-SY5Y neuroblastoma cells, respectively ($n = 3$ for each)

number of species analyzed together with an apparent emphasis on contiguous sequences with near perfect homology likely prevented this previous analysis from identifying the larger territories of high conservation described in the current study.

The five upstream regions defined in the present study contain both previously and newly identified binding sites for transcription factors that regulate *Th* transcription during development and homeostasis. NURR1 is a key regulator of *Th* expression in several brain regions, and our study showed that NURR1 is recruited to the E2, E3, and E5 regions. NURR1 is a bifunctional regulator of *Th* transcription, and whether it either activates or represses transcription is influenced by its interaction with other transcription factors [21, 23]. Thus, the recruitment of NURR1 to E2, E3, and E5 regions is expected to make the regulatory role of these regions bifunctional as well, which suggests these regions can act as either distal enhancers or repressors depending on specific transcription factor expression profile of the cell.

NURR1 is also reported to bind the *Th* proximal promoter [46, 47], but a previous analysis of nucleotide conservation

within the *Th* proximal promoter found that this site was poorly conserved outside of rodents [10]. In addition, the position of this putative site in the promoter potentially conflicts with the binding of general transcription factors, such as TBP. In light of these observations, the findings from the present study suggest that the targeting of the E2, E3, and E5 regions is the conserved mechanism by which NURR1 regulates *Th* transcription.

The findings in this study suggest that the E5 region is important for regulating *Th* transcription in midbrain neurons. This region can recruit the transcription factors (NURR1, PITX3, and FOXA2; this study) and co-activator proteins (DJ1 and MTA1; [24]) that are established regulators of the midbrain dopaminergic phenotype. Other studies have also indicated that the sequences within the E2 regions are also important for driving *Th* expression in the developing midbrain by also recruiting NURR1, PITX3, and FOXA2 [21–23], which suggests that regulatory protein complexes assembled on the E5 region may work coordinately with those assembled on the E2 region. The role of PITX3 or FOXA2 on

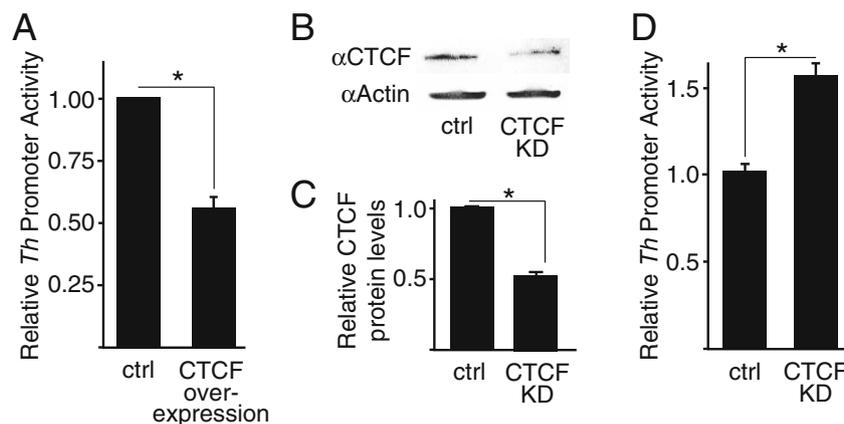


Fig. 6 CTCF regulates *Th* promoter activity. **a** Luciferase reporter activity with the 4.5 kb rat *Th* promoter in SH-SY5Y cells with over-expressing CTCF ($n = 3$). **b, c** Western blots and quantitation of protein levels, respectively, for SH-SY5Y cells with partial *Ctcf* expression knocked down by RNAi ($n = 3$ for each). **d** luciferase reporter activity

with the 4.5 kb rat *Th* promoter in SH-SY5Y cells with *Ctcf* expression knocked down by RNAi ($n = 3$). For all studies, asterisks indicate a significant difference ($p \leq 0.01$) and “n.s.” indicates non-significant difference

these regulatory regions within the midbrain, however, is limited to the substantia nigra since the loss of either PITX3 or FOXA2 only disrupts development of dopaminergic neurons in the substantia nigra [25, 48–51]. By contrast, NURR1 is critical for all ventral midbrain dopaminergic neurons [18, 19], and thus, other transcription factors may interact with NURR1 on the E2, E3, and E5 regions in midbrain regions outside the substantia nigra to regulate *Th* transcription. Further studies are required to establish whether these conserved upstream regions are targeted by other transcription factors that regulate specification and maintenance of midbrain dopaminergic phenotypes, such as LMX1A/B, EN1, NGN2, and OTX2 (reviewed in [52]).

In the OB, our studies indicate that MEIS2, PBX1, and PAX6 regulate *Th* transcription, in part, by collectively targeting the E4 upstream region. Our findings also suggest that the E3 and E5 regions may also contribute to regulation of *Th* in the OB by recruiting PAX6. The presence of a homeodomain motif in E3 suggest that PAX6 could directly bind this region, but the absence of either a paired domain or homeodomain recognition motif in E5 suggests that PAX6 associates with this region indirectly through interactions with other DNA-binding proteins. An important goal for future studies is to establish whether the association of PAX6 to the *Th* upstream regions is mediated by direct binding to the *Th* genomic DNA or through protein-protein interactions with factors, like MEIS2, that are also recruited to these regions. The high conservation of the MEIS-PBX binding motifs within E4 suggests that this region mediates the conserved regulation of *Th* expression by MEIS and PBX proteins in the mammalian OB, whereas the previously identified MEIS-PBX site may be an important co-regulator of *Th* expression specifically in rodents. It should also be noted that the one of the conserved MEIS-PBX sites overlap with AP-1 binding site that is responsible for induction of *Th* expression in response to glucocorticoid receptor activation [53]. AP-1 also has a highly conserved site in the *Th* proximal promoter that is an important modulator of stimulus-induced transcription [10, 54, 55]. This further suggests that the E4 region is an important regulator of *Th* expression in the olfactory system, and future studies will address with MEIS, PBX, or PAX6 proteins bound to E4 interact with AP-1 bound to either E4 or the proximal promoter to drive *Th* transcription.

The present study identified CTCF as a novel and region-specific regulator of *Th* transcription. The CTCF recognition sequence in the E1 region is strongly conserved and the ChIP assays showed that CTCF occupied this region preferentially in the forebrain. Since CTCF can mediate interactions between distant genomic regions [41], CTCF may be important for bringing E1 into the proximity of other *Th* distal regulatory regions. CTCF is a bifunctional regulator of transcription and its ability to either repress or enhance transcription is context-dependent [41]. Our functional studies indicate that CTCF is a

repressor of *Th* expression, and this property may be important for restricting *Th* transcription to specific neuronal subpopulations in the forebrain [42, 56]. Since CTCF is broadly expressed in the brain and other tissues [57], further studies are required to establish how brain region-specific recruitment to the E1 region is achieved.

The conserved upstream regions identified in this study significantly advance our understanding of the genomic DNA regulatory architecture for the mammalian *Th* gene. Moreover, our findings that these regions recruit transcription factors that are established regulators of *Th* expression provides insight into conserved mechanisms that regulate specification and maintenance of catecholaminergic neuronal phenotypes in the mammalian brain. These mechanisms, however, also likely include contributions from regulatory regions downstream of the transcription start site. Previous studies with transgenic mice showed that downstream regions can augment reporter expression levels [58, 59]. The first 2 kb downstream of the human *Th* gene was found to be sufficient for activating transcription in specific brain regions, but it did not suppress ectopic reporter gene expression in non-catecholaminergic regions [58]. This suggests that the combined input from upstream and downstream regulatory regions is required to drive *Th* expression at high levels specifically in catecholaminergic neurons.

Acknowledgements We thank Dr. Kaoru Saijo (Univ. California, Berkeley) for the NURR1 expression plasmid and Dr. Mary Donohoe (Burke Medical Research Institute) for the CTCF expression plasmid.

Compliance with Ethical Standards

Conflicts of Interest The authors declare that they have no conflicts of interest.

References

1. Kobayashi K, Morita S, Sawada H, Mizuguchi T, Yamada K, Nagatsu I, Hata T, Watanabe Y et al (1995) Targeted disruption of the tyrosine hydroxylase locus results in severe catecholamine depletion and perinatal lethality in mice. *J Biol Chem* 270(45):27235–27243. <https://doi.org/10.1074/jbc.270.45.27235>
2. Zhou QY, Quaipe CJ, Palmiter RD (1995) Targeted disruption of the tyrosine hydroxylase gene reveals that catecholamines are required for mouse fetal development. *Nature* 374(6523):640–643. <https://doi.org/10.1038/374640a0>
3. Kobayashi K, Noda Y, Matsushita N, Nishii K, Sawada H, Nagatsu T, Nakahara D, Fukabori R et al (2000) Modest neuropsychological deficits caused by reduced noradrenaline metabolism in mice heterozygous for a mutated tyrosine hydroxylase gene. *J Neurosci* 20(6):2418–2426
4. Willemsen MA, Verbeek MM, Kamsteeg EJ, de Rijk-van Andel JF, Aeby A, Blau N, Burlina A, Donati MA et al (2010) Tyrosine hydroxylase deficiency: a treatable disorder of brain catecholamine biosynthesis. *Brain* 133(Pt 6):1810–1822. <https://doi.org/10.1093/brain/awq087>

5. Smeets WJ, Gonzalez A (2000) Catecholamine systems in the brain of vertebrates: new perspectives through a comparative approach. *Brain Res Brain Res Rev* 33(2–3):308–379. [https://doi.org/10.1016/S0165-0173\(00\)00034-5](https://doi.org/10.1016/S0165-0173(00)00034-5)
6. Kessler MA, Yang M, Gollomp KL, Jin H, Iacovitti L (2003) The human tyrosine hydroxylase gene promoter. *Brain research. Mol Brain Res* 112(1–2):8–23
7. Kim TE, Park MJ, Choi EJ, Lee HS, Lee SH, Yoon SH, Oh CK, Lee BJ et al (2003) Cloning and cell type-specific regulation of the human tyrosine hydroxylase gene promoter. *Biochem Biophys Res Commun* 312(4):1123–1131. <https://doi.org/10.1016/j.bbrc.2003.11.029>
8. Schimmel JJ, Crews L, Roffler-Tarlov S, Chikaraishi DM (1999) 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):1–14. [https://doi.org/10.1016/S0169-328X\(99\)00234-X](https://doi.org/10.1016/S0169-328X(99)00234-X)
9. Yang C, Kim HS, Seo H, Kim KS (1998) Identification and characterization of potential cis-regulatory elements governing transcriptional activation of the rat tyrosine hydroxylase gene. *J Neurochem* 71(4):1358–1368
10. Wang M, Banerjee K, Baker H, Cave JW (2015) Nucleotide sequence conservation of novel and established -regulatory sites within the tyrosine hydroxylase gene promoter. *Front Biol (Beijing)* 10(1):74–90. <https://doi.org/10.1007/s11515-014-1341-z>
11. Liu J, Merlie JP, Todd RD, O'Malley KL (1997) Identification of cell type-specific promoter elements associated with the rat tyrosine hydroxylase gene using transgenic founder analysis. *Brain Res Mol Brain Res* 50(1–2):33–42. [https://doi.org/10.1016/S0169-328X\(97\)00163-0](https://doi.org/10.1016/S0169-328X(97)00163-0)
12. Min N, Joh TH, Kim KS, Peng C, Son JH (1994) 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(2):281–289. [https://doi.org/10.1016/0169-328X\(94\)90011-6](https://doi.org/10.1016/0169-328X(94)90011-6)
13. Sasaoka T, Kobayashi K, Nagatsu I, Takahashi R, Kimura M, Yokoyama M, Nomura T, Katsuki M et al (1992) Analysis of the human tyrosine hydroxylase promoter-chloramphenicol acetyltransferase chimeric gene expression in transgenic mice. *Brain research. Mol Brain Res* 16(3–4):274–286
14. Iacovitti L, Wei X, Cai J, Kostuk EW, Lin R, Gorodinsky A, Roman P, Kusek G et al (2014) The hTH-GFP reporter rat model for the study of Parkinson's disease. *PLoS One* 9(12):e113151. <https://doi.org/10.1371/journal.pone.0113151>
15. Matsushita N, Okada H, Yasoshima Y, Takahashi K, Kiuchi K, Kobayashi K (2002) Dynamics of tyrosine hydroxylase promoter activity during midbrain dopaminergic neuron development. *J Neurochem* 82(2):295–304. <https://doi.org/10.1046/j.1471-4159.2002.00972.x>
16. Brudno M, Do CB, Cooper GM, Kim MF, Davydov E, Green ED, Sidow A, Batzoglou S (2003) LAGAN and Multi-LAGAN: efficient tools for large-scale multiple alignment of genomic DNA. *Genome Res* 13(4):721–731. <https://doi.org/10.1101/gr.926603>
17. Frazer KA, Pachter L, Poliakov A, Rubin EM, Dubchak I (2004) VISTA: computational tools for comparative genomics. *Nucleic Acids Res* 32(Web Server issue):W273–W279. <https://doi.org/10.1093/nar/gkh458>
18. Saucedo-Cardenas O, Quintana-Hau JD, Le WD, Smidt MP, Cox JJ, De Mayo F, Burbach JP, Conneely OM (1998) Nurr1 is essential for the induction of the dopaminergic phenotype and the survival of ventral mesencephalic late dopaminergic precursor neurons. *Proc Natl Acad Sci U S A* 95(7):4013–4018. <https://doi.org/10.1073/pnas.95.7.4013>
19. Zetterstrom RH, Solomin L, Jansson L, Hoffer BJ, Olson L, Perlmann T (1997) Dopamine neuron agenesis in Nurr1-deficient mice. *Science* 276(5310):248–250. <https://doi.org/10.1126/science.276.5310.248>
20. Vergano-Vera E, Diaz-Guerra E, Rodriguez-Traver E, Mendez-Gomez HR, Solis O, Pignatelli J, Pickel J, Lee SH et al (2015) Nurr1 blocks the mitogenic effect of FGF-2 and EGF, inducing olfactory bulb neural stem cells to adopt dopaminergic and dopaminergic-GABAergic neuronal phenotypes. *Dev Neurobiol* 75(8):823–841. <https://doi.org/10.1002/dneu.22251>
21. Kim TE, Seo JS, Yang JW, Kim MW, Kausar R, Joe E, Kim BY, Lee MA (2013) Nurr1 represses tyrosine hydroxylase expression via SIRT1 in human neural stem cells. *PLoS One* 8(8):e71469. <https://doi.org/10.1371/journal.pone.0071469>
22. Yi SH, He XB, Rhee YH, Park CH, Takizawa T, Nakashima K, Lee SH (2014) Foxa2 acts as a co-activator potentiating expression of the Nurr1-induced DA phenotype via epigenetic regulation. *Development* 141(4):761–772. <https://doi.org/10.1242/dev.095802>
23. Jacobs FM, van Erp S, van der Linden AJ, von Oerthel L, Burbach JP, Smidt MP (2009) Pitx3 potentiates Nurr1 in dopamine neuron terminal differentiation through release of SMRT-mediated repression. *Development* 136(4):531–540. <https://doi.org/10.1242/dev.029769>
24. Reddy SD, Rayala SK, Ohshiro K, Pakala SB, Kobori N, Dash P, Yun S, Qin J et al (2011) Multiple coregulatory control of tyrosine hydroxylase gene transcription. *Proc Natl Acad Sci U S A* 108(10):4200–4205. <https://doi.org/10.1073/pnas.1101193108>
25. Ferri AL, Lin W, Mavromatakis YE, Wang JC, Sasaki H, Whittsett JA, Ang SL (2007) Foxa1 and Foxa2 regulate multiple phases of midbrain dopaminergic neuron development in a dosage-dependent manner. *Development* 134(15):2761–2769. <https://doi.org/10.1242/dev.000141>
26. Wederell ED, Bilenky M, Cullum R, Thiessen N, Dagpinar M, Delaney A, Varhol R, Zhao Y et al (2008) Global analysis of in vivo Foxa2-binding sites in mouse adult liver using massively parallel sequencing. *Nucleic Acids Res* 36(14):4549–4564. <https://doi.org/10.1093/nar/gkn382>
27. Merika M, Thanos D (2001) Enhanceosomes. *Curr Opin Genet Dev* 11(2):205–208. [https://doi.org/10.1016/S0959-437X\(00\)00180-5](https://doi.org/10.1016/S0959-437X(00)00180-5)
28. Ptashne M, Gann A (2002) *Genes & signals*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York
29. Agoston Z, Heine P, Brill MS, Grebbin BM, Hau AC, Kallenborn-Gerhardt W, Schramm J, Gotz M et al (2014) Meis2 is a Pax6 cofactor in neurogenesis and dopaminergic periglomerular fate specification in the adult olfactory bulb. *Development* 141(1):28–38. <https://doi.org/10.1242/dev.097295>
30. Grebbin BM, Hau AC, Gross A, Anders-Maurer M, Schramm J, Koss M, Wille C, Mittelbronn M et al (2016) Pbx1 is required for adult subventricular zone neurogenesis. *Development* 143(13):2281–2291. <https://doi.org/10.1242/dev.128033>
31. Kohwi M, Osumi N, Rubenstein JL, Alvarez-Buylla A (2005) Pax6 is required for making specific subpopulations of granule and periglomerular neurons in the olfactory bulb. *J Neurosci* 25(30):6997–7003. <https://doi.org/10.1523/JNEUROSCI.1435-05.2005>
32. Dellovade TL, Pfaff DW, Schwanzel-Fukuda M (1998) Olfactory bulb development is altered in small-eye (Sey) mice. *J Comp Neurol* 402(3):402–418. [https://doi.org/10.1002/\(SICI\)1096-9861\(19981221\)402:3<402::AID-CNE8>3.0.CO;2-0](https://doi.org/10.1002/(SICI)1096-9861(19981221)402:3<402::AID-CNE8>3.0.CO;2-0)
33. Ninkovic J, Pinto L, Petricca S, Lepier A, Sun J, Rieger MA, Schroeder T, Cvekl A et al (2010) The transcription factor Pax6 regulates survival of dopaminergic olfactory bulb neurons via crystallin alphaA. *Neuron* 68(4):682–694. <https://doi.org/10.1016/j.neuron.2010.09.030>
34. Sun J, Rockowitz S, Xie Q, Ashery-Padan R, Zheng D, Cvekl A (2015) Identification of in vivo DNA-binding mechanisms of Pax6 and reconstruction of Pax6-dependent gene regulatory networks

- during forebrain and lens development. *Nucleic Acids Res* 43(14): 6827–6846. <https://doi.org/10.1093/nar/gkv589>
35. Kim SM, Yang JW, Park MJ, Lee JK, Kim SU, Lee YS, Lee MA (2006) Regulation of human tyrosine hydroxylase gene by neuron-restrictive silencer factor. *Biochem Biophys Res Commun* 346(2): 426–435. <https://doi.org/10.1016/j.bbrc.2006.05.142>
 36. Suo H, Wang P, Tong J, Cai L, Liu J, Huang D, Huang L, Wang Z et al (2015) NRSF is an essential mediator for the neuroprotection of trichostatin A in the MPTP mouse model of Parkinson's disease. *Neuropharmacology* 99:67–78. <https://doi.org/10.1016/j.neuropharm.2015.07.015>
 37. Yang JW, Choi EY, Park MJ, Lee MA (2011) Expression of tyrosine hydroxylase is epigenetically regulated in neural stem cells. *Biochem Biophys Res Commun* 414(4):712–718. <https://doi.org/10.1016/j.bbrc.2011.09.141>
 38. Palm K, Belluardo N, Metsis M, Timmusk T (1998) Neuronal expression of zinc finger transcription factor REST/NRSF/XBR gene. *J Neurosci* 18(4):1280–1296
 39. Mori N, Mizuno T, Murai K, Nakano I, Yamashita H (2002) Effect of age on the gene expression of neural-restrictive silencing factor NRSF/REST. *Neurobiol Aging* 23(2):255–262
 40. Consortium EP (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature* 489(7414):57–74. <https://doi.org/10.1038/nature11247>
 41. Ong CT, Corces VG (2014) CTCF: an architectural protein bridging genome topology and function. *Nat Rev Genet* 15(4):234–246. <https://doi.org/10.1038/nrg3663>
 42. Baker H, Kobayashi K, Okano H, Saino-Saito S (2003) Cortical and striatal expression of tyrosine hydroxylase mRNA in neonatal and adult mice. *Cell Mol Neurobiol* 23(4–5):507–518. <https://doi.org/10.1023/A:1025015928129>
 43. Yamamoto K, Ruuskanen JO, Wullimann MF, Vernier P (2010) Two tyrosine hydroxylase genes in vertebrates. New dopaminergic territories revealed in the zebrafish brain. *Mol Cell Neurosci* 43(4): 394–402. <https://doi.org/10.1016/j.mcn.2010.01.006>
 44. Filippi A, Mahler J, Schweitzer J, Driever W (2010) Expression of the paralogous tyrosine hydroxylase encoding genes th1 and th2 reveals the full complement of dopaminergic and noradrenergic neurons in zebrafish larval and juvenile brain. *J Comp Neurol* 518(4):423–438. <https://doi.org/10.1002/cne.22213>
 45. Chen YC, Priyadarshini M, Panula P (2009) Complementary developmental expression of the two tyrosine hydroxylase transcripts in zebrafish. *Histochem Cell Biol* 132(4):375–381. <https://doi.org/10.1007/s00418-009-0619-8>
 46. Iwawaki T, Kohno K, Kobayashi K (2000) Identification of a potential nurr1 response element that activates the tyrosine hydroxylase gene promoter in cultured cells. *Biochem Biophys Res Commun* 274(3):590–595. <https://doi.org/10.1006/bbrc.2000.3204>
 47. Kim KS, Kim CH, Hwang DY, Seo H, Chung S, Hong SJ, Lim JK, Anderson T et al (2003) Orphan nuclear receptor Nurr1 directly transactivates the promoter activity of the tyrosine hydroxylase gene in a cell-specific manner. *J Neurochem* 85(3):622–634. <https://doi.org/10.1046/j.1471-4159.2003.01671.x>
 48. Smidt MP, Smits SM, Bouwmeester H, Hamers FP, van der Linden AJ, Hellemons AJ, Graw J, Burbach JP (2004) Early developmental failure of substantia nigra dopamine neurons in mice lacking the homeodomain gene Pitx3. *Development* 131(5):1145–1155. <https://doi.org/10.1242/dev.01022>
 49. Hwang DY, Ardayfio P, Kang UJ, Semina EV, Kim KS (2003) Selective loss of dopaminergic neurons in the substantia nigra of Pitx3-deficient aphakia mice. *Brain Res Mol Brain Res* 114(2): 123–131. [https://doi.org/10.1016/S0169-328X\(03\)00162-1](https://doi.org/10.1016/S0169-328X(03)00162-1)
 50. van den Munckhof P, Luk KC, Ste-Marie L, Montgomery J, Blanchet PJ, Sadikot AF, Drouin J (2003) Pitx3 is required for motor activity and for survival of a subset of midbrain dopaminergic neurons. *Development* 130(11):2535–2542. <https://doi.org/10.1242/dev.00464>
 51. Stott SR, Metzakopian E, Lin W, Kaestner KH, Hen R, Ang SL (2013) Foxa1 and foxa2 are required for the maintenance of dopaminergic properties in ventral midbrain neurons at late embryonic stages. *J Neurosci* 33(18):8022–8034. <https://doi.org/10.1523/JNEUROSCI.4774-12.2013>
 52. Arenas E, Denham M, Villaescusa JC (2015) How to make a midbrain dopaminergic neuron. *Development* 142(11):1918–1936. <https://doi.org/10.1242/dev.097394>
 53. Sheela Rani CS, Soto-Pina A, Iacovitti L, Strong R (2013) Evolutionary conservation of an atypical glucocorticoid-responsive element in the human tyrosine hydroxylase gene. *J Neurochem* 126(1):19–28. <https://doi.org/10.1111/jnc.12294>
 54. Liu N, Cigola E, Tinti C, Jin BK, Conti B, Volpe BT, Baker H (1999) Unique regulation of immediate early gene and tyrosine hydroxylase expression in the odor-deprived mouse olfactory bulb. *J Biol Chem* 274(5):3042–3047. <https://doi.org/10.1074/jbc.274.5.3042>
 55. Nagamoto-Combs K, Piech KM, Best JA, Sun B, Tank AW (1997) Tyrosine hydroxylase gene promoter activity is regulated by both cyclic AMP-responsive element and AP1 sites following calcium influx. Evidence for cyclic amp-responsive element binding protein-independent regulation. *J Biol Chem* 272(9):6051–6058. <https://doi.org/10.1074/jbc.272.9.6051>
 56. Saino-Saito S, Sasaki H, Volpe BT, Kobayashi K, Berlin R, Baker H (2004) Differentiation of the dopaminergic phenotype in the olfactory system of neonatal and adult mice. *J Comp Neurol* 479(4): 389–398. <https://doi.org/10.1002/cne.20320>
 57. Phillips JE, Corces VG (2009) CTCF: master weaver of the genome. *Cell* 137(7):1194–1211. <https://doi.org/10.1016/j.cell.2009.06.001>
 58. Choi EY, Yang JW, Park MS, Sun W, Kim H, Kim SU, Lee MA (2012) Transgenic mice expressing yellow fluorescent protein under control of the human tyrosine hydroxylase promoter. *J Neurosci Res* 90(10):1949–1959. <https://doi.org/10.1002/jnr.23085>
 59. Kaneda N, Sasaoka T, Kobayashi K, Kiuchi K, Nagatsu I, Kurosawa Y, Fujita K, Yokoyama M et al (1991) Tissue-specific and high-level expression of the human tyrosine hydroxylase gene in transgenic mice. *Neuron* 6(4):583–594. [https://doi.org/10.1016/0896-6273\(91\)90061-4](https://doi.org/10.1016/0896-6273(91)90061-4)
 60. Akahoshi E, Yoshimura S, Uruno S, Ishihara-Sugano M (2009) Effect of dioxins on regulation of tyrosine hydroxylase gene expression by aryl hydrocarbon receptor: a neurotoxicology study. *Environ Health* 8(1):24. <https://doi.org/10.1186/1476-069X-8-24>
 61. Penkov D, Mateos San Martin D, Fernandez-Diaz LC, Rossello CA, Torroja C, Sanchez-Cabo F, Warnatz HJ, Sultan M et al (2013) Analysis of the DNA-binding profile and function of TALE homeoproteins reveals their specialization and specific interactions with Hox genes/proteins. *Cell Rep* 3(4):1321–1333. <https://doi.org/10.1016/j.celrep.2013.03.029>