OXFORD

Cerebral Cortex, 2016; 1–16

doi: 10.1093/cercor/bhw290
Original Article

ORIGINAL ARTICLE

Genetic Labeling of Nuclei-Specific Thalamocortical Neurons Reveals Putative Sensory-Modality Specific Genes

Henrik Gezelius[1], Verónica Moreno-Juan[1], Cecilia Mezzera[1,4], Sudhir Thakurela[2], Luis Miguel Rodríguez-Malmierca[1], Jelena Pistolic[3], Vladimir Benes[3], Vijay K. Tiwari[2], and Guillermina López-Bendito[1]

[1]Instituto de Neurociencias de Alicante, Universidad Miguel Hernández-Consejo Superior de Investigaciones Científicas (UMH-CSIC), 03550 Sant Joan d'Alacant, Spain, [2]Institute of Molecular Biology (IMB), Ackermannweg 4, D-55128 Mainz, Germany, [3]EMBL, GeneCore, Meyerhofstr. 1, D-69117 Heidelberg, Germany, and [4]Current address: Champalimaud Neuroscience Programme, Champalimaud Centre for the Unknown, 1400-038 Lisbon, Portugal

Address correspondence to Guillermina López-Bendito, Instituto de Neurociencias de Alicante, Universidad Miguel Hernández-Consejo Superior de Investigaciones Científicas (UMH-CSIC), 03550 Sant Joan d'Alacant, Spain. Email: g.lbendito@umh.es

Abstract

The thalamus is a central brain structure with topographically ordered long-range axonal projections that convey sensory information to the cortex via distinct nuclei. Although there is an increasing knowledge about genes important for thalamocortical (TC) development, the identification of genetic landmarks of the distinct thalamic nuclei during the embryonic development has not been addressed systematically. Indeed, a more comprehensive understanding of how the axons from the individual nuclei find their way and connect to their corresponding cortical area is called for. Here, we used a genetic dual labeling strategy in mice to purify distinct principal sensory thalamic neurons. Subsequent genome-wide transcriptome profiling revealed genes specifically expressed in each nucleus during embryonic development. Analysis of regulatory regions of the identified genes revealed key transcription factors and networks that likely underlie the specification of individual sensory-modality TC connections. Finally, the importance of correct axon targeting for the specific sensory-modality population transcriptome was evidenced in a *Sema6A* mutant, in which visual TC axons are derailed at embryonic life. In sum, our data determined the developmental transcriptional profile of the TC neurons that will eventually support sensory processing.

Key words: brain development, gene expression, regional patterning, sensory system, thalamus

Introduction

The thalamus comprises several anatomically distinct nuclei. Each of the principal sensory-modality thalamic nuclei receives specific information from the periphery and projects topographically to the corresponding sensory cortical area (López-Bendito and Molnár 2003). Thus, the dorsal lateral geniculate nucleus (dLGN) conveys visual input from the retina to the primary visual cortex (V1), the ventrobasal complex (VB) conveys somatosensory information to the primary somatosensory cortex (S1), and the medial geniculate nucleus (MGN), directs auditory input to the primary auditory area (A1). By the time of sensory input onset in the mouse, around postnatal day 10

(P10), the distinct thalamic nuclei are anatomically well defined (Yuge et al. 2011). Interestingly, the specificity of the connectivity is established already at birth (Erzurumlu and Gaspar 2012), before sensory input is received, suggesting that intrinsic factors are responsible for establishing the early thalamocortical (TC) topographical axonal connectivity. This is similar to the case of the developing cerebral cortex where areal identity is established before arrival of cortical afferents by region-specific gene expression (Rakic 1988; Grove and Fukuchi-Shimogori 2003; Geschwind and Rakic 2013).

A number of well-known axon guidance molecules as well as transcription factors (TFs) expressed in the thalamus or in territories along the TC axonal path are known to be important for the correct wiring (López-Bendito and Molnár 2003; Garel and López-Bendito 2014). However, there is still much left to learn on the formation of the distinct nuclei and how the axons from those neurons find their respective cortical target. The vast majority of the thalamic neurons are generated in the ventricular zone between embryonic day (E) 10.5 and E15.5, with neurons of the different nuclei becoming postmitotic during discrete times (Angevine 1970). In the last few years it has been a considerable increase in the knowledge of the factors that specify the fate of thalamic projection neurons (Nakagawa and Shimogori 2012; Price et al. 2012; Song et al. 2015), Hornas opposed to that of GABAergic interneurons or habenular neurons, which are generated from adjacent ventricular zone areas. Still, the mechanisms by which differential gene expression within the thalamic progenitor domains give rise to distinct thalamic neuronal populations remains largely sparse (Vue et al. 2007; Ebisu et al. 2016). In postmitotic cells, however, specific TFs have been found that have a transient expression in distinct thalamic nuclei and are important to achieve a correct thalamic neuronal identity (Chen et al. 2009; Horng et al. 2009; Li et al. 2012) and/or TC connectivity (Seibt et al. 2003; Marcos-Mondéjar et al. 2012). One of the most studied TFs is *Gbx2*, which has been used as marker for the developing thalamus (Rubenstein et al. 1994) and it is expressed by all thalamic neurons at an early postmitotic stage (Chen et al. 2009). The laboratory of Li generated a versatile mutant mouse that concomitantly expresses a tamoxifen-conditional Cre recombinase protein (CreER) and a green fluorescent protein (GFP) under the control of the *Gbx2* regulatory sequences (Chen et al. 2009). This mouse represents a means to target distinct populations of thalamic neurons in a tamoxifen dependent manner.

In this study, we have crossed the $Gbx2^{CreER}$ line to a red fluorescent reporter mouse line ($R26^{tdTomato}$) and tagged the neurons of the principal thalamic nuclei at embryonic stages. Using this dual fluorescent strategy, the expression profile of each population was determined by microarray analysis at 2 time points important for TC axon growth, at E14.5 when the axons are traveling through the territory and interact with reciprocally growing corticothalamic axons, and at E18.5, when the first axons are entering into the cortical plate. Further, we analysed the promoter regions of the newly identified genes to reveal key TFs and involved networks that may be fundamental in determining the development and establishment of sensory-modality specific TC. Finally, to test whether the sensory-modality nucleus-specific transcriptome is set entirely cell autonomously or could be influenced by the axonal targeting behavior, we performed a genome-wide analysis on the thalamus of a *Sema6A* knockout mouse. In sum, our study reveals the developmental profile of the transcriptome related to each principal thalamic nucleus and suggests novel candidate genes important to establish TC connectivity.

Materials and Methods

Animals

In vivo cell labeling was obtained by crosses of $Gbx2^{CreERT2\text{-}ires\text{-}EGFP}$ (here denoted as $Gbx2^{CreER}$; Chen et al. 2009) and $R26^{tdTomato}$ (*loxP-stop-loxP-tdTomato* in the *Rosa26* locus, Ai14, Jackson Labs Strain 007914, Bar Harbor, ME; Madisen et al. 2010) mice. The morning of plug detection was designated as E0.5. Tamoxifen (Tam) was delivered at a dose of 3 mg by gavage feeding to the pregnant female at E9.0 or E9.5 as indicated. $Sema6A^{-/-}$ has been described before (Mitchell et al. 2001; Little et al. 2009), littermate heterozygous animals were used as controls since their axons are equal to wild type (Little et al. 2009). For in situ hybridization, experiments on wild type ICR or *Sema6A* mutant mice were performed. Animals of both sexes were used alike for all analyses. All animal procedures were approved by the Committee of Animal Research at the University Miguel Hernández and were carried out according to Spanish and European Union regulations.

Thalamic Cell Isolation and Sorting

Timed pregnant females were opened at selected time points. Cell isolation was done essentially as described before (Catapano et al. 2001). Additional details can be found in the Supplementary Material.

Expression Analysis of Thalamic Labeled Populations

Genome-wide expression analysis was done with Affymetrix Mouse Gene arrays with biological triplicates of each thalamic population. For each biological replicate, cells were collected from 2 to 3 (E18.5) or 1 to 2 E14.5 experiments (litters) and pooled to obtain a sufficient amount of RNA. The RNA was extracted from the isolated cells with RNeasy mini kit (Qiagen) according to the manufacturer's instruction, except that the last wash was with 80% ethanol. RNA was eluted in 30 µl RNase-free water and further concentrated by speed-vac. RNA integrity was verified and concentrations quantified by Agilent Bioanalyzer Pico Chip (Agilent). Biotinylated cDNA was synthesized from 10 ng of total RNA using the NuGen Ovation Pico WTA System V2 kit. Following fragmentation and labeling of 5 µg of ssDNA (NuGen Encore Biotin Module), cDNA samples were hybridized for 16 h at 45 °C on Affymetrix Mouse Gene 2.0 ST Arrays. Arrays were washed and stained in the Affymetrix Fluidics Station 450 and scanned using the GeneChip Scanner 3000 7G. Quality of array data was assessed using Expression Console (v 1.3) software (Affymetrix) prior to importing and analyzing the data in GeneSpring 12.6 (Agilent Technologies). Data were analyzed with settings for Affymetrix Expression experiment type, Exon analysis type and normalization using the Robust Multi-array Averaging (RMA) 16 summarization algorithm. Differentially expressed genes were determined after filtering the data sets for background level of fluorescence (as determined by Expression Console). Principal component analyses were performed on data that was filtered for statistical significant genes between pairs of conditions (pair comparisons of yellow vs. red, yellow vs. green and red vs. green for E18.5, and red vs. green for E14.5) using a one-way ANOVA and a cutoff of 0.05. Samples of

the same population are group together. Differentially expressed genes were determined by comparing a selected population against the remaining populations of the experiment, filtering for statistically significant genes (t-test, $P < 0.05$) testing for fold change (FC) greater than 1.5 between populations. Correction for multiple testing was not used, as current methods were too stringent for our data. Scatter plots were generated to visualize the differentially expressed (FC > 1.5, $P < 0.05$) genes. Genes were annotated from the probes with NetAffx NA33 (2010-10-01) against mouse genome (mm10), with most genes represented by a single probeset.

Immunohistochemistry and In situ Hybridization

Immunohistochemistry and in situ hybridization were done using standard procedures. The probes used for in situ hybridization are indicated in Table 1. Additional details can be found in the Supplementary Material.

Motif Prediction

As the number of genes in each population was very limited, we used an approach that tries to identify presence of TF motifs at promoters of given set of genes. We used ChIP-Enricher and looked for enrichment of motifs at −2000 and +200 base pairs with respect to the TSS (Chen et al. 2013).

Gene Network Identification

For creation of gene networks we used the Genomatix package. In brief, for each population a list of genes was provided for network creation using known interactions, coexpression, cocitations, and literature mining. Only TFs that were expressed within the population (as indicated by microarray) with their predicted targets were included. A basic network was first built using these genes. Then genes were connected for all possible interactions from literature. This resulted in a more connected network. Genes that were not connected were later removed.

DiI Tracing in *Sema6A* Embryos

For axonal tracing, brains were dissected out and fixed >12 h with 4% paraformaldehyde (PFA) in phosphate-buffered saline 0.01M. Small DiI (1,1′-dioctadecyl 3,3,3′,3′-tetramethylindocarbocyanine perchlorate; Invitrogen) crystals were inserted into the dLGN region of the thalamus after removing the caudal part of the brain. The dyes were allowed to diffuse at 37°C in PFA solution for 4–6 days. Vibratome sections (100 μm) were then counterstained with the fluorescent nuclear dye DAPI (Sigma-Aldrich).

Microdissection, RNA Isolation, and Affymetrix Microarray in *Sema6A* Mouse

To collect tissue from the dLGN nucleus in the *Sema6A* control and knockout mice, P0 animals were sacrificed and their brain was dissected out in RNase-free conditions to prevent RNA degradation. Vibratome sections (200 μm) were obtained and collected in ice-cold oxygenated aCSF (117 mM NaCl, 4.7 mM KCl, 1.2 mM $MgCl_2$ $6H_2O$, 2.5 mM $CaCl_2$ $2H_20$, 1.2 mM NaH_2PO_4, 25 mM $NaHCO_3$, and 0.45% D-glucose), and the dLGN thalamic nucleus was rapidly microdissected under a microscope. The tissue was maintained overnight at 4 °C in RNA-Later (Sigma) and stored at −80 °C for subsequent RNA extraction. For microarray hybridization, RNA was extracted from pooled thalamic tissue using the RNeasy Mini Kit (Qiagen), including a DNaseI step. Three replicates of control and 2 of mutant tissues were used for microarray analysis. Complementary RNAs (2 rounds of amplification) were hybridized to Affymetrix GeneChip Mouse Genome arrays 430 v2, and the signal intensities were analyzed using Partek Genomics suites (Partek) and Matlab (The MathWorks Inc.). The data were normalized using RMA and changes in gene expression >1.5-fold with a P value <0.05 (one-way ANOVA) were considered to reflect a significant difference in expression.

Table 1 Probes used for in situ hybridizations

Gene ID	Accession	nt	Source
Ascl1	NM_008553.4	480–2259	S. Martínez
Cck	NM_001284508.1	56–854	IMAgenes
Crabp2	NM_007759.2	62–859	P. Chambon (Ruberte et al. 1992)
Ebf1	NM_001290709.1	231–972	R. Grosschedl (Garel et al. 2002)
Galnt14	NM_027864.2	217–2180	IMAgenes
Gbx2	NM_010262.3	440–1448	J. Rubenstein (Marcos-Mondéjar et al. 2012)
Hs6st2	NM_001077202.1	1416–1982	T. Pratt (Conway, Price, et al. 2011)
Hs6st3	NM_015820.3	165–316 & 504–667	T. Pratt (Conway, Price, et al. 2011)
Igfbp4	NM_010517.3	6–2069	K. Kullander (Enjin et al. 2010)
Jam2	NM_023844.5	462–1136	J. Sanes (Kim et al. 2008)
Ldb2	NM_010698.4	248–1375	I. Bach (Bach et al. 1999)
Lef1	NM_010703.4	1247–1905	J. Galcerán (Galcerán et al. 2000)
Lhx2	NM_010710.4	277–878	S. Retaux (Marcos-Mondéjar et al. 2012)
Pknox2	NM_148950.3	176–1570	X. Guo (Zhou et al. 2013)
Pou2f2	NM_001163556.1	1432–1891	T. Shimogori (Shimogori et al. 2010)
Sp9	NM_001005343.2	2042–2167	J. Rubenstein (Long et al. 2009)
Tox	NM_145711.4	82–2851	IMAgenes
Tshz1	NM_001081300.1	3671–2840	L. Fasano (Caubit et al. 2000)
Vgf	NM_001039385.1	1539–1861	S.R. Salton (Hahm et al. 1999)
Zfhx4	NM_030708.2	2148–3118	Y. Komine (Komine et al. 2012)

Gene IDs, accession numbers for the mRNA and nucleotides (nt) covered for the antisense probes used for in situ hybridizations in this study. Source indicates providing scientist or company, with a reference when the clone previously has been used.

Results

Dual Genetic Labeling and Isolation of Embryonic Thalamic Nuclei

To unravel genes expressed in specific thalamic nuclei during the embryonic development, we developed a genetic strategy in mice to label and isolate nuclei-specific TC cells. We took advantage of the TF Gbx2, which is expressed by all early postmitotic thalamic neurons (Chen et al. 2009). A $Gbx2^{CreERT2\text{-}ires\text{-}EGFP}$ mouse (here denoted as $Gbx2^{CreER}$) was crossed with a $R26^{tdTomato}$ reporter mouse (Madisen et al. 2010). In the double-mutant mouse, inducible Cre recombinase (CreER), and enhanced GFP (EGFP) follow the current *Gbx2* expression. Tam assists nuclear entry of the CreER fusion protein, which initiates constitutive expression of the red reporter gene tdTomato by removal of an upstream stop cassette situated between 2 lox-sites (Fig. 1A). Only cells with *Gbx2* expression temporally coinciding with the presence of Tam will be labeled in red. The *Gbx2* is expressed at different developmental time points in distinct nuclei (Chen et al. 2009) and we determined that delivering Tam at embryonic day 9.0–9.5 (E9.0–9.5) selectively labeled the principal sensory thalamic nuclei in red (Fig. 1B–D′′ and see Supplementary Fig. 1B) (Chen et al. 2009). The visual dLGN and the somatosensory VB later downregulated *Gbx2* and thus lost the GFP and became red only by E18.5 (Fig. 1C′′). The auditory MGN, and the lateroposterior (LP) populations persistently expressed GFP together with tdTomato and thus were labeled in yellow (Fig. 1C′′–D′′). There were also nonprincipal thalamic cells that did not yet express *Gbx2* by the time of Tam delivery but which later started to express it and thus were only GFP positive at E18.5 (e.g., medial dorsal nucleus, MDN; Fig. 1C and see Supplementary Fig. 1B), here collectively denoted as nonprincipal thalamic nuclei. The thalamic neurons differently labeled by the paradigm described above were isolated and separated by fluorescence activated cell sorting (FACS) at E18.5, based on the red and green fluorescence intensities (Fig. 1E and see Supplementary Fig. 1C–H). The isolated cells were collected and the RNA was extracted. Thus, this strategy allowed us to isolate the RNA from dLGN and VB (red fluorescence), the MGN and to a small extent LP (yellow) and nonprincipal thalamic neurons (green).

Genes Enriched in Distinct Sensory-Related Thalamic Nuclei

The genes expressed in the distinct thalamic nuclei were analysed by Affymetrix GeneChip microarray after extracting RNAs from red, yellow, and green cells at E18.5 (Fig. 2A). To identify genes differently expressed between the nuclei, normalized expression profiles were compared between the populations. First, the results given by the microarray analysis allowed us to verify the strategy used to isolate nuclei-specific genes since we found the TF *Gbx2* to be expressed in both the MGN and nonprincipal populations, as previously described (Jones and Rubenstein 2004; Chen et al. 2009; indicated in Fig. 2B). Indeed, *Gbx2* was present among the 51 transcripts enriched (FC > 1.5, $P < 0.05$) in the group of green and yellow cells (Fig. 2B,C). Looking on the opposite end, this comparison also allowed us to identify 69 transcripts specifically enriched in the red population (FC > 1.5, $P < 0.05$). Genes among those with the most distinctive expression were chosen to corroborate their restricted expression patterns by in situ hybridization at E18.5. Genes specific for the auditory MGN and nonprincipal MDN populations were *Gbx2*, *Lhx2*, *Ldb2* (also known as *CLIM-1*), and *Tox* (Fig. 2C). Among the genes enriched in the red (dLGN and VB) population, we identified *Jam2*, *Galnt14*, *Hs6st3*, and *Zfhx4* in both dLGN and VB nuclei, while *Hs6st2* and *Sp9* were found to be specifically expressed in dLGN and *Cck* more restricted to VB (Fig. 2D). A number of genes related to dopaminergic neurons (e.g., *Th*, *Slc6a3* also known as dopamine transporter and *Slc18a2* also known as Vmat2) were also enriched in this population (data not shown). We next compared genes expressed in the yellow cells to those of red and green combined (Fig. 3), and we identified 15 presumptive MGN-enriched genes (FC > 1.5, $P < 0.05$) (Fig. 3B; left scatter plot). In a third comparison, the genes expressed in the green cells were compared with those of red and yellow combined (principal sensory nuclei) (Fig. 3B, right scatter plot). Most of the genes enriched in the yellow population showed strong expression in MGN. Whereas *Crabp2* and *Ascl1* (also known as *Mash1*) were exclusively expressed in this nucleus, *Pknox2* and *Tshz1* expressions were also found at lower levels in additional nuclei (Fig. 3C). Among the genes collectively enriched in the principal sensory thalamic nuclei (red and yellow), we confirmed the expression patterns for *Ebf1*, *Pou2f2*, *Lef1 Igfbp4*, *Vgf*, and *Vegfc* (Fig. 3D). These results demonstrate that the dual genetic labeling and gene profiling indeed identified genes expressed specifically in each of the principal sensory thalamic nuclei as well as genes common to them.

Embryonically Enriched Nuclei-Specific Thalamic Genes

Next, we unraveled genes whose expression were developmentally restricted to principal thalamic nuclei and thus could contribute to the acquisition of their specificity and connectivity programs. To this end, we further exploited the genetic strategy of labeling by analysing the gene profile of genetically fate mapped thalamic neurons at E14.5. At this stage, TC axons from principal thalamic nuclei are extending towards their cortical targets (Garel and López-Bendito 2014). Like above, delivery of Tam at E9.0–E9.5 induced expression of the fluorescent red protein tomato in the principal sensory nuclei, which allowed us to determine their anatomical boundaries at this early developmental stage (Fig. 4A–C′′). At E14.5, *Gbx2* started to be downregulated in the dLGN and VB (Fig. 4 and data not shown), yet the GFP was still present in these cells. Therefore, only very few red-only cells could be separated at this stage. Anyhow, despite the fact that the GFP makes these red E14.5 cells to appear yellow we called them red following our fate-mapping paradigm. Thalami of double-mutant embryos exposed to Tam at E9.5 were dissociated and all the red cells were collected and separated from the green-only positive ones by FACS (Fig. 4D). By comparing the genes expressed in the red and the green-only cells at E14.5, a list of 586 transcripts were found to be enriched in the red population compared with the green (Fig. 4E, FC > 1.5, $P < 0.05$). Given the permanent nature of the red label by the genetic fate mapping used, the cells that are red at E14.5 (or at any other stage) are the same cells that will be either red or yellow at E18.5 (Fig. 4F). When this list of genes was compared with the genes enriched in each of the populations at E18.5, 20 transcripts were preserved in the red population, 4 genes were preserved in the yellow population, and 8 additional transcripts were preserved in the principal population over embryonic development (Fig. 4G). The distribution of expression of these presumptive nuclei-specific genes within the developing thalamus at E14.5 was examined by in situ hybridization. At this early stage the different nuclei cannot yet be irrefutably distinguished as indicated by the

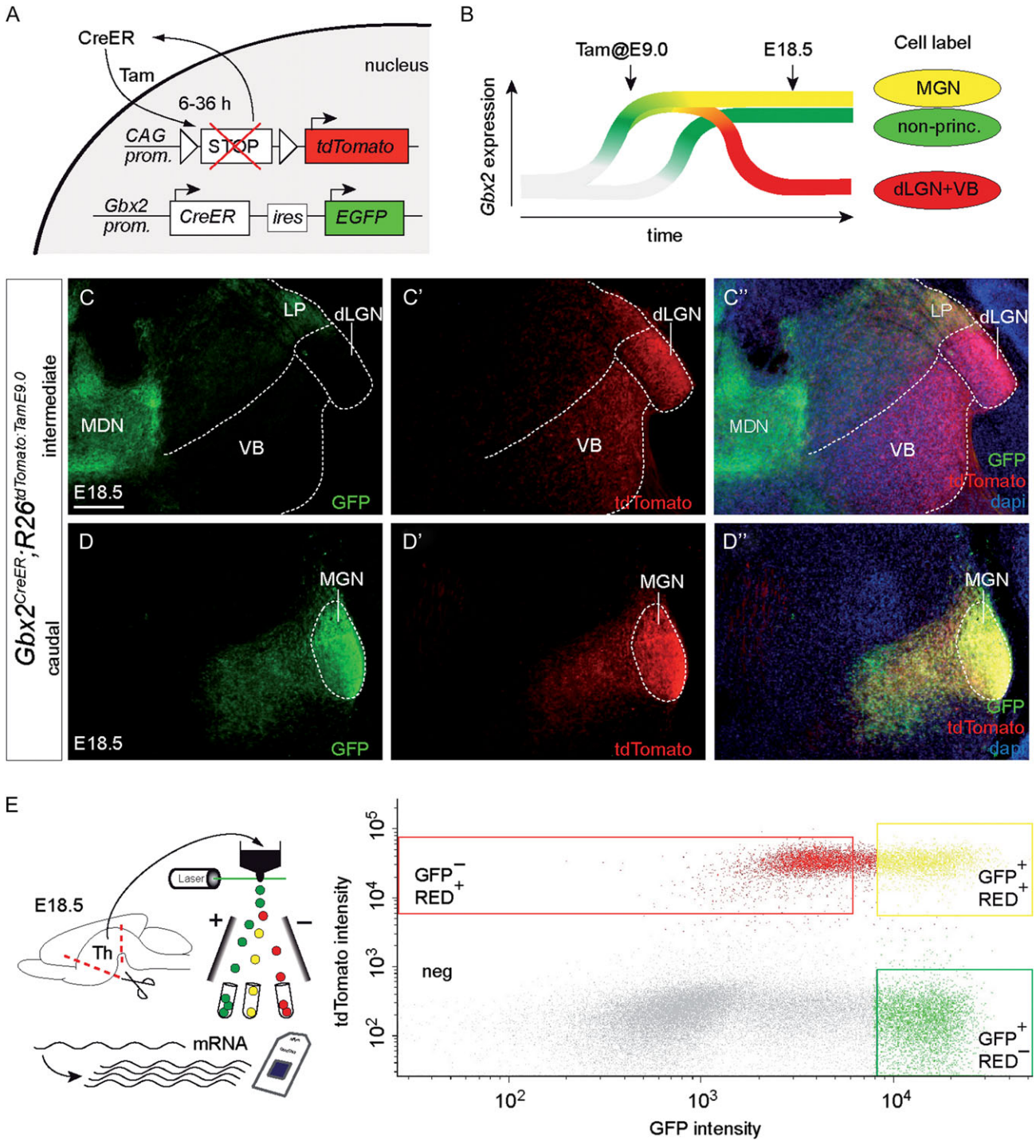

Figure 1. Isolation of neurons from distinct thalamic nuclei at E18.5. (*A*) Scheme of genetic strategy for dual labeling in vivo. EGFP follows current *Gbx2* expression while cells with *Gbx2* expression coinciding with presence of Tam are labeled red by induction of tdTomato expression. (*B*) Scheme of labeling of distinct thalamic nuclei. The *Gbx2* is expressed at different developmental time points in distinct nuclei and thus delivery of Tam at embryonic day 9.0 (E9.0) labeled principal thalamic nuclei red. Because the dLGN and VB later downregulate *Gbx2* and thus loose the GFP they are only red by the time of analysis at E18.5. Those cells that keep the GFP are the mainly MGN and those cells that did not yet express *Gbx2* by the time of Tam delivery which are nonprincipal thalamic neurons. (*C–D*) Fluorescent labels of the thalamic cells in $Gbx2^{CreER};R26^{tdTomato;Tam\ E9.0}$ mutant embryos at intermediate (*C–C''*) and caudal (*D–D''*) thalamic levels at E18.5. dLGN and VB were labeled in red only, MGN and to a small extent LP are expressing both green and red and hence were yellow. The nonprincipal MDN expressed only GFP, also seen at more rostral levels (Supplementary Fig. 1B). (*E*) Scheme of FACS of dissociated thalamic neurons followed by RNA extraction, amplification, and microarray analysis. Example of FACS population selection (right). Each dot corresponds to a cell. Green versus red fluorescence intensities were used to separate the tagged cells. The colored boxes indicate cells selected from each group, cells expressing both green and red were labeled yellow. Note that gaps were left between groups to avoid ambiguous labels. Scale bar represents 100 μm.

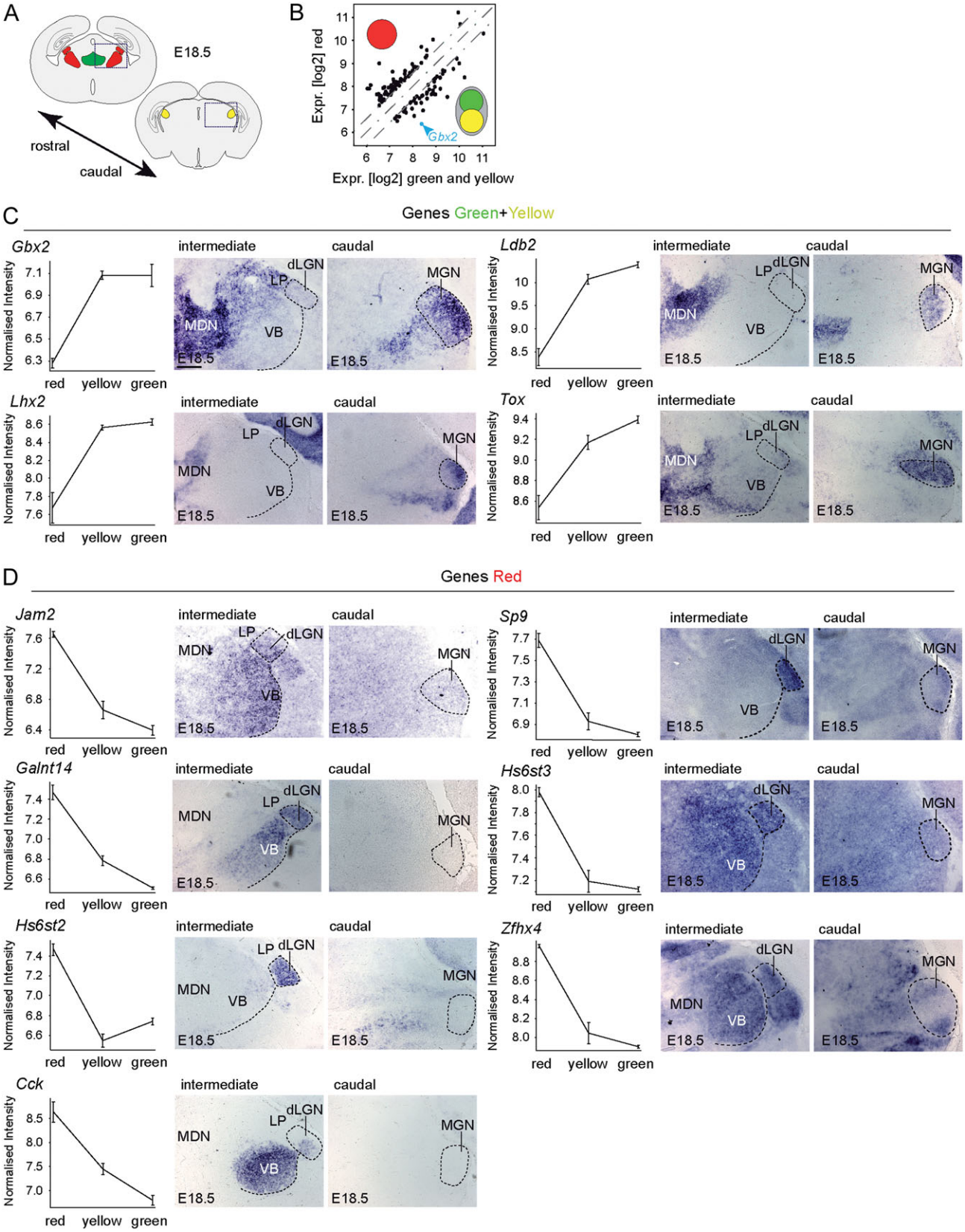

Figure 2. Genes expressed in distinct thalamic nuclei at late embryonic stage. (*A*) Scheme of coronal sections at E18 with colors indicating the labeling of principal thalamic nuclei in *Gbx2*^{CreER};*R26*$^{tdTomato:Tam\ E9.0}$ at intermediate (left) and caudal (right) levels. Boxes indicate approximate areas shown in (*C*–*D*). (*B*) Scatter plots of genes differentially expressed between thalamic nuclei. Sorted red cells were compared with the green and yellow cells combined. Only genes with FC > 1.5 ($P < 0.05$) are plotted. Stippled lines indicate FC 1.5. (*C*–*D*) Expression levels revealed by mean [log2] normalized intensity values from the microarray data (left, error bar

expression patterns by *Gbx2* and *Ebf1* (see Supplementary Fig. 2A,B), the later being enriched in principal neurons at E18.5 (Fig. 3). The genes *Hs6st2* and *Hs6st3*, which at E18.5 were selectively expressed in dLGN, displayed a lateral expression at rostro-intermediate thalamic levels at E14.5 (Fig. 4H and see Supplementary Fig. 2C). On the contrary, *Crabp2*, *Ascl1*, and *Tshz1* that later became specific for the MGN, were found laterally at more caudal levels were the presumptive MGN will be formed (Fig. 4H and see Supplementary Fig. 2D). In summary, we have identified genes that show restricted expression pattern in the developing thalamus and thus are presumptive markers of sensory-related thalamic nuclei throughout development, for example *Hs6st2* in the visual dLGN and *Crabp2* in the auditory MGN (Fig. 4H).

Enriched TFs and Regulatory Networks for Principal Thalamic Nuclei Genes

We next aimed to further identify TFs that are key for determining the identity and expression profile of the neurons of each thalamic nucleus. Using a bioinformatics approach, each group of distinctive expressed genes of the different populations were screened for TF binding motifs in their promoter regions. Furthermore, in order to put this in a more biological context, we looked for known interactions between the identified enriched TFs and their potential targets by performing a gene network analysis for each of the populations (Figs 5 and 6). This network assembling included not only the predicted and expressed TFs but also the predicted targets that were particularly expressed in each population (blue boxes vs. yellow boxes, respectively; Figs 5 and 6). Although these networks only included the TFs and their potential targets that are already known to interact, they revealed that the novel and thalamic nuclei-specific TFs we found are connected to described biological networks. Nevertheless, our bioinformatics analysis provided the entire list of predicted downstream target genes for the set of TFs enriched in the distinct populations, which suggests additional interactions (Tables in Figs 5C,F and 6C,F).

When looking at the regulatory sequences of the genes in the different populations at E18.5, the enriched genes of the dLGN and VB (red) populations predicted 46 TF motifs ($P < 0.05$) (Fig. 5A), out of which 14 were expressed within the population according to the microarray analysis (colored in blue). A network building using these expressed motifs along with their potential targets showed a large number of connections (Fig. 5B,C). The distinctive MGN (yellow) genes predicted 34 TF motifs and thus the identified network was very different from that of the dLGN/VB population (Fig. 5D–F).

The genes expressed in all principal thalamic nuclei (red and yellow) at E18.5 revealed 35 TF motifs that were also organized into a highly interconnected network, which was very distinct from other thalamic populations (Fig. 6A–C). Finally, we screened for TF motifs in the 15 genes specifically expressed in principal thalamic neurons at E18.5 (red and yellow) and also at E14.5 (red) during the development, which revealed motifs for 25 TFs with 7 of those being expressed within the population (Fig. 6D). Network analyses for this population showed both cytosolic and nuclear interactions between the identified TFs and their potential targets (Fig. 6E,F). Next, we went one population at the time and screened the list of motifs to identify any "hub" TF that is also specifically expressed within the population. In 2 of the populations we identified such hub TFs. We found *Jun* to be specifically expressed by dLGN/VB (red population) at E18.5 (Fig. 5B) and *Lef1* in the case of the principal thalamic (yellow and red) population at E18.5 (Fig. 6B). These results suggest that *Jun* and *Lef1* could play a "master" role in regulating other genes enriched in the distinct thalamic populations studied.

Detailed and comparative analysis of networks from different populations revealed that genes that are specifically expressed in dLGN/VB (red) at E18.5, MGN (yellow) at E18.5 and developing principal thalamic nuclei (yellow and red at E18.5 and red at E14.5) populations have distinct set of TFs that regulate them resulting in a unique interaction network in each population. Overall, our bioinformatics analysis of the expression profiles of distinct genetically labeled populations revealed a set of TFs that are likely important for establishing the transcriptional profile of principal thalamic neurons, which ultimately determines their connectivity and functional properties.

Nuclei-Specific Gene Expression Profile Might Depend on Cortical Connectivity

Our results so far unravel the complete transcriptional gene profile and network interactions for each sensory-modality specific thalamic nuclei. Transcriptional gene programs are known to control developmental processes such as neuronal specification and axon guidance (e.g., Seibt et al. 2003; Li et al. 2012). However, it is also potential that the gene profile of the neurons of a given nucleus is influenced by their axonal connectivity (Lin et al. 1998; Haase et al. 2002; Vrieseling and Arber 2006). To test this possibility, we sought out for an in vivo mouse model that has an embryonic nucleus-specific derailment of TC connectivity. In *Sema6A*$^{-/-}$ mice, the axons of the thalamic dLGN neurons are misguided during development due to the disruption of the transmembrane axon guidance molecule semaphoring 6 A and only a very small proportion of them reach the V1 cortical area prenatally (Little et al. 2009) (Fig. 7A). Thus, the *Sema6A*$^{-/-}$ mouse represents a means to test whether the gene programs and network identified are changed in a model in which most of the visual TC axons are not connected to their normal target area. Hence, mRNA from dLGN of control and homozygous embryos were isolated and microarray was performed (Fig. 7B). Several transcripts showed an altered expression (FC > 1.5, $P < 0.05$) between dLGN of control and *Sema6A*$^{-/-}$ animals, with the *Sema6A* transcripts being among the most downregulated (Fig. 7C). Note that a different array with more probes per transcripts was used in this analysis. When the genes with altered expression were compared with the genes enriched in the dLGN and VB at E18.5 (red population; Fig. 2), 23% (14 out of 60) of the specific genes were downregulated in the dLGN of *Sema6A*$^{-/-}$ mice. Remarkably, almost all of the altered red-specific genes (12/14) code for membrane-associated proteins, such as Hs6st2 (Habuchi et al. 2000) and Jam2 (Palmeri et al. 2000). In situ hybridization for *Hs6st2* and *Jam2* confirmed their reduced expression in the dLGN of *Sema6A*$^{-/-}$ mice (Fig. 7D). In sum, our study has identified the

represent SEM) and in situ hybridization confirmation of expression of selected genes from each group. Intermediate (middle) and caudal thalamus (right). (C) *Gbx2*, *Lhx2*, *Ldb2*, and *Tox* were restricted to the yellow (MGN) and green (MDN, nonprincipal) populations. (D) Selected genes enriched in the red (dLGN and VB) populations. *Jam2* and *Galnt14* were expressed in both VB and dLGN. *Hs6st2* was specifically expressed in dLGN, while *Cck* showed expressions more restricted to VB. The TF *Sp9* showed strong expression in the visual LGN nuclei. *Hs6st3* was expressed in both VB and dLGN. The *Zfhx4* expression was strong in VB and dLGN and some additional areas, with minor expression in the MGN and MDN. Scale bar represents 100 μm.

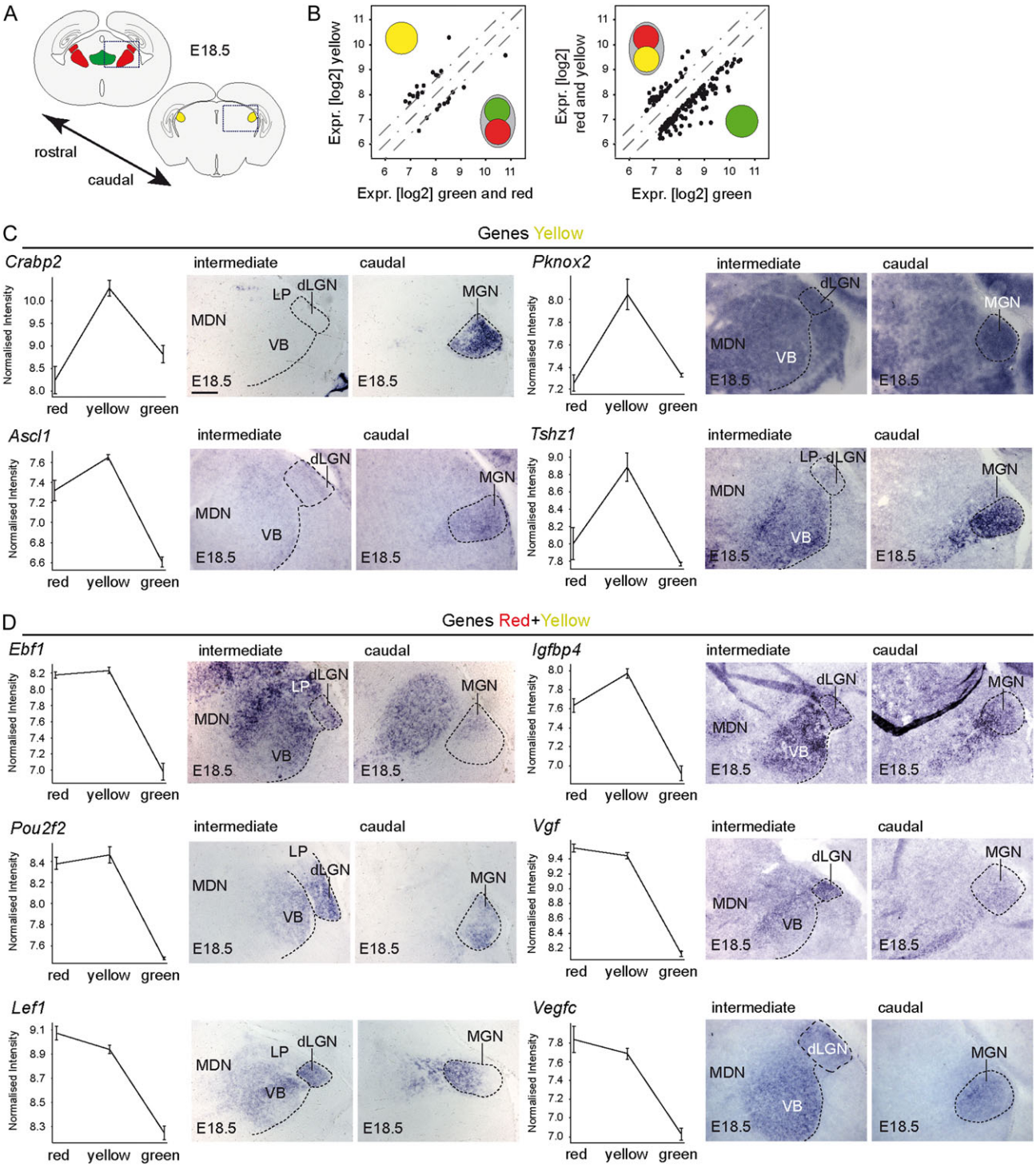

Figure 3. Genes expressed in principal thalamic nuclei at late embryonic stage. (*A*) Scheme of coronal sections at E18 as in Fig. 2. (*B*) Scatter plots of genes differentially expressed between thalamic nuclei. Yellow cells were compared with red and green combined (left) and red and yellow combined cells were compared with green (right). Only genes with FC > 1.5 ($P < 0.05$) are plotted. Stippled lines indicate FC 1.5. (*C–D*) Expression levels from the microarray data (left) and in situ hybridization confirmation of expression of selected genes from each group. Intermediate (middle) and caudal thalamus (right). (*C*) Two of the genes enriched in the yellow population, *Crabp2* and *Ascl1*, showed expression restricted to MGN. Genes with enriched, but not exclusive, expression were *Pknox2* and *Tshz1*. (*D*) In situ hybridization for selected genes, *Ebf1*, *Pou2f2*, *Lef1*, *Igfbp4*, *Vgf*, *and Vegfc* enriched in red and yellow principal populations (dLGN, VB, and MGN). Scale bar represents 100 μm.

transcriptional gene profile of sensory-modality TC neurons during prenatal development and suggests that a correct axonal connectivity is necessary to establish and/or maintain this gene imprinting.

Discussion

Identification of TFs uniquely expressed by distinct sensory modality-related thalamic neurons is crucial for understanding the intrinsic mechanisms controlling programs of TC

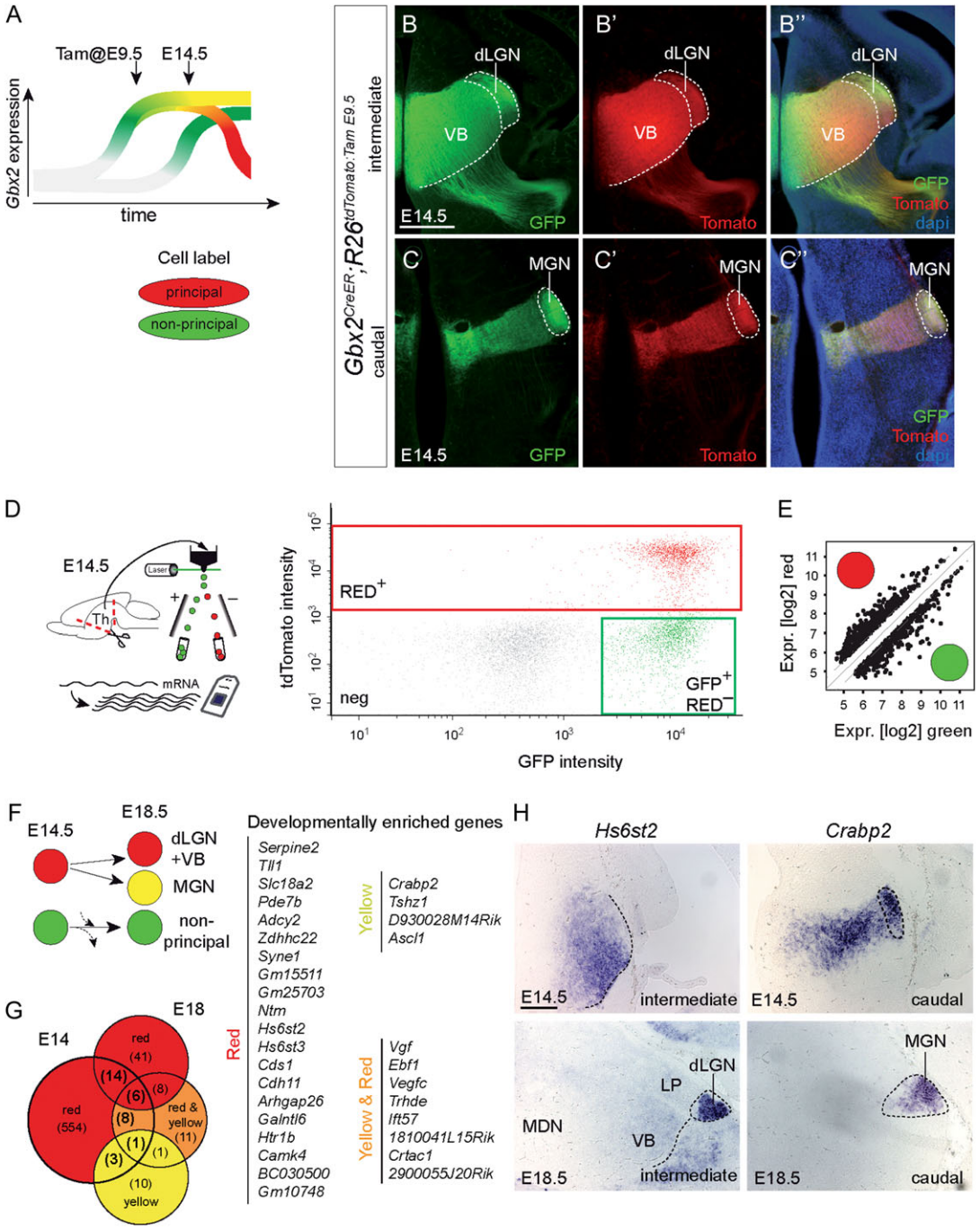

Figure 4. Identification of genes expressed in principal thalamic neurons at embryonic development. (*A*) Delivery of Tam at E9.5 label principal sensory thalamic neurons in red. At the time of analysis, E14.5, the green fluorescence is still present in most of those cells at this stage and additional nonprincipal thalamic neurons have started to express *Gbx2* and thus are green. (*B–C*) Fluorescent labels of the thalamic cells in *Gbx2*^{CreER};*R26*$^{tdTomato;Tam\ E9.5}$ mutant embryos at E14.5 at intermediate (*B*) and caudal (*C*) thalamic levels. Putative dLGN, VB, and MGN are labeled in red. Green is exclusive in medial nonprincipal thalamic nuclei. (*D*) Illustration of FACS procedure. Example of FACS of E14.5 double-mutant embryos (right). Green versus red fluorescence intensity was used to separate the cells. The colored boxes indicate cells selected from each group. Note that almost no red-only cells could be isolated at this stage. (*E*) Genes differentially expressed between sorted red and green thalamic cells. Only genes with FC > 1.5 ($P < 0.05$) are plotted. Stippled lines indicate FC 1.5. (*F*) Schema representing the scenario: cells expressing red at E14.5 will later give rise to either red or yellow cells at E18.5, while cells with only GFP expression at E14.5 could be different from the green population at E18.5. (*G*) Venn diagram showing overlaps between the lists of genes enriched in principal (red) population at E14.5 and the different principal populations (red and/or yellow) at E18.5 (left). The numbers in parenthesis indicate the number of genes in each category. There were 20 genes in common between the red populations at E14.5 and E18.5 and 4 genes in common between E14.5 red and E18.5 yellow. When including the genes enriched in red and yellow populations, 8 additional transcripts were found enriched also in the red population at E14.5. (*H*) In situ hybridization for 2 of the developmentally enriched genes at E14.5 and E18.5, *Hs6st2* at intermediate (left) and *Crabp2* at caudal levels (right).

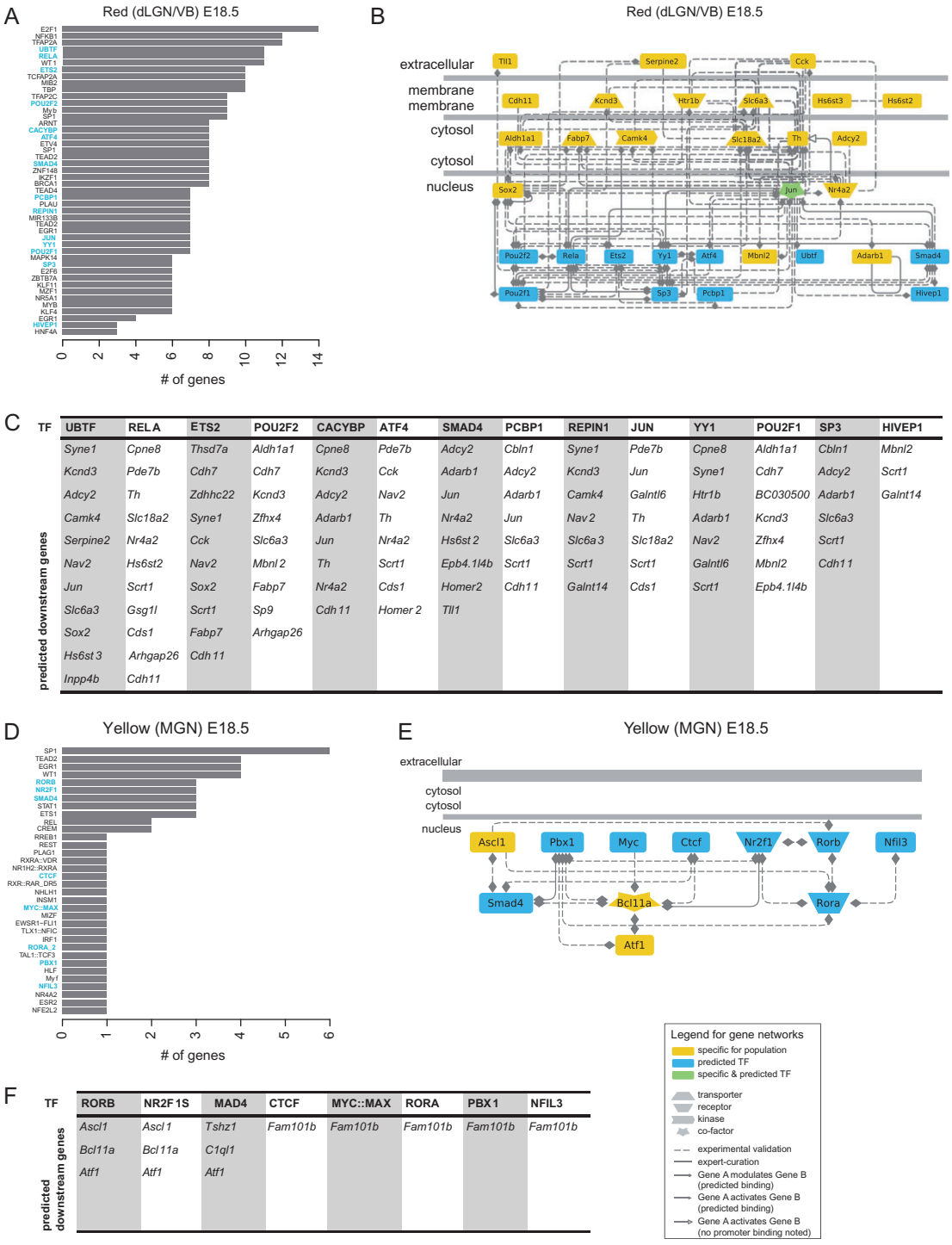

Figure 5. TF binding motifs and gene interaction networks in the distinct principal thalamic populations. (*A*) Bar plot showing predicted TFs at promoters of the genes specifically expressed in dLGN/VB (red) population at E18.5. The *x*-axis shows the number of genes that were found to have a motif for the respective factor on the *y*-axis. Only those factors that were found to be significant ($P < 0.05$) are plotted. TFs expressed in this population according to the microarray analysis are highlighted in blue. (*B*) Gene network using the expressed significant TFs and their potential targets for dLGN/VB (red) population at E18.5. (*C*) TF predicted by motifs analysis of genes enriched in the red population (top row) and the downstream genes having each motif. Only TFs that are expressed within the population and their targets are shown. (*D*) Bar plot showing predicted TFs at promoters of the genes specifically expressed in MGN (yellow) population at E18.5. (*E*) Gene network using the expressed significant TFs and their potential targets for MGN (yellow) population at E18.5. (*F*) TF predicted by motifs analysis of genes enriched in the yellow population (top row) and the downstream genes having each motif. Only TFs that are expressed within the population and their targets are shown.

connectivity. Here, we have developed a genetic dual fluorescent labeling strategy to identify genes with selective expression in developing thalamic neurons. First, the strategy was validated as *Gbx2* and *Lhx2*, which are well-known genes expressed in auditory MGN and medial nonprincipal thalamic populations (Nakagawa and O'Leary 2001; Suzuki-Hirano et al. 2011;

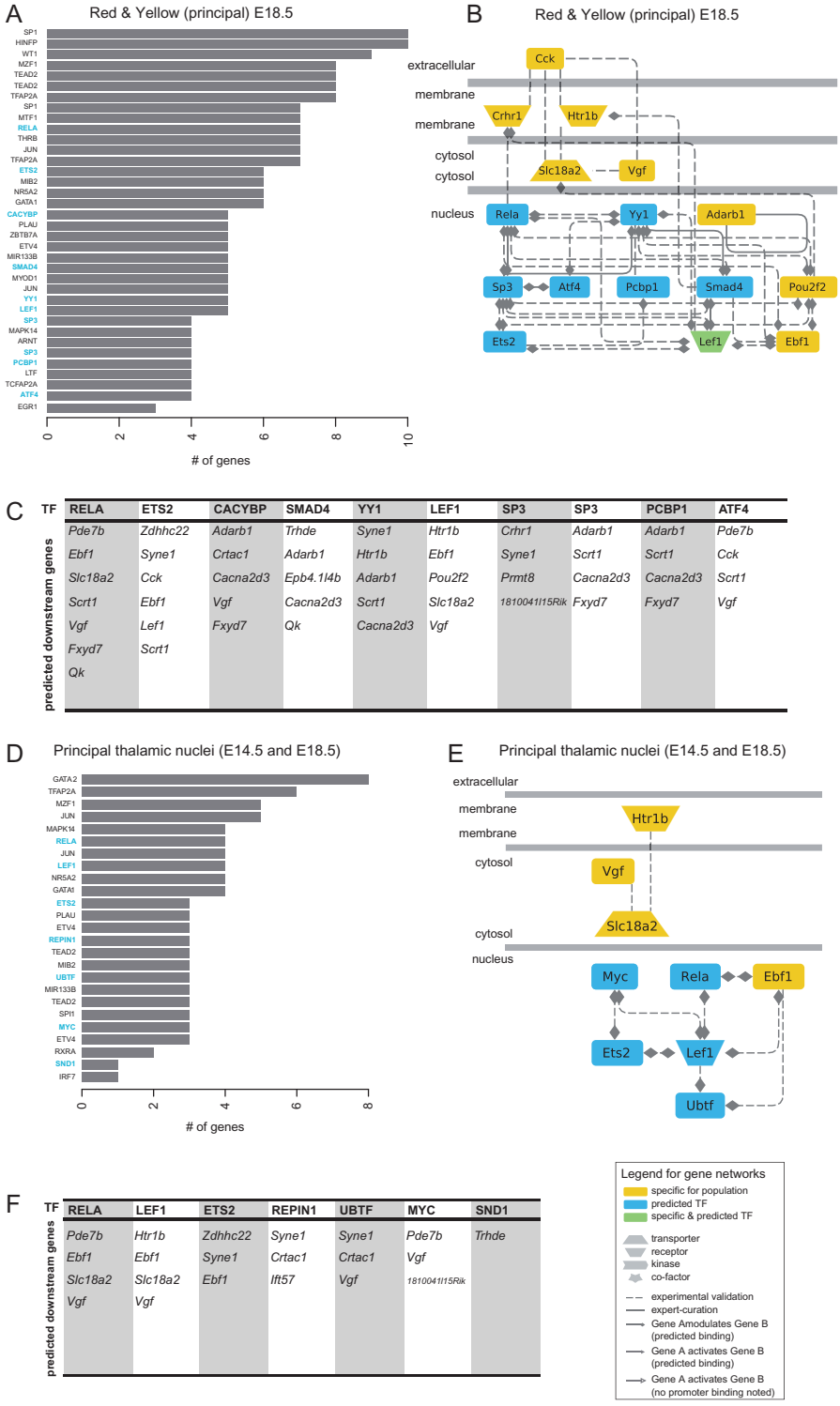

Figure 6. TF binding motifs and gene interaction networks in principal thalamic neuronal populations. (*A*) Bar plot showing predicted TFs at promoters of the genes specifically expressed in principal dLGN/VB/MGN (red and yellow) populations at E18.5. The *x*-axis shows the number of genes that were found to have a motif for the respective factor on *y*-axis. Only those factors that were found to be significant ($P < 0.05$) are plotted. TFs expressed in this population according to the microarray analysis are highlighted in blue. (*B*) A gene network created by using the significant TFs and their potential targets for this population. (*C*) TF predicted by motifs analysis of genes enriched in the population (top row) and the downstream genes having each motif. Only TFs that are expressed within the population and their targets are shown. (*D*) Bar plot showing predicted TFs at promoters of the genes specifically expressed in principal thalamic nuclei at both E14.5 and E18.5. (*E*) Gene network using the expressed significant TFs and their potential targets for principal thalamic nuclei at both E14.5 and E18.5. (*F*) TF predicted by motifs analysis of genes enriched in the population (top row) and the downstream genes having each motif. Only TFs that are expressed in to principal (yellow and red) populations at both E14.5 and E18.5 and their targets are shown.

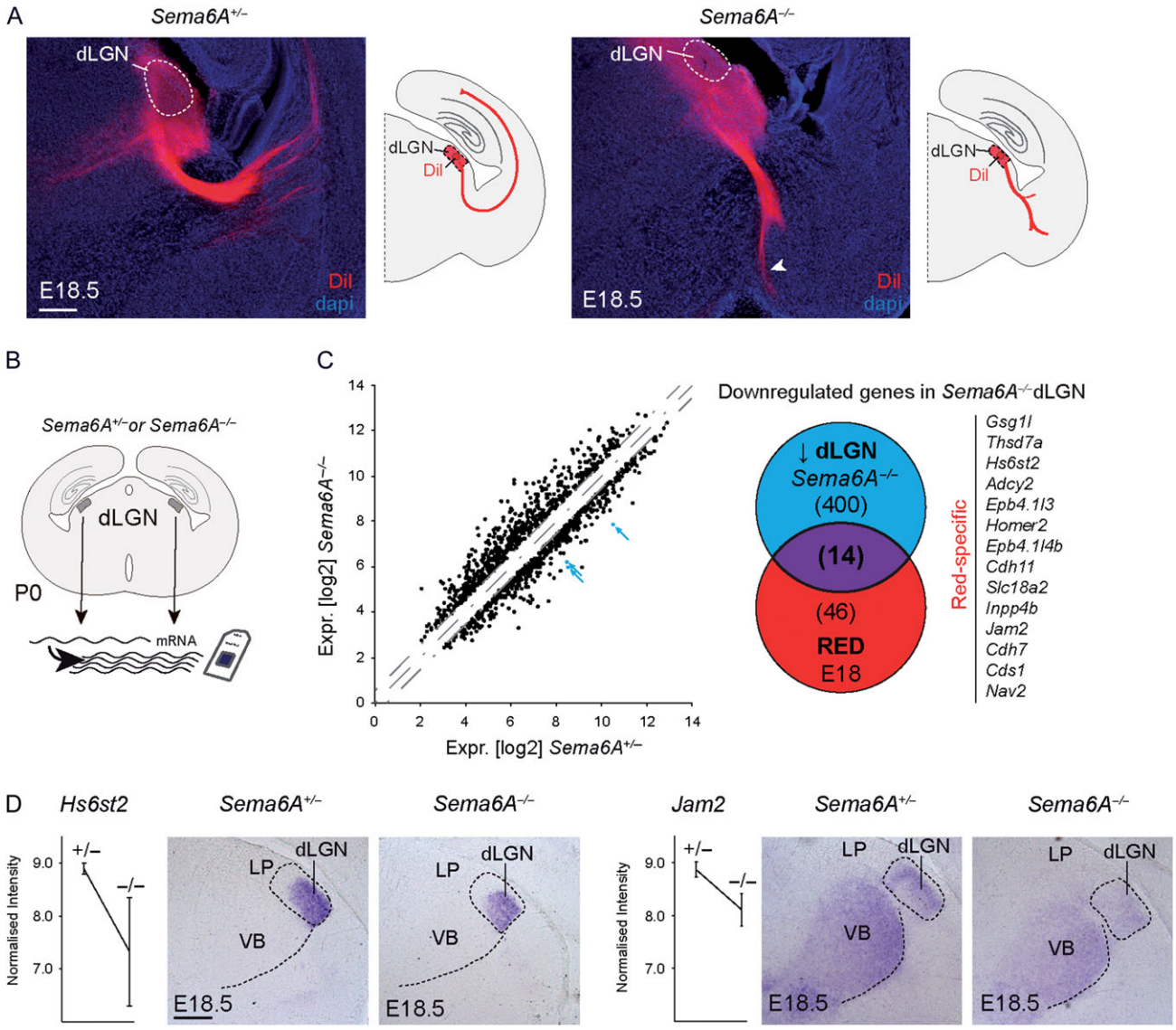

Figure 7. Misguided dLGN axons in Sema6A mutants concur with changes in nuclei-specific gene expression. (*A*) Visual dLGN axons are derailed in a *Sema6A*$^{-/-}$ in ventral telencephalon at embryonic day E18.5. Placements of small DiI crystals in the dLGN show the normal projection to V1 cortex in control but misguided axons in *Sema6A*$^{-/-}$ mutant embryos. (*B*) The dLGN dissected out from fresh sections for genome-wide analysis of gene expression. (*C*) Scatter plots of transcripts differentially expressed between control and *Sema6A*$^{-/-}$ animals. Only transcripts with FC > 1.5 ($P < 0.05$) are plotted. The blue arrows indicate the probes of the *Sema6A* transcripts. Venn diagram of overlap between downregulated genes in dLGN and genes enriched in the red E18.5 population (dLGN+VB from Fig. 2). The 14 red-specific genes that were downregulated in the dLGN of *Sema6A*$^{-/-}$ are listed to the right. (*D*) Relative expression levels revealed by normalized intensity [log2] values from the microarray and in situ hybridization verifying the reduced expression of *Hs6st2* and *Jam2* in the thalamus of *Sema6A*$^{-/-}$ embryos. Scale bars represents 250 µm.

Marcos-Mondéjar et al. 2012), were enriched in these neurons in our transgenic mice. Moreover, genes with expression restricted to each principal sensory nucleus or developmentally restricted to principal sensory thalamic nuclei were revealed suggesting that these genes might play important roles in specifying thalamic neurons to connect with primary cortical areas.

Previous efforts have provided a number of genes that could be important for TC development or TC nuclei specification (Horng et al. 2009; Suzuki-Hirano et al. 2011; Yuge et al. 2011; Price et al. 2012). For example, a microarray screen of principal thalamic nuclei in neonatal mice was able to unravel, for the first time, genes with a restrictive expression in visual dLGN and auditory MGN thalamic neurons (Horng et al. 2009). Although this study provided the evidence for a nuclei-specific gene expression pattern in the thalamus, the fact that the analysis was performed postnatally limits the possible implication of these genes in early development. At the postnatal stages analysed, TC axons have already targeted and invade their corresponding cortical target areas and thus, genes involved in early nuclei specification and/or axon guidance might already have been downregulated. On the contrary, our 2-stage embryonic analysis of nuclei-specific TC genes might provide a relevant list of potential candidates genes involved in early thalamic neuronal specification. Along this line, using in situ hybridization, the laboratory of Shimogori has provided a list of several genes potentially involved in the early specification of thalamic nuclei (Suzuki-Hirano et al. 2011; Yuge et al. 2011), all stemming from a previous work on transcriptional profiling of the developing hypothalamus (Shimogori 2010). Consistently with this, some of the genes we found in our

unbiased screen were also described in this study to have a nuclei-specific pattern of expression, as for example *Slitrk6, Drd2, Kitl (Steel), Cdh6*, and *Ror* α in the population of principal neurons at E14.5, and Shox2, in the nonprincipal nuclei at E14.5 and at E18.5.

In addition to published studies, the extensive expression databases such as the Allen Brain Atlases (ABA; http://brain-map.org; Lein et al. 2007), the GenePaint (Visel et al. 2004), or the BGEM (Magdaleno et al. 2006) are good tools for looking at expression patterns of genes within the developing thalamus. Further, in silico methods using these databases have identify nuclei-specific genes in the adult thalamus or combinations of genes expressed in the developing thalamic neurons (Price et al. 2012; Nagalski et al. 2016). However, although possible for the adult, these databases fail to provide new genes expressed a priori in restricted embryonic thalamic nuclei.

Here, we developed a genetic strategy to isolate murine principal thalamic nuclei genes at distinct embryonic stages using the *Gbx2* promoter. Following a genome-wide analysis, we identified genes differently expressed between visual dLGN, somatosensory VB, and auditory MGN nuclei. Furthermore, taking advantage of the genetic labeling strategy following *Gbx2* expression, we could for the first time, lineage trace and unravel developmentally enriched principal thalamic nuclei genes as well as determine the gene profile for the 3 principal sensory-related thalamic nuclei. All cells labeled and analysed in this study are derived from *Gbx2*-progeny and thus are thalamic neurons (Chen et al. 2009; Li et al. 2012). As some of the genes with expression restricted to a single nucleus at E18.5 already showed specific patterns of expression at E14.5, it is tempting to speculate that these genes could be specifying and further be used as markers of those nuclei, already before they are anatomically distinguishable and have connected to their corresponding cortical area. Interestingly, some of the genes we found enriched in specific populations are expressed in the thalamus already at E12.5 in a *Gbx2*-dependent fashion (Mallika et al. 2015).

In our study we confirmed a number of genes, such as *Hs6st2* and *Crabp2,* whose expression was previously described as enriched in the dLGN or MGN, respectively, at neonatal stages (Horng et al. 2009) and already present at E12.5 (Mallika et al. 2015). We also found novel genes with an enriched expression in a specific sensory modality-related thalamic nucleus. The heparan sulfate 6-O-sulphotransferases *Hs6st2* and *Hs6st3* were expressed in the dLGN. Interestingly 2 orthologs, *Hs2st* and *Hs6st1,* are known to be important for axon guidance of retinal axons and corpus callosum development (Pratt et al. 2006; Conway, Howe, et al. 2011), raising the possibility that these molecules could also play a role in the connectivity of dLGN axons to the visual cortex. In addition, *Jam2* that was observed to be enriched in the dLGN and VB (red) populations at E18.5, has been implicated in tight junctions of endothelial cells (Weber et al. 2007). Whether this gene may also play a role in neuronal development or specification of axonal connectivity remains to be elucidated. Remarkably, we found that *Hs6st2* and *Jam2* are 2 of the genes significantly reduced in the misguided dLGN neurons in the *Sema6A*$^{-/-}$ suggesting that these genes may play a role in TC connectivity.

Another gene that we found to be specific for the dLGN and VB (red) populations is cholecystokinin (*Cck*). This neuropeptide has been previously described to be expressed in the rodent thalamus among other brain regions (Schiffmann and Vanderhaeghen 1991). However, while in the adult *Cck* is expressed in many thalamic nuclei, we found it to be specific for the visual and somatosensory thalamic nuclei at E18.5, opening the possibility of a function of *Cck* in the development and/or cortical targeting of these neurons. Indeed, CCK has the ability to depolarize the thalamic recipient layer 6 neurons (Chung et al. 2009) and also has an effect on intrathalamic oscillations (Cox et al. 1997). We previously demonstrated that developmental changes in spontaneous activity in thalamic neurons determine TC axonal extension during development (Mire et al. 2012). Thus, it is possible that *Cck* could play a role in the modulation of this spontaneous activity in the embryonic thalamus, and thus, influence TC pathfinding.

Regarding the auditory MGN neurons, we found *Crabp2* and *Tshz1* genes to be specifically expressed in this nucleus, extending previous finding (Horng et al. 2009). The role of either *Crabp2* or *Tshz1* in TC development has not been addressed before. Surprisingly, we found that *Ascl1* is enriched in the yellow MGN population at E18.5, as well as, in the red population at E14.5. This is unexpected since this gene is well known to be expressed in domains of progenitor cells of the diencephalon that later will form inhibitory neurons (Vue et al. 2007). *Ascl1* is a TF known to be important for TC axon pathfinding to the cortex (Tuttle et al. 1999); however, the TC axon guidance defect seen in the *Ascl1* null mice has so far been attributed to a non-cell autonomous role of this gene due to early changes in the patterning of ventral thalamic territories (Tuttle et al. 1999). Thus, our results open the possibility that *Ascl1* also may play a cell-autonomous role in the specification and/or TC guidance of auditory thalamic neurons. Finally, our strategy allowed us to identify a number of genes explicitly common to the principal sensory thalamic nuclei (red and yellow) populations at E18.5, exemplified by *Lef1*, *Pou2f2*, and *Ebf1*. Some of these genes where previously known to be expressed in the thalamus, such as *Lef1* (Jones and Rubenstein 2004), while others like *Pou2f2* are for the first time identified as being unique of principal thalamic nuclei. Interestingly, both *Ebf1* and *Pou2f2* have binding motifs for *Lef1* in their promoter regions raising the possibility of *Lef1* being a key TF for principal thalamic nuclei specification, as recently also suggested (Nagalski et al. 2016).

Screening the promoter regions of the genes found revealed additional putative important TFs. We identified *Jun* specifically expressed in the dLGN and VB (red) populations at E18.5 and that regulates several genes specific for these neurons. Interestingly, *Jun* is an activity dependent gene, which suggest that spontaneous embryonic neuronal activity might be an important feature for setting the transcriptional profile of the principal thalamic populations. Moreover, our motif analysis revealed potential novel interactions. For example, in the promoter region of the dLGN-specific gene *Sp9* there is a motif for POU2F2 or the MGN-specific gene *Tshz1* there is a motif for SMAD4. Unfortunately, a limitation in our bioinformatics analyses is that the database used to screen these motifs did not contain any consensus binding sequence for Gbx2, and thus, the genes with binding sites for this TF could not be determined. In an attempt to find genes regulated by Gbx2, a recent study performed a microarray analysis in a *Gbx2* knockout E12.5 thalamus and identified several genes with a Gbx2-dependent expression (Mallika et al. 2015). Interestingly, some of these genes we found to be persistently expressed in principal thalamic neurons such as *Hs6st2*, *Slc18a2*, and *Vegfc*. Still, a possible *Gbx2* regulation might be indirect. Furthermore, it should be taken into consideration that most of the microarray analysis performed at early embryonic brain stages, as the one we did here, are based on low input RNA samples. This might represent a technical limitation as variability is inherited and

could generate undesired false positive and false negative results. Future studies should thus be focused in experimentally validating these and others predicted interactions and thus also determine the role of each TF as activator or repressor of the respective target gene. More interesting still would be to determine the interactions and combination of TFs needed to specify neurons of each thalamic nucleus. Reaching this point of knowledge would make it possible to reveal how TFs specify certain features of the thalamic cells and their axonal pathfinding to reach the correct cortical area.

Our study support that thalamic nuclei specification takes place intrinsically before anatomical boundaries are present in the thalamus and TC neurons are connected by either afferent input or to their final target areas. However, a given transcriptional profile could also be influence by a correct axonal targeting. We used the *Sema6A*$^{-/-}$ mouse to address this question. Indeed, several of the genes identified enriched in the dLGN and VB populations at E18.5 have a reduced expression level in the dLGN of *Sema6A*$^{-/-}$ when compared with controls. Intriguingly, nearly all specific genes downregulated in the dLGN of the *Sema6A*$^{-/-}$ are membrane associated and include several adhesion molecules, which suggests a bidirectional interaction between the TC and cortical neurons involved in setting the final synaptic connectivity of the target cortical area. One alternative explanation of the altered expression could be through a cell autonomous effect of the Sema6A protein, though there is no evidence of such mechanism for this semaphorin (Suto et al. 2005). However, the importance of a correct targeting for maintenance of specific TC neuronal connectivity has been recently described (Zembrzycki et al. 2013). Thus, the transcriptome of a nucleus, though largely set intrinsically, is also to some extent affected by the neuron's connectivity.

In sum, here we present genes and networks plausibly important for the specification, development and/or connectivity of the distinct thalamic neuronal populations. Moreover, the nuclei-specific novel genes revealed here will help to provide new tools, like Cre-mice, for in vivo targeting of specific sensory-related thalamic nuclei. Such mice would be key both for linage-tracing studies and also in combination with the rapidly expanding tool-box of optogenetic (Deisseroth 2015) and drug-induced genetic manipulations (Roth 2016) that will allow for functional studies of with high precision.

Supplementary Material

Supplementary material can be found at: http://www.cercor.oxfordjournals.org/.

Authors' Contributions

H.G. and G.L.-B. conceived the project and wrote the manuscript. H.G. designed and performed the conditional labeling of thalamic nuclei, FACS experiments and analysed the data. V.M.-J. and C.M. performed the microarray experiments in the *Sema6A* mutant mouse; G.L.-B. and V.M.-J. performed and analysed the tracing experiments in the *Sema6A* knockout mouse. S.T. and V.T. performed the bioinformatics analyses. L.M.R.-M. performed the in situ hybridization experiments. J.P. analysed microarray data (1st level analysis) with V.B.'s supervision. G.L.-B. supervised the project. All authors revised and approved the final manuscript.

Funding

H.G. held postdoc fellowships from the Swedish Research Council and the Swedish Brain Foundation. Supported by the European Commission Grants ERC-2009-StG_20081210 and ERC-2014-CoG 647012 and the Spanish MINECO BFU2012-34298 and BFU2015-64432-R. G.L.-B. is an EMBO YIP Investigator and a FENS-Kavli Scholar. The Instituto de Neurociencias is a "Centre of Excellence Severo Ochoa".

Notes

We are grateful to Prof. James Li for sharing the *Gbx2*^{CreER} mice and various colleagues for providing clones for making in situ probes stated in Table 1. We thank Prof. Paola Arlotta for advice on FACS protocol, Dr Denis Jabaudon for guidance on thalamic RNA extraction as well as Rafael Susin and Antonio Caler Escribano for technical assistance. We are thankful Prof. Esther Stoeckli for constructive comments on the manuscript as well as to members of the G. López-Bendito laboratory for stimulating discussions and comments. The NCBI GEO accession numbers for the microarray data reported in this paper are GSE68450 and GSE79683. *Conflict of Interest*: None declared.

References

Angevine JB Jr. 1970. Time of neuron origin in the diencephalon of the mouse. An autoradiographic study. J Comp Neurol. 139:129–187.

Bach I, Rodriguez-Esteban C, Carriere C, Bhushan A, Krones A, Rose DW, Glass CK, Andersen B, Izpisua Belmonte JC, Rosenfeld MG. 1999. RLIM inhibits functional activity of LIM homeodomain transcription factors via recruitment of the histone deacetylase complex. Nat Genet. 22:394–399.

Catapano LA, Arnold MW, Perez FA, Macklis JD. 2001. Specific neurotrophic factors support the survival of cortical projection neurons at distinct stages of development. J Neurosci. 21:8863–8872.

Caubit X, Core N, Boned A, Kerridge S, Djabali M, Fasano L. 2000. Vertebrate orthologues of the Drosophila region-specific patterning gene teashirt. Mech Dev. 91:445–448.

Chen L, Guo Q, Li JY. 2009. Transcription factor Gbx2 acts cell-nonautonomously to regulate the formation of lineage-restriction boundaries of the thalamus. Development. 136:1317–1326.

Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, Clark NR, Ma'ayan A. 2013. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics. 14:128.

Chung L, Moore SD, Cox CL. 2009. Cholecystokinin action on layer 6b neurons in somatosensory cortex. Brain Res. 1282:10–19.

Conway CD, Howe KM, Nettleton NK, Price DJ, Mason JO, Pratt T. 2011. Heparan sulfate sugar modifications mediate the functions of slits and other factors needed for mouse forebrain commissure development. J Neurosci. 31:1955–1970.

Conway CD, Price DJ, Pratt T, Mason JO. 2011. Analysis of axon guidance defects at the optic chiasm in heparan sulphate sulphotransferase compound mutant mice. J Anat. 219:734–742.

Cox CL, Huguenard JR, Prince DA. 1997. Peptidergic modulation of intrathalamic circuit activity in vitro: actions of cholecystokinin. J Neurosci. 17:70–82.

Deisseroth K. 2015. Optogenetics: 10 years of microbial opsins in neuroscience. Nat Neurosci. 18:1213–1225.

Ebisu H, Iwai-Takekoshi L, Fujita-Jimbo E, Momoi T, Kawasaki H. 2016. Foxp2 regulates identities and projection patterns of thalamic nuclei during development. Cereb Cortex. Advance Access published July 6, 2016. doi: 10.1093/cercor/bhw187.

Enjin A, Rabe N, Nakanishi ST, Vallstedt A, Gezelius H, Memic F, Lind M, Hjalt T, Tourtellotte WG, Bruder C, et al. 2010. Identification of novel spinal cholinergic genetic subtypes disclose Chodl and Pitx2 as markers for fast motor neurons and partition cells. J Comp Neurol. 518:2284–2304.

Erzurumlu RS, Gaspar P. 2012. Development and critical period plasticity of the barrel cortex. Eur J Neurosci. 35:1540–1553.

Galcerán J, Miyashita-Lin EM, Devaney E, Rubenstein JL, Grosschedl R. 2000. Hippocampus development and generation of dentate gyrus granule cells is regulated by LEF1. Development. 127:469–482.

Garel S, López-Bendito G. 2014. Inputs from the thalamocortical system on axon pathfinding mechanisms. Curr Opin Neurobiol. 27:143–150.

Garel S, Yun K, Grosschedl R, Rubenstein JL. 2002. The early topography of thalamocortical projections is shifted in Ebf1 and Dlx1/2 mutant mice. Development. 129:5621–5634.

Geschwind DH, Rakic P. 2013. Cortical evolution: judge the brain by its cover. Neuron. 80:633–647.

Grove EA, Fukuchi-Shimogori T. 2003. Generating the cerebral cortical area map. Annu Rev Neurosci. 26:355–380.

Haase G, Dessaud E, Garces A, de Bovis B, Birling M, Filippi P, Schmalbruch H, Arber S, deLapeyriere O. 2002. GDNF acts through PEA3 to regulate cell body positioning and muscle innervation of specific motor neuron pools. Neuron. 35: 893–905.

Habuchi H, Tanaka M, Habuchi O, Yoshida K, Suzuki H, Ban K, Kimata K. 2000. The occurrence of three isoforms of heparan sulfate 6-O-sulfotransferase having different specificities for hexuronic acid adjacent to the targeted N-sulfoglucosamine. J Biol Chem. 275:2859–2868.

Hahm S, Mizuno TM, Wu TJ, Wisor JP, Priest CA, Kozak CA, Boozer CN, Peng B, McEvoy RC, Good P, et al. 1999. Targeted deletion of the Vgf gene indicates that the encoded secretory peptide precursor plays a novel role in the regulation of energy balance. Neuron. 23:537–548.

Horng S, Kreiman G, Ellsworth C, Page D, Blank M, Millen K, Sur M. 2009. Differential gene expression in the developing lateral geniculate nucleus and medial geniculate nucleus reveals novel roles for Zic4 and Foxp2 in visual and auditory pathway development. J Neurosci. 29:13672–13683.

Jones EG, Rubenstein JL. 2004. Expression of regulatory genes during differentiation of thalamic nuclei in mouse and monkey. J Comp Neurol. 477:55–80.

Kim IJ, Zhang Y, Yamagata M, Meister M, Sanes JR. 2008. Molecular identification of a retinal cell type that responds to upward motion. Nature. 452:478–482.

Komine Y, Takao K, Miyakawa T, Yamamori T. 2012. Behavioral abnormalities observed in Zfhx2-deficient mice. PLoS One. 7:e53114.

Lein ES, Hawrylycz MJ, Ao N, Ayres M, Bensinger A, Bernard A, Boe AF, Boguski MS, Brockway KS, Byrnes EJ, et al. 2007. Genome-wide atlas of gene expression in the adult mouse brain. Nature. 445:168–176.

Li K, Zhang J, Li JY. 2012. Gbx2 plays an essential but transient role in the formation of thalamic nuclei. PLoS One. 7: e47111.

Lin JH, Saito T, Anderson DJ, Lance-Jones C, Jessell TM, Arber S. 1998. Functionally related motor neuron pool and muscle sensory afferent subtypes defined by coordinate ETS gene expression. Cell. 95:393–407.

Little GE, López-Bendito G, Runker AE, Garcia N, Pinon MC, Chedotal A, Molnar Z, Mitchell KJ. 2009. Specificity and plasticity of thalamocortical connections in Sema6A mutant mice. PLoS Biol. 7:e98.

Long JE, Swan C, Liang WS, Cobos I, Potter GB, Rubenstein JL. 2009. Dlx1&2 and Mash1 transcription factors control striatal patterning and differentiation through parallel and overlapping pathways. J Comp Neurol. 512:556–572.

López-Bendito G, Molnar Z. 2003. Thalamocortical development: how are we going to get there? Nat Rev Neurosci. 4: 276–289.

Madisen L, Zwingman TA, Sunkin SM, Oh SW, Zariwala HA, Gu H, Ng LL, Palmiter RD, Hawrylycz MJ, Jones AR, et al. 2010. A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. Nat Neurosci. 13:133–140.

Magdaleno S, Jensen P, Brumwell CL, Seal A, Lehman K, Asbury A, Cheung T, Cornelius T, Batten DM, Eden C, et al. 2006. BGEM: an in situ hybridization database of gene expression in the embryonic and adult mouse nervous system. PLoS Biol. 4:e86.

Mallika C, Guo Q, Li JY. 2015. Gbx2 is essential for maintaining thalamic neuron identity and repressing habenular characters in the developing thalamus. Dev Biol. 407:26–39.

Marcos-Mondéjar P, Peregrin S, Li JY, Carlsson L, Tole S, López-Bendito G. 2012. The lhx2 transcription factor controls thalamocortical axonal guidance by specific regulation of robo1 and robo2 receptors. J Neurosci. 32:4372–4385.

Mire E, Mezzera C, Leyva-Diaz E, Paternain AV, Squarzoni P, Bluy L, Castillo-Paterna M, Lopez MJ, Peregrin S, Tessier-Lavigne M, et al. 2012. Spontaneous activity regulates Robo1 transcription to mediate a switch in thalamocortical axon growth. Nat Neurosci. 15:1134–1143.

Mitchell KJ, Pinson KI, Kelly OG, Brennan J, Zupicich J, Scherz P, Leighton PA, Goodrich LV, Lu X, Avery BJ, et al. 2001. Functional analysis of secreted and transmembrane proteins critical to mouse development. Nat Genet. 28:241–249.

Nagalski A, Puelles L, Dabrowski M, Wegierski T, Kuznicki J, Wisniewska MB. 2016. Molecular anatomy of the thalamic complex and the underlying transcription factors. Brain Struct Funct. 221:2493–2510.

Nakagawa Y, O'Leary DD. 2001. Combinatorial expression patterns of LIM-homeodomain and other regulatory genes parcellate developing thalamus. J Neurosci. 21:2711–2725.

Nakagawa Y, Shimogori T. 2012. Diversity of thalamic progenitor cells and postmitotic neurons. Eur J Neurosci. 35:1554–1562.

Palmeri D, van Zante A, Huang CC, Hemmerich S, Rosen SD. 2000. Vascular endothelial junction-associated molecule, a novel member of the immunoglobulin superfamily, is localized to intercellular boundaries of endothelial cells. J Biol Chem. 275:19139–19145.

Pratt T, Conway CD, Tian NM, Price DJ, Mason JO. 2006. Heparan sulphation patterns generated by specific heparan sulfotransferase enzymes direct distinct aspects of retinal axon guidance at the optic chiasm. J Neurosci. 26:6911–6923.

Price DJ, Clegg J, Duocastella XO, Willshaw D, Pratt T. 2012. The importance of combinatorial gene expression in early mammalian thalamic patterning and thalamocortical axonal guidance. Front Neurosci. 6:37.

Rakic P. 1988. Specification of cerebral cortical areas. Science. 241:170–176.

Roth BL. 2016. DREADDs for neuroscientists. Neuron. 89: 683–694.

Rubenstein JL, Martínez S, Shimamura K, Puelles L. 1994. The embryonic vertebrate forebrain: the prosomeric model. Science. 266:578–580.

Ruberte E, Friederich V, Morriss-Kay G, Chambon P. 1992. Differential distribution patterns of CRABP I and CRABP II transcripts during mouse embryogenesis. Development. 115:973–987.

Schiffmann SN, Vanderhaeghen JJ. 1991. Distribution of cells containing mRNA encoding cholecystokinin in the rat central nervous system. J Comp Neurol. 304:219–233.

Seibt J, Schuurmans C, Gradwhol G, Dehay C, Vanderhaeghen P, Guillemot F, Polleux F. 2003. Neurogenin2 specifies the connectivity of thalamic neurons by controlling axon responsiveness to intermediate target cues. Neuron. 39:439–452.

Shimogori T, Lee DA, Miranda-Angulo A, Yang Y, Wang H, Jiang L, Yoshida AC, Kataoka A, Mashiko H, Avetisyan M, et al. 2010. A genomic atlas of mouse hypothalamic development. Nat Neurosci. 13:767–775.

Song H, Lee B, Pyun D, Guimera J, Son Y, Yoon J, Baek K, Wurst W, Jeong Y. 2015. Ascl1 and Helt act combinatorially to specify thalamic neuronal identity by repressing Dlxs activation. Dev Biol. 398:280–291.

Suto F, Ito K, Uemura M, Shimizu M, Shinkawa Y, Sanbo M, Shinoda T, Tsuboi M, Takashima S, Yagi T, et al. 2005. Plexin-a4 mediates axon-repulsive activities of both secreted and transmembrane semaphorins and plays roles in nerve fiber guidance. J Neurosci. 25:3628–3637.

Suzuki-Hirano A, Ogawa M, Kataoka A, Yoshida AC, Itoh D, Ueno M, Blackshaw S, Shimogori T. 2011. Dynamic spatiotemporal gene expression in embryonic mouse thalamus. J Comp Neurol. 519:528–543.

Tuttle R, Nakagawa Y, Johnson JE, O'Leary DD. 1999. Defects in thalamocortical axon pathfinding correlate with altered cell domains in Mash-1-deficient mice. Development. 126: 1903–1916.

Visel A, Thaller C, Eichele G. 2004. GenePaint.org: an atlas of gene expression patterns in the mouse embryo. Nucleic Acids Res. 32:D552–D556.

Vrieseling E, Arber S. 2006. Target-induced transcriptional control of dendritic patterning and connectivity in motor neurons by the ETS gene Pea3. Cell. 127:1439–1452.

Vue TY, Aaker J, Taniguchi A, Kazemzadeh C, Skidmore JM, Martin DM, Martin JF, Treier M, Nakagawa Y. 2007. Characterization of progenitor domains in the developing mouse thalamus. J Comp Neurol. 505:73–91.

Weber C, Fraemohs L, Dejana E. 2007. The role of junctional adhesion molecules in vascular inflammation. Nat Rev Immunol. 7:467–477.

Yuge K, Kataoka A, Yoshida AC, Itoh D, Aggarwal M, Mori S, Blackshaw S, Shimogori T. 2011. Region-specific gene expression in early postnatal mouse thalamus. J Comp Neurol. 519:544–561.

Zembrzycki A, Chou SJ, Ashery-Padan R, Stoykova A, O'Leary DD. 2013. Sensory cortex limits cortical maps and drives top-down plasticity in thalamocortical circuits. Nat Neurosci. 16: 1060–1067.

Zhou W, Zhu H, Zhao J, Li H, Wan Y, Cao J, Zhao H, Yu J, Zhou R, Yao Y, et al. 2013. Misexpression of Pknox2 in mouse limb bud mesenchyme perturbs zeugopod development and deltoid crest formation. PLoS One. 8:e64237.