

SOX2 and NR2F1 coordinate the gene expression program of the early postnatal visual thalamus

Linda Serra^{1*}, Anna Nordin^{2,3*}, Mattias Jonasson^{2,3}, Carolina Marenco¹, Francesca Gullo¹, Sergio Ottolenghi¹, Federico Zambelli⁴, Michèle Studer⁵, Giulio Pavesi⁴, Claudio Cantù^{2,3§}, Silvia K. Nicolis^{1§} and Sara Mercurio^{1§}

¹Department of Biotechnology and Biosciences, University of Milano-Bicocca, piazza della Scienza 2, 20126 Milano, Italy

²Wallenberg Centre for Molecular Medicine, Linköping University, Linköping, Sweden

³Department of Biomedical and Clinical Sciences, Division of Molecular Medicine and Virology, Faculty of Medicine and Health Sciences, Linköping University, Linköping, Sweden

⁴Department of Biosciences, University of Milano, Milan, Italy.

⁵Institute of Biology Valrose (IBV), University Côte d'Azur, 06108 Nice, France.

§corresponding authors: sara.mercurio@unimib.it, silvia.nicolis@unimib.it, claudio.cantu@liu.se

*joint first authors

Abstract

The thalamic dorsolateral geniculate nucleus, (dLGN) receives visual input from the retina via the optic nerve, and projects to the cortical visual area, where eye-derived signals are elaborated. The transcription factors SOX2 and NR2F1 are directly involved in the differentiation of dLGN neurons, based on mouse work and patient mutations leading to vision defects. However, whether they regulate each other, or control common targets is still unclear. By RNA-seq analysis of neonatal dLGN from thalamo-specific *Sox2* and *Nr2f1* mouse mutants, we found a striking overlap of deregulated genes. Among them, VGF, a cytokine transported along thalamic-cortical axons is strongly downregulated in both mutants. CUT&RUN analysis of SOX2 binding in dLGN chromatin identified a binding pattern characteristic of the dLGN. Collectively, the SOX2 and NR2F1-coregulated genes, and cognate SOX2 binding sites, contribute as a basis to understand the gene regulatory network driving the differentiation and connectivity of thalamic neurons.

Keywords: Sox2, Nr2f1, vision, thalamus, dLGN, CUT&RUN, Vgf

Introduction

Inherited diseases affecting vision importantly contribute to significant disability in human populations. Most frequently, the retina and the eye are involved, but other parts of the visual axis, such as thalamic nuclei and visual cortical areas can also be affected¹⁻³. A wide variety of genes are known to be mutated in such diseases, ranging from transcription factors to cell membrane receptors, to secreted signalling molecules, to second messengers in signal transduction, etc^{1,2}. Interestingly, transcription factors expressed during early development with pleiotropic functions are found to be mutated in specific cases, which raises the issue of which genes among their targets are functionally important to cause the pathological phenotype. Studies in humans and mice show that mutations of two genes, *NR2F1* and *SOX2* cause important vision defects^{1,4,5}: *NR2F1* mutations are responsible for the Bosch-Boonstra-Schaaf optic atrophy syndrome (OMIM #615722; BBSOAS), whereas *SOX2* mutations cause micro- or anophthalmia (OMIM #206900, Microphthalmia, syndromic 3; Optic nerve hypoplasia and abnormalities of the central nervous system). Thalamo-specific conditional knockout (cKO) mouse mutants show that inactivation of *Sox2* or *Nr2f1* affects, albeit at different degrees of severity, proper differentiation of the thalamic dorsolateral geniculate nucleus (dLGN) and the visual cortex, also causing alterations of the topographic connections relaying information to and from these functional locations⁶⁻¹⁰. The dLGN nucleus is crucial for conveying information from the retina to the primary visual cortex. Mouse mutants, conditionally deleted in the dLGN for either *Sox2* or *Nr2f1* using a ROR α -Cre transgene, result in hypomorphic visual thalamus postnatally, reduced thalamic projections to the primary visual cortex^{7,9}, significant abnormalities of the visual cortex, and important alterations of retino-thalamic connections⁹. We reasoned that dysregulation of a common set of target genes of *NR2F1* and *SOX2* might be the underlying factor causing phenotypic similarity following their thalamic (dLGN) inactivation. If this is correct, it might point to a gene set universally shared during development of the dLGN and the visual axis, as well as common to the visual phenotype of *Nr2f1* and *Sox2* mutants.

In the present work, we obtained RNA-seq data from wild-type as well as *Nr2f1* or *Sox2* dLGN mutants and identified over 500 dysregulated genes in common. We show that most of the top downregulated genes in *Nr2f1* mutants are also significantly downregulated in *Sox2* mutants (44 out of 50); likewise, the large majority of the top downregulated genes in *Sox2* mutants are significantly downregulated also in *Nr2f1* mutants. These results suggest that many of the downregulated genes identified in this study contribute to at least some of the phenotypic alterations observed in mouse mutants. Functional enrichment analysis showed that deregulated genes are highly enriched in differentiated neuronal functions (axon guidance molecules, synaptic proteins, etc.). Deconvolution analysis of RNA-seq data, based on the comparison with single-cell RNA-seq data previously obtained on wild-type visual thalamus¹¹, revealed a substantial reduction in cells with a transcriptional identity of glutamatergic neurons already at early developmental stages, preceding phenotypic defects. Moreover, we employed the Cleavage Under Targets and Release Using Nuclease (CUT&RUN;¹²) approach to identify the *in vivo* genome-wide direct SOX2 binding sites in visual thalamic nuclei. This analysis allowed us to establish: i) the first full set of SOX2 binding sites in functionally relevant differentiated neurons, ii) the subset of targets that are transcriptionally regulated by SOX2, and iii) those that likely depend on the interplay between SOX2 and NRF1. Finally, we also identified a general SOX/NRF binding consensus sequence characteristic of the dLGN that might be used by several other transcription factors involved in the differentiation of thalamic neurons.

Results

***Sox2* and *Nr2f1* deletion in the developing visual thalamus causes deregulation of gene expression that precedes thalamic defects**

To unravel the origin of the common phenotype observed in *Sox2* and *Nr2f1* thalamo-specific mouse mutants, we first stained the dLGN at postnatal day 0 (P0) (Fig. 1A) with anti-SOX2 and anti-NR2F1 antibodies and observed high co-expression of the two proteins in thalamic differentiated neurons (Fig. 1A). Then, to identify the gene regulatory network downstream of SOX2 and NR2F1 in the visual thalamus, we performed RNA-seq experiments on *ex vivo* dissected dLGN from *Sox2* or *Nr2f1* thalamic mutants (obtained via ROR α -Cre deletion) and their control littermates at postnatal day 0 (P0), before the appearance of overt morphological impairments^{7,9} (Fig. 1B). Three independent pools of mutant and control dissected visual thalami for each mutant line were processed. We identified 2284 differentially expressed genes (DEG) with FDR < 0.01 following *Sox2* conditional inactivation, of which 1002 downregulated and 1282 upregulated (Fig. 1C; Table S1). In addition, with the same thresholds, we identified 1081 genes differentially expressed following *Nr2f1* conditional inactivation, of which 493 genes downregulated and 588 upregulated (Fig. 1C; Table S1). Notably, neither is *Nr2f1* dysregulated in *Sox2* mutant dLGN, nor is *Sox2* significantly dysregulated in the *Nr2f1* mutant, ruling out mutual regulation of the two genes (Table S1, see also ⁹). This was also confirmed by immunofluorescence showing no changes in the number and distribution of NR2F1-expressing cells in the absence of SOX2 (Fig. 1A), suggesting that the two proteins likely do not regulate each other, but control a common set of genes.

Many genes are co-regulated by SOX2 and NR2F1

We previously showed that important effectors of SOX2 function in NSC self-renewal and differentiation were among the most highly expressed and the most highly down-regulated genes in *Sox2*-mutated cells¹³⁻¹⁵. Importantly, genes dysregulated in *Sox2* mutant dLGN strikingly differed from those dysregulated in NSC¹³. We thus focused on the most down- or up-regulated genes in both thalamic mutants (Tables 1,2; Tables S2, S3). We first asked whether there are common genes deregulated in both *Sox2* and *Nr2f1* thalamic mutants. Fig. 1C shows that 514 genes significantly change their expression levels in both mutants, nearly all in the same direction (UP, or DOWN). This figure by far exceeds the number of genes expected by pure chance (more than four times the expected value for a random overlap, and probability of having a similar overlap by chance < 10⁻¹⁰⁰). Moreover, almost all the 50 genes in *Nr2f1* mutants with the most significant degree of downregulation are also significantly downregulated in *Sox2* mutants (44 out of 50); similarly, the majority of the 50 most downregulated genes in *Sox2* mutants are significantly downregulated also in *Nr2f1* mutants (30 out of 50) (Tables 1,2). Overall, these results clearly point to a SOX2 and NR2F1 co-regulated gene regulatory network in the early postnatal visual thalamus.

Genes regulated by SOX2 and NR2F1 are enriched in functions related to neuronal differentiation and connectivity

To get insights into the collective functions of dysregulated genes, we performed Gene Ontology (GO) and functional enrichment analyses, focusing our attention on the genes significantly varying their expression in both *Sox2* and *Nr2f1* mutants (Fig. 1D). Results showed a striking enrichment in functions related to neuronal differentiation, and in particular to neuronal connectivity, activity and synaptic plasticity (Fig. 1D). Among the differentially expressed genes (DEG) represented in these

categories, regulators in axon guidance (*Efna5*, *EphA5*, *EphA7*, *Sema7A*), specifically in retinal axon guidance (*Nrp1*, *Alcam*, *EphB1*), in glutamatergic synapses (*Grid1*, *Cdh8*), as well as transcription factors involved in visual development and axonogenesis (*Sox5*), could be identified.

Notably, among genes commonly downregulated in *Sox2* and *Nr2f1* mutants (Tables 1,2), *Vgf* deserved a particular attention, as it represented the most highly expressed, and one of the most strongly downregulated genes identified in both mutants (Tables 1,2). *Vgf* encodes a diffusible cytokine, transported along thalamo-cortical axons until the axon terminals. Its function is important for the development of cortical layer 4, onto which it acts instructively to maintain the appropriate numbers of layer 4 neurons, in particular within the somatosensory and visual cortical areas¹⁶. We thus looked at layer 4 development in thalamic *Sox2* mutants, using *in situ* hybridization for the layer 4 marker *RORβ*, the same marker previously used by¹⁶ (Fig. 2A). Indeed, we observed a reduction of the *RORβ* signal in the visual area (Fig. 2A), as observed in *Vgf*-knock-out mice¹⁶. This suggests that VGF may be a relevant downstream mediator of SOX2 thalamic function onto the development of cortical layer 4 in the visual area.

Sox5, encoding a transcription factor key to the development of cortical neuronal connectivity^{17,18}, was also strongly downregulated in both mutants (Tables 1,2). Immunofluorescence studies with an anti-SOX5 antibody showed that the numbers of cells expressing SOX5 protein at high levels typical of wild-type cells were strongly diminished in both mutants (Fig. 2B).

Deconvolution analysis identifies specific cell types affected by *Sox2* or *Nr2f1* deletion

A recent single-cell RNA-seq (scRNA-seq) analysis of the postnatal developing visual thalamus¹¹, defined specific cell types by their distinct transcriptional identity. Specific gene expression patterns characterizing excitatory neurons, inhibitory neurons, oligodendrocytes, astrocytes, endothelial cells, pericytes and microglia could be defined, in the normal (wild type) situation. We thus performed a deconvolution analysis of our RNA-seq data, based on these single-cell transcriptional profiles at P5, the earliest time point analyzed by¹¹(Fig. 3), in order to estimate the abundance of each cell type in our samples. This showed a relevant reduction of the estimated fraction of glutamatergic neurons, in both *Sox2* and (to a lesser extent) *Nr2f1* mutant dLGNs (Fig. 3A,B). Interneurons, pericytes and endothelial cells, conversely, were slightly increased, in both mutants. Oligodendrocytes were slightly reduced in *Sox2* mutants and increased in *Nr2f1* mutants. Of note, these changes were detected prior to overt abnormalities in the cell type composition of the mutant thalami, as previously shown by immunofluorescence and *in situ* hybridization only a week later, at P7 and/or P8⁹.

CUT&RUN identifies SOX2 binding sites in the P0 visual thalamus

We set out determine whether the functionally relevant deregulated genes are direct targets of SOX2 by establishing the genome-wide binding profile of this transcription factor. To overcome the anticipated technical difficulties due to the scant cell number obtained by dissection of this structure, we pooled dLGNs from P0 wild-type newborns and subjected them to CUT&RUN targeting SOX2¹² (Fig. 4A). Several thousand SOX2 peaks were identified in two independent experiments; their overlap further defined a group of 717 high-confidence, reproducible binding events (Fig. 4B). SOX2 peaks were found in promoter and intronic, as well as intergenic regions, consistent with the genomic binding pattern identified in other cellular systems¹³⁻¹⁵(Fig. 4C). Two types of global analyses supported the dataset: a) Motif analysis revealed SOX2 as the top enriched motif, along with other SOX factors, likely due to motif similarities (Fig. 4D), increasing our confidence of the specificity of the identified peak regions. b) Gene ontology of peak-associated genes (by GREAT,¹⁹) revealed enrichment for primarily neuronal biological processes (Fig. 4E), supporting involvement in the development of neurons and their connectivity.

We compared the genome-wide physical occupancy of SOX2 in the dLGN to our previously identified binding of this factor in brain-derived neural stem cells (NSC) performed both with CUT&RUN and with ChIP-seq^{14,15}. Among the 717 high-confidence targets, 248 were shared with NSC (Fig. 4F). This allowed us to define a subset of 469 dLGN specific SOX2 targets (Fig. 4F), which likely underlay a different role of SOX2 in this differentiated cell population. Of note, while the shared targets almost only displayed motif enrichment for SOX factors (Fig. 4G), the dLGN-specific SOX2 peak regions contained an enrichment for the ROR α consensus sequence, highly similar to the NR2F1 binding site – a motif that is not enriched in the neural stem cell subset (Fig. 4H). Examples of peaks shared with the neural stem cells and unique to the dLGN can be seen in Fig. 4I,L; note the presence of peaks in the proximity of the *Vgf* (Fig. 4I) and intronically within *Sox5* (Fig. 4L) regions, which are among the most strongly downregulated genes in *Sox2* mutant (see above and Table 1), suggesting their direct regulation by SOX2.

GREAT annotated the 717 binding events to a total of 1102 genes (Fig. 4J). This allowed us to overlap our CUT&RUN with the genes identified to be expressed or differentially expressed when *Sox2* is inactivated (RNA-sequencing data, Fig. 1C; Table S1). We found that 784 of the 1102 peak-associated genes were expressed in dLGNs at P0 (TPM > 5). Of these, 247 (145 down, 92 up) were dysregulated in SOX2 mutant dLGNs (< 0.01 FDR) (Fig. 4J), providing robust evidence of SOX2 direct regulation in differentiated neurons.

To understand whether these genes could also be regulated by NR2F1, we overlapped these differentially expressed SOX2 targets with the genes dysregulated in *Nr2f1* mutant dLGNs. This led to a core signature of 79 SOX2 and NR2F1 coregulated genes, which display known literature-based physical and functional interactions (Fig. 4K, STRING²⁰ diagram of core interacting SOX2/NR2F1 coregulated genes, disconnected nodes removed). Of note, this list includes several genes that are among the top 100 up- or down- regulated upon SOX2 deletion in the dLGN, including *Dmrtb1*, *Sgk1*, *Cnr2*, *Sox5*, *Gsg1l*, *Frrs1l*, *Sox13*, *Flt3*, *Neurog2*, *Hk3* (Fig. 4L).

Discussion

The conditional deletion of *Sox2*, and of *Nr2f1*, from the developing visual thalamus leads to important, significantly overlapping defects in the thalamus itself, as well as in the thalamus-connected visual cortex^{7,9}. We show that more than 1000 genes are dysregulated in each of the two mutants. Among them, an important subset of genes downregulated in both *Sox2* and *Nr2f1* mutants points to the existence of a shared network of genes essential for the proper development of the dLGN and its connections. In principle, the genes downregulated in the *Sox2*- and the *Nr2f1*-mutant dLGNs should explain three types of phenotypic defects observed in mutant mice^{7,9}: (i) the reduction in the number of neurons (glutamatergic) in the mutant visual thalami (at postnatal day 7); (ii) the alterations in terms of amount and distribution of thalamo-cortical and cortico-thalamic connectivity; (iii) the alterations in the retino-geniculate connections.

Indeed, gene ontology (GO) analysis of the genes dysregulated in both mutants reveals a highly significant enrichment of categories related to functions that develop abnormally in the mutants. These include, as the most significantly enriched categories, biological processes such as axonogenesis, neuron projection morphogenesis, chemical synapses transmission, axon development, retinal ganglion cell axon guidance (Fig. 2D), and related cellular components (Fig. 2D). It is likely that, collectively, genes forming these categories underlie, by their altered regulation, the above-mentioned defects.

Potential key contributors of SOX2- and NR2F1-dependent vision development

As shown in [Tables 1,2](#), a large proportion of genes highly expressed in WT cells, and highly down-regulated in *Sox2* mutants, are also down-regulated in *Nr2f1* mutants; the reciprocal is also true. In contrast, among the genes up-regulated in *Sox2* mutants, just a few are upregulated also in *Nr2f1* mutants ([Tables S2, S3](#)). This suggests that the functional association of SOX2 and NR2F1 is mostly involved in positive regulation (activation) of gene transcription.

Vgf

Among the genes at the top of the list, *Vgf* is the most highly expressed and among the most strongly downregulated ones, among those downregulated in both *Sox2* and *Nr2f1* mutants ([Table 1](#)). *Vgf* encodes a signaling molecule, transported along thalamo-cortical axons until the axon terminals in the L4 layer of the cortex, where it acts instructively to maintain the appropriate neuron numbers within the visual and somatosensory cortical areas¹⁶. The role of VGF in the layer was initially demonstrated by ablating the thalamo-cortical neuronal connections, thus depriving the cortex of sufficient VGF. Further, the knock-out of *Vgf* strongly reduced layer 4 in both the somatosensory and visual cortices. This phenotype could then be rescued to normality by transgenic expression of VGF in the developing cortex¹⁶. We observe ([Fig. 2A](#)) that *RORβ*, marking cortical layer 4, is moderately reduced in the visual cortex of *Sox2* mutants, mirroring, at least to an extent, what is found in *Vgf* mutants¹⁶. Knowing that VGF expression is strongly reduced in the mutant dLGN ([Table 1,2](#)), and that axonal connections reaching the visual cortex are also reduced in both mutants^{7,9}, and in agreement with the abnormalities in the primary (V1) and higher order (V^{H0}) visual cortical areas in both thalamic mutants^{7,9}, we hypothesize that the observed cortical layer defects of the *Sox2* mutant ([Fig. 2A](#)) might be the consequence of a reduced amount of VGF reaching the axon terminals in the visual cortex. Thus, VGF might be a relevant mediator of the function of SOX2 during area- and layer-specific cortical development. Furthermore, *Sox2* is also down-regulated in the somatosensory VP (Vento-Posterior) thalamic nucleus in *Sox2* thalamo-specific conditional mutants, leading to moderate abnormalities of the somatosensory cortical area⁹. We thus looked at the somatosensory cortex layer 4 and observed a *RORβ* reduction (less marked than in the visual area) also in this area ([Fig. 2A](#)), in agreement with the above hypothesis.

Sox5

SOX5, a transcription factor, is significantly downregulated in both mutants at the mRNA and protein level ([Table 1,2](#); [Fig. 2B](#)). In humans, the *SOX5* gene is associated to “Optic nerve hypoplasia bilateral, autosomal dominant” (OMIM #165550). As shown in *Sox5*-null mice^{18 17} this gene controls important aspects of neuronal connectivity, such as the development of cortico-fugal neurons, including cortical neurons projecting to thalamic neurons. Misrouting of subplate and layer 6 cortico-thalamic axons to the hypothalamus is also observed¹⁷. These data suggest the possibility that the reduction of SOX5 in the dLGN of *Sox2* and *Nr2f1* mutants affects axonogenesis.

Hs6st2, Hs6st3

The Heparan sulphate 6-O-sulfotransferase enzyme modifies the sulfation status of heparan sulphate proteoglycans (HSPG)²¹, extracellular matrix proteins, with covalently linked polysaccharidic chains, which are polymerized by specific enzymes and further modified by sulfation to obtain ample structural and functional diversity²². HSPG interact with a variety of proteins, implicated in cell proliferation and differentiation, adhesion, migration, and other processes²¹. It is intriguing that three of the genes (*Hs6st1, 2* and *3*) encoding isoforms of the enzyme are downregulated in *Sox2* and/or *Nr2f1* mutant dLGNs. Axon guidance or axon extension defects abnormalities, and eye defects, have been reported in association with mutation of *Hs6st* genes²¹⁻²⁵. We hypothesize that SOX2- and NR2F1-dependent expression of HS6ST enzymes,

acting on the sulfation status of HSPGs, might play a role in the development of axonal connectivity of thalamic neurons.

Rora

Rora, also named *Nr1f1*, is one of the most downregulated genes in *Sox2* mutants, although not in *Nr2f1* mutants (Tables 1,2); yet, it may have a special significance also in the perspective of the present study, focusing on a SOX2 and NR2F1 co-regulated gene expression network. *Rora* encodes a close homolog of NR2F1, and its product NR1F1 binds to a DNA sequence closely related to that of NR2F1, with an identical core binding site (<https://www.genecards.org/cgi-bin/carddisp.pl?gene=RORA>; <https://www.genecards.org/cgi-bin/carddisp.pl?gene=NR2F1>). *Rora* spontaneous (staggered mouse) or engineered mutations cause ataxia and cerebellar neurodegeneration, with synaptic arrangement and immature morphology of cerebellar neurons (Purkinje) ^{26,27}. *RORα* defects are connected in humans with “Intellectual development disorder with or without epilepsy, or cerebellar ataxia, OMIM #600825. If the activity of RORα on some genes partially overlaps with that of NR2F1, it is possible that, in the *Sox2* mutant, the reduced activity of RORα leads to down-regulation of a set of genes closely related to those activated by NR2F1, explaining some common phenotypic defects observed in *Sox2* and *Nr2f1* mutants.

Mechanisms of dysregulation of gene expression in mutant thalamus

The observed dysregulation of several genes in the mutant thalami could be due to variations in the abundance of specific cell types, as well as to variations in gene expression levels within specific cell types normally expressing SOX2 and NR2F1. They may also be due to indirect effects, whereby decreased expression of genes downregulated following *Sox2* or *Nr2f1* knockout affects the expression of other genes not directly targeted by SOX2 or NR2F1 themselves. Evidence for direct effects of *Sox2* or *Nr2f1* deficiency is provided by CUT&RUN, which identified several hundreds of chromatin sites bound by SOX2 in replicate experiments (Fig. 4).

The variation in cell numbers in the mutants is indeed important. In fact, in previous work, we detected a reduction in the number of neurons by P7 following *Sox2* thalamic deletion ⁹. In the present work, we detect a specific reduction in the fraction of cells bearing a transcriptional program characteristic of projection neurons, already at P0 (Fig. 3, deconvolution). This reduction is likely to be rooted in gene expression abnormalities within glutamatergic neurons themselves, which normally express SOX2 and NR2F1, and not in a decreased number of cells. Accordingly, several of our downregulated genes (e.g. *Vgf*, *Sox5*, *Rora*) are reported to be prevalently expressed in neurons in published datasets of scRNA-seq analyses ²⁸. On the other hand, the large relative increase in GABAergic interneurons (that do not express SOX2 in the thalamus) ⁹ might be due to indirect effects, such as a change in the fate of part of the mutant neurons, or an increased interneuron colonization of the mutant thalamus, or increased survival of the interneurons in the mutant thalamus.

dLGN-specific SOX2 binding sites differ from those shared with neural stem cells

The comparison of SOX2 binding sites detected by CUT&RUN in the visual thalamus (present paper) with those previously observed in neural stem cells by both CUT&RUN and ChIPseq ¹³⁻¹⁵ allowed us to compare the most represented SOX2 binding sequence motifs specific to the thalamus with those specific to NSC. SOX2 binding sites are among the most represented sites in both datasets, as expected (Fig. 4D, 4G, 4H). Remarkably, however, the thalamus-specific SOX2-binding sites showed, among the most represented, the binding site for RORα, also called NR1F1. As discussed above, NR1F1 and NR2F1, highly homologous, recognize similar DNA sequences with a core reported to be identical (<https://www.genecards.org/cgi-bin/carddisp.pl?gene=RORA>; <https://www.genecards.org/cgi-bin/carddisp.pl?gene=NR2F1>). This suggests that the molecular

basis for the sharing of co-regulated target genes between SOX2 and NR2F1 might lie in the co-binding of the two factors to regulatory DNA sequences.

Overall, our work identifies a common transcriptional program driven by SOX2 and NR2F1 in the visual thalamus, highlighting a small subset of genes commonly downregulated in both mutants, and potentially important for explaining abnormalities of retina-thalamus-cortex connections. This paves the way to the precise identification, by functional transgenic studies, of genes mediating the common Sox2 and Nr2f1 functions in the visual thalamus. It will also open the way to defining the molecular mechanisms mediating SOX2 and NR2F1 interactions in transcriptional regulation.

Table 1
Genes most down-regulated in Sox2-mutant visual thalamus

Gene	Av. Exp. Sox2 WT TPM	Av. Exp. Sox2 MUT TPM	Av. Exp. Sox2 WT/ Av. Exp. Sox2 MUT	log fold change	DEG in Nr2f1 MUT
9030625G05Rik	8.1	1.05	7.69	-2.895780452	YES
Hdc	21.82	2.97	7.35	-2.813501459	
Fam83f	9.84	1.6	6.15	-2.550812898	YES
Ptpn22	4.52	0.84	5.4	-2.371974499	YES
Tuba8	9.12	1.75	5.2	-2.313158562	
Hs6st2	109.4	22.54	4.85	-2.230994808	
Pcdhac1*	14.92	3.04	4.9	-2.224814845	YES
Drd5	10.05	2.23	4.51	-2.107744688	YES
Flt3	11.74	2.72	4.31	-2.044844104	YES
Wnt9b	50.83	12.19	4.17	-1.997868077	YES
Hs3st1	121.89	29.48	4.13	-1.986194868	
Rora	62.54	16.24	3.85	-1.955983475	
Nexn	7.58	1.92	3.95	-1.927547838	YES
Neurog2	12.08	3.05	3.96	-1.925343754	YES
Hs6st3*	48.32	12.18	3.97	-1.92121348	YES
Vgf	882.88	231.32	3.82	-1.873781013	YES
Muc15	6.6	1.72	3.84	-1.870164907	YES
Sp9*	269.34	78.82	3.42	-1.710807965	
Lgi2	35.57	10.41	3.42	-1.707749283	
Slc18a2	218.62	64.43	3.39	-1.701630392	
Col9a1	9.37	3	3.13	-1.698584742	
Cpne9	43.55	13.27	3.28	-1.655753749	YES
Efna5	37.94	11.68	3.25	-1.638069788	
Cd47	129.27	41.23	3.14	-1.586910507	YES
Tmem132d	20.26	6.61	3.07	-1.551311996	YES
Galnt14	34.04	11.15	3.05	-1.548701903	
Drd1	4.73	1.55	3.04	-1.543540066	
2900005J15Rik	6.47	2.13	3.04	-1.539759249	
Chrm3	16.91	5.67	2.98	-1.512314896	YES
Adgrl2	88.59	29.81	2.97	-1.510982218	YES
Tmem132b*	70.73	23.74	2.98	-1.510191854	YES
Edaradd	9.05	3.05	2.97	-1.507508032	
Frrs1l	71.66	24.07	2.98	-1.506989596	YES
Zmat4*	25.97	8.78	2.96	-1.50054383	YES
Kcnn1	45.79	15.68	2.92	-1.484704706	
Tmem163	142.8	48.96	2.92	-1.484221758	YES
Ephx4	9.52	3.3	2.88	-1.469356885	YES
Trhde	12.26	4.25	2.89	-1.461311822	

Sox13	86.24	30.48	2.83	-1.440847204	YES
Cdca7	51.7	18.3	2.83	-1.43834782	
Dmrtb1*	111.05	39.94	2.78	-1.416756694	YES
Cpne4	34.3	12.39	2.77	-1.406373382	YES
Slc6a4	58.61	21.26	2.76	-1.402243263	YES
Fzd8	17.56	6.36	2.76	-1.399855218	
Sox5	25.07	9.1	2.75	-1.397136566	YES
Dkk4	6.33	2.33	2.71	-1.374842774	YES
Il22	7.01	2.6	2.69	-1.372466032	
Camk4	88.09	32.64	2.7	-1.367842536	YES
Gm15417	17.04	6.48	2.63	-1.346050718	YES
Dusp5	8.72	3.31	2.63	-1.335628473	

Table 2
Genes most down-regulated in *Nr2f1*-mutant visual thalamus

Gene	Av. Exp. <i>Nr2f1</i> WT TPM	Av. Exp. <i>Nr2f1</i> MUT TPM	Av. Exp. <i>Nr2f1</i> WT / Av. Exp. <i>Nr2f1</i> MUT	log fold change	DEG in <i>Sox2</i> MUT
Cck	592.67	105.75	5.6	-2.494874416	YES
Nexn	5	0.95	5.28	-2.417211004	YES
Tmem132d	27.05	5.84	4.64	-2.222822518	YES
Slc6a4	41.84	11.42	3.66	-1.880306789	YES
Fam163a	30.74	8.83	3.48	-1.807413828	YES
Cpne4	44.25	13.24	3.34	-1.754921497	YES
Dlk1	38.85	12.04	3.23	-1.731410795	YES
Rasd1	9.85	3.13	3.15	-1.652614906	YES
Stxbp6	35.08	11.5	3.05	-1.621502112	YES
Vgf	855.03	282.91	3.02	-1.60661903	YES
Zmat4*	24.53	8.21	2.99	-1.591740499	YES
Wnt9b	57.22	19.28	2.97	-1.581768251	YES
Dkk4	7.87	2.74	2.88	-1.519344111	YES
9030625G05Rik	4.6	1.62	2.83	-1.508843997	YES
Pex5l	8.52	3.07	2.77	-1.496385596	YES
Epha7	34.01	11.79	2.89	-1.436854461	YES
Asic2	85.02	32.42	2.62	-1.426127408	YES
Hr	14.71	5.53	2.66	-1.419102634	YES
Ak5	48.21	19.44	2.48	-1.32240026	YES
Ccer2*	5.99	2.51	2.39	-1.26669584	YES
Itga9	8.92	3.89	2.3	-1.209569337	YES
Hs6st3*	64.49	28.25	2.28	-1.207364056	YES
Fam189a1*	95.17	43.11	2.21	-1.159193496	YES
Smyd1	46.41	21.22	2.19	-1.143516085	YES
Epha4	85.04	39.26	2.17	-1.127404784	YES

Abcc8	13.12	6.14	2.14	-1.106531723	YES
Cdh8	101.92	48.02	2.12	-1.098249211	YES
Nefm	353.74	167.17	2.12	-1.090279287	YES
Chrm3	22.76	10.83	2.1	-1.089514412	YES
Pcdhac1*	17.48	8.31	2.1	-1.089091981	YES
lsm1	4.45	2.16	2.06	-1.064831627	
Muc15	7.49	3.64	2.06	-1.052316375	YES
Trpc3	27.24	13.31	2.05	-1.045132272	YES
Ube2q1	235.63	115.61	2.04	-1.041974849	YES
Chrna7	9.75	4.81	2.02	-1.0302295	
Mag	13.17	6.51	2.02	-1.029770911	YES
Cntnap5a	13.99	6.96	2.01	-1.016787279	YES
Tmem132b*	94.79	47.41	2	-1.015256615	YES
Samd5*	23.94	12.02	1.99	-1.010167121	YES
Gal	25.19	12.54	2.01	-1.008787519	
Hs3st5	16.51	8.29	1.99	-1.008469418	YES
Dcdc2a	7	3.54	1.98	-0.986267357	YES
Mef2c	31.56	16.19	1.95	-0.978036224	
Sv2b	9.49	4.99	1.9	-0.945580396	
BC089491	9.83	5.16	1.9	-0.942591105	
Cdh7	12.1	6.39	1.9	-0.93529364	YES
Dmrtb1*	113.78	60.29	1.89	-0.932789973	YES
Vangl1	34.34	18.29	1.88	-0.925797074	YES
Neurog2	13.95	7.52	1.86	-0.90664983	YES
St3gal1	20.68	11.17	1.85	-0.902451943	YES

Av. Exp.: Average expression. TPM: transcripts per million. DEG: differentially expressed gene. *: T-dark genes (genes about whose function very little is known; see <https://pharos.nih.gov/>, <https://pharos.nih.gov/targets>). YES in column 5 marks those genes that are also significantly dysregulated in the “other” mutant, i.e. *Nr2f1* mutant for tables listing genes dysregulated in *Sox2* mutants, and vice versa; YES genes vary, in the “other” mutant, in the same direction (downregulated in Table 1,3; upregulated in Table 2,4). **YES** (in bold) means that the gene is among the 50 most down- or up-regulated genes also in the “other” mutant.

Figure legends

Figure 1

RNA-seq of the visual thalamus (dLGN) of *Sox2* and *Nr2f1* thalamic mutants identifies many genes differentially expressed in both mutants.

A. dLGN immunofluorescence with antibodies recognizing SOX2 (green) and NR2F1 (red) at P0 (top; wild type) and P8 (bottom; wild type control, Con, and thalamic mutant, cKO). Note that *Sox2* thalamic loss (*Sox2* cKO in the P8 panel) does not detectably affect NR2F1 expression. Scale bar top: 200um; scale bar bottom: 50um. **B.** Sections comprising the dLGN at P0 used for RNA-seq analyses, before (left) and after (right) dissection. Scale bar: 200um. **C.** Numbers of significantly dysregulated genes, identified in RNAseq experiments, in *Sox2* and *Nr2f1* mutants, and in both mutants are shown. **D.** Gene Ontology (GO) analysis of genes differentially expressed in both *Sox2* and *Nr2f1* mutants (DEG; the 514 genes in B) reveals enrichment in categories involved in neuronal development. The GO Biological Processes, Cellular Component and Molecular Function categories which are significantly enriched within the indicated mutants are shown.

Figure 2

***Sox2* thalamic cKO affects the development of cortical layer 4, and thalamic SOX5 protein expression.**

A. *In situ* hybridization with a *Rorb* probe, marking layer 4, on sagittal brain sections of thalamic *Sox2* mutants (*Sox2* cKO) versus controls carrying wild type *Sox2* (Con) at P8 shows a reduction in the region corresponding to the visual (V, V1) and somatosensory (S1) areas (enlarged details). The results shown are representative of n=3 mutant and n=3 control brains analysed. Scale bar: 600um. **B.** Immunofluorescence with antibodies recognizing SOX5 (red) on coronal sections of E18.5 brains of controls, *Sox2* cKO and *Nr2f1* cKO. DAPI (blue) marks nuclei. At this stage, the dLGN is only modestly reduced in the mutants^{7,9}. Arrows point to SOX5-positive cells within the dLGN region. Note the reduction of the SOX5-positive area in the *Sox2* and *Nr2f1* mutant dLGN. SOX5 positivity is not altered in the adjacent hippocampal region in mutants. The results shown are representative of n=3 mutant and n=3 control brains analysed. Scale bar: 200um.

Figure 3

Deconvolution analysis documents the loss of projection neurons transcriptional identities in *Sox2* and *Nr2f1* mutants.

The cell types indicated are those defined by their scRNA-seq transcriptional identity in ¹¹. The data for all three mutant and control samples are shown numerically in (A) and graphically in (B,C).

Figure 4

CUT&RUN of SOX2 binding in the P0 visual thalamus.

A. Schematic depiction of the CUT&RUN experimental design. 2 independent biological replicates for SOX2 and an anti-HA negative control were performed from pools of 6 dorsal lateral geniculate nuclei (dLGN) from 3 brains.

B. Venn diagram overlap and signal intensity plots of the two SOX2 replicates, showing enrichment over the control in all peak regions.

C. Peak region annotation by HOMER, showing that SOX2 binds primarily promoter, intronic and intergenic regions in dLGN.

D. HOMER known motifs for dLGN SOX2 peaks. The SOX2 motif is the most highly ranked, followed by other SOX factors.

E. Gene ontology enrichment of biological processes for genes associated by GREAT to SOX2 dLGN peaks. Dot size shows number of peak associated genes, dot color represents -log₁₀ FDR (FDR < 0.05), and the x axis represents fold enrichment. The top 20 terms are shown.

Enrichment terms include neuron development related processes. **F.** Pie chart depicting the distribution of SOX2 dLGN peaks in unique peaks and those shared with neural stem cell (NS) SOX2 datasets.

G. HOMER known motifs for dLGN and NS shared peaks. NFY and SOX2 are the top motifs. **H.** HOMER known motifs for dLGN unique peaks. Top motifs include ZIC3 and ROR α . Sox2 is ranked 10th.

I. CUT&RUN tracks as visualized in Integrative Genome Viewer

(IGV), showing both dLGN and NS shared peaks (left) and dLGN unique peaks (right). **J.** Schematic depiction of CUT&RUN and RNA-seq overlap, showing genes that are transcribed (> 5 TPM) and those that are differentially expressed (DEG) in *Sox2* mutant dLGN (FDR < 0.01). **K.** STRING map of interactions between SOX2, NR2F1, and the 79 SOX2 dLGN direct targets that are dysregulated in both *Sox2* and *Nr2f1* mutant mice. Confidence was set on default (0.4), disconnected nodes were removed, and line thickness represents confidence. Red nodes are those known to be involved in neurogenesis, while blue nodes are not included in the neurogenesis set and could represent novel genes important to the generation of neurons in the visual thalamus. **L.** CUT&RUN tracks showing SOX2 dLGN peaks near the important targets *Vgf* (left), *Sox5* (center, intronic region), and *Sox13* (right). While the peaks near *Vgf* and *Sox5* are unique to the dLGN datasets, the *Sox13* peak is also bound by SOX2 in NS.

Materials and Methods

Mouse strains

Sox2 Mutant mice were obtained by crossing Sox2Flox (Favaro et al., 2009) with Rora-Cre (Chou et al., 2013) mouse lines.

Nr2f1 mutant mice were generated by crossing COUP-TF1Flox (Armentano et al., 2007) with Rora-Cre (Chou et al., 2013) mouse lines (note: COUP-TF1 is the old name of Nr2f1).

Genotyping was performed with the following primers (Chou et al., 2013; Mercurio et al. 2019):

Rora-Cre IRES Forward: 5'AGGAATGCAAGGTCTGTTGAAT 3'; Rora-Cre IRES Reverse: 5' TTTTCAAAGGAAAACCACGTC 3';

Sox2 Flox Forward: 5'AAGGTACTGGGAAGGGACATT 3'; Sox2 Flox Reverse: 5'AGGCTGAGTCGGGTCAATTA 3';

COUP-TF1 Flox Forward 5'-CTGCTGTAGGAATCCTGTCTC-3'; COUP-TF1 Flox Reverse: 5'-AATCCTCCTCGGTGAGAGTGG-3' and 5'- AAGCAATTTGGCTTCCCCTGG-3'.

The day of vaginal plug was defined as embryonic day 0 (E0) and the day of birth as postnatal day 0 (P0).

We note that Rora is the gene to the 3' of which the Cre gene has been inserted (as IRES-Cre knocked-in into the 3'UTR of the gene), in the Cre transgene driving Sox2 and Nr2f1 deletion in our thalamic mutants. In principle, *a priori*, the Cre insertion into the Rora locus per se might have lowered Nr1f1-Rora expression. However, we also note that the significant (about 4 times) reduction of expression is observed in Sox2 mutants, but not Nr2f1 mutants, where expression levels are unchanged with respect to control. We thus think it is unlikely that Cre insertion plays a role in the expression reduction of Rora in Sox2 mutants.

Brain extraction and dissection for RNA-sequencing

Brains from Sox2 thalamic mutants and controls, and Nr2f1 thalamic mutants and controls, at postnatal day 0 (P0) were removed from the skulls in ice-cold PBS1x, coronally embedded in low melt agarose 4% in PBS1x and kept on ice. Brains were then sectioned in ice-cold and sterile PBS1x with a vibratome (Leica VT1000s). 200µm brain sections were collected in ice cold PBS1x. Sections including the dLGN were identified under a stereoscopic microscope and the dLGN was quickly dissected with sterile chirurgic scalpels on a glass slide. dLGNs were present usually in two sequential sections. Excised dLGNs were collected with a pipette, collected in sterile tubes, snap-frozen in liquid nitrogen and stored at -80°C until RNA extraction. All the sections were imaged before and after dissection of dLGNs.

RNA extraction and RNA sequencing

RNA from snap-frozen dLGNs was extracted with the RNeasy Micro Kit (QIAGEN) with some precautions. Each dissected dLGN weighed about 0,3 mg. dLGNs samples were taken out of the -80°C and left few seconds on the bench before tissue homogenization. Each dLGN sample was homogenized in 350 µl of RLT buffer containing β-mercaptoethanol (β-ME) as indicated by the manufacturer (10 µl of β-ME per 1 ml of RLT buffer). Tissue was homogenized by sequential passages through needles of descending diameter. Precisely we used a 1cc syringe and three different sterile needles: 18G, 22G and 26G. First, the tissue was passed 10 times through the larger diameter needle (18G), then 10 times through the intermediate diameter needle (22G) and finally 10-12 times through the thinnest one (26G), in order to correctly lysate and disrupt the tissue. Subsequently, the homogenized sample was vortexed for 30 seconds and centrifuged for 3 minutes at 13200 rpm to pellet the debris. RNA was extracted from the homogenized supernatant as indicated by the manufacturer's instructions.

RNA sequencing was performed on three independent samples for both mutant and control dLGN. Each sample was composed of dLGNs from three animals of the same genotype pooled together. Genotypes were: for mutants, Sox2flox/flox, or Nr2f1 flox/flox, plus RORalpha-Cre transgene; for

controls, we used littermates of the respective mutants (Sox2 or Nr2f1 mutants), carrying two intact copies of Sox2 (Sox2^{flox/flox} or Sox2^{flox/+}), or Nr2f1 (Nr2f1^{flox/flox} or Nr2f1^{flox/+}), and no Cre transgene. For each sample sequenced we obtained at least 150 ng of high quality total RNA (RIN \geq 8), and thus 150 ng were used for library preparation. Library preparation was performed with Nugen Universal + mRNAseq kit, followed by sequencing in a HiSeq 4000 (Illumina), paired-reads 2x 75 bp. The KEGG 2019 Mouse database was used to analyse RNA-sequencing data. The Gene Ontology analysis reported in Fig. 1 was done using Enrichr (<https://maayanlab.cloud/Enrichr/0>).

RNA-seq raw data analysis and deconvolution analysis

Sequence reads were mapped with STAR²⁹ against the mouse RefSeq transcriptome, version April 2019, retrieved from the UCSC Genome Browser Database³⁰. Read counts and subsequent normalized transcripts per million (TPM) were computed with RSEM (rsem-calculate-expression)³¹.

Differential expression analysis was performed with edgeR³². Initial read counts were normalized by trimmed mean of M values (TMM), with default parameters. Differentially expressed genes were identified by the quasi-likelihood (QL) F-test of edgeR (glmQLFfit and glmQLFtest functions, with default parameters). We selected as differentially expressed all genes with an adjusted p-value (FDR) < 0.01 .

Bulk RNA-Seq deconvolution analysis was performed with MuSiC³³, with as input raw count tables for bulk RNA-Seq and raw count tables at P5 for single cell RNA-Seq samples. Single cell clusters and cell type annotations were retrieved from¹¹.

For Tables 1-4, DEG were initially selected having expression levels higher than 4 TPM (at least in wild type, for downregulated genes; at least in mutant, for upregulated genes). Average expression values were then calculated from triplicate samples in Table S1. The top 50 most down- or up-regulated genes (with greatest log-fold change in mutant versus wild type) are shown for both mutants.

Immunohistochemistry

To detect SOX2 and NR2F1, slides were washed 2X for 10 minutes in PBS1X and then unmasked in Na Citrate 0.1M-Citric acid 0.1M pH6 solution for 2 minutes. Sections were washed in PBS1X for 10 minutes and incubated 1 h with pre-blocking solution (Sheep serum 5%, Tween-20 0,3% in PBS1X) at room temperature. Sections were then incubated over night at 4°C in blocking solution (Sheep serum 1%, Tween-20 0,1% in PBS1X) containing primary antibodies. The following primary antibodies were used: anti-SOX2 diluted 1:500 (R&D AB2018, mouse) and anti-NR2F1 diluted 1:1000 (Abcam Ab181137, rabbit). Sections were then washed 2X for 10 minutes in PBS1X and incubated for 1h at room temperature with blocking solution containing the following fluorescent secondary antibodies: anti-mouse IgG Alexa Fluor 488 and anti-rabbit IgG Alexa Fluor 594 (Thermo Fisher) diluted 1:500. Slides were then washed 2X for 10 minutes in PBS1X and mounted.

To detect SOX5, slides were treated as above, but without the unmasking step. The anti-SOX5³⁴ antibody (a gift from A. Morales) was diluted 1:500, and the anti-rabbit Alexa Fluor 594 (Thermo Fisher) was diluted 1:1000.

In situ hybridization

In situ hybridization was performed as in (Mercurio et al. 2019). Briefly, brains at P7 brains were dissected and fixed overnight in paraformaldehyde (PFA) 4% in PBS (Posphate Buffered Saline) 1X. The fixed tissue was cryoprotected in a series of sucrose solutions (15%, 30%) in PBS 1X and then embedded in OCT (Killik, Bio-Optica) and stored at -80°C. Brains were sectioned (20 μ m) with a cryostat, placed on a slide (Super Frost Plus 09-OPLUS, Menzel) and stored at -80°C. Slides

were then defrosted, fixed in formaldehyde 4% in PBS for 10 minutes (min), washed 3 times for 5 min in PBS 1X, incubated for 10 min in acetylation solution (for 200 ml: 2.66 ml triethanolamine, 0.32 ml HCl 37%, 0.5 ml acetic anhydride 98%) with constant stirring and then washed 3 times for 5 min in PBS1X. Slides were placed in a humid chamber and covered with prehybridization solution (50% formamide, 5X SSC, 0.25 mg/ml tRNA, 5X Denhardt's, 0.5 µg/ml salmon sperm) for at least 2 hours (h) and then incubated in hybridization solution (fresh prehybridization solution containing the digoxigenin (DIG)-labelled RNA probe of interest) overnight at 65°C. Slides were washed 5 min in 5X SSC, incubated 2 times in 0.2X SSC for 30 min at 65°C, washed 5 min in 0.2X SSC at room temperature and then 5 min in Maleic Acid Buffer (MAB, 100 mM maleic acid, 150 mM NaCl pH 7.5). The slides were incubated in blocking solution (10% sheep serum, 2% blocking reagent (Roche), 0.3% Tween-20 in MAB) for at least 1 h at room temperature, then covered with fresh blocking solution containing anti-DIG antibody Roche © 1:2000 and finally placed overnight at 4°C. Slides were washed in MAB 3 times for 5 min, in NTMT solution (100 mM NaCl, 100 mM Tris-HCl pH 9.5, 50 mM MgCl₂, 0.1% Tween-20) 2 times for 10 min and then placed in a humid chamber, covered with BM Purple (Roche), incubated at 37°C until desired staining was obtained (1-6 h), washed in water for 5 min, air dried and mounted with Eukitt (Sigma). A DIG-labelled *RORβ* probe was used³⁵.

CUT&RUN

CUT&RUN was performed as described in ³⁶. dLGNs were dissected (as described above in Brain extraction and dissection for RNA-sequencing), suspended in 1.5 ml of cold CUT&RUN wash buffer (HEPES pH 7.5 [20 mM], NaCl [150 mM], Spermidine [0.5 mM], Roche Complete Protease Inhibitor EDTA-Free (Cat. # COEDTAFRO, Roche) and then manually dissociated by gentle pipetting. Cells were collected by centrifugation at 600g for 3 minutes and the supernatant carefully removed, and then washed 2 more times by resuspension in 1 ml wash buffer. 4 µl of ConA bead slurry (prepared in binding buffer as described), were added per pair of dLGNs and the cells and beads were incubated for 10 min on a rotator. While the rest of the dissections proceeded, bead bound cells were kept in storage buffer (wash buffer with EDTA [2 mM]). 3 brains (6 dLGNs) were collected per sample. Once all dLGNs were collected, the beads were collected on the magnet and resuspended in 150 µl antibody buffer (wash buffer with EDTA [2 mM] and 0.025 % digitonin). Antibodies were added at 1:100 dilution and incubation proceeded ON at 4 °C. Antibodies used included anti-SOX2 (ABIN2855074, antibodies online), anti-SOX2 (ABIN2855073, antibodies online), and anti-HA (05-902R, Merck). One SOX2 sample was with only ABIN2855074, the other was with both SOX2 antibodies together. The two SOX2 biological replicates were performed independently from different litters of mice. The next day samples were washed twice with dig-wash (wash buffer with 0.025% digitonin) and then resuspended in 150 µl dig-wash containing pA-MNase (New England Biolabs 700 ng/ml, received as a gift from Steven Henikoff) at 1:200 dilution and rotated for 1 hr. Samples were washed twice and then resuspended in 100 µl dig-wash and equilibrated in ice. 2 µl 100 mM CaCl₂ was added, and digestion proceeded for 30 min in ice. 100 µl 2X STOP buffer (NaCl [340 mM], EDTA [20 mM], EGTA [4 mM], digitonin [0.05%], RNase A [100 µg/ml], glycogen [50 µl/ml]) was added to stop the digestion reaction and incubated at 37 °C for 30 min. Samples were pelleted at 16000 g for 5 min and placed on the magnet rack. The supernatant was harvested, and beads were discarded. DNA purification was performed with phenol chloroform. Library preparation was performed using the KAPA HyperPrep Kit (Cat. #KK8504, KAPA Biosystems) according to manufacturer's instructions, using KAPA DUI adapters at 0.15 µM. Libraries were sequenced with the Illumina NextSeq 550 using the High-Output 75 cycles kit v2.5 (Cat. #20024906, Illumina), 36 base pair pair-end.

CUT&RUN Data Analysis

Reads were trimmed to remove adapters, artifacts and repeats of poly [AT]₃₆, [C]₃₆ and [G]₃₆ with bbmap bbdduk ³⁷ (version 38.18). Alignment was performed to the mm10 mouse genome with bowtie

³⁸ (version 1.0.0) with the options `-v 1 -m 1 -X 500`. Samtools ³⁹ (version 1.11) was used for deduplication and to remove and incorrectly paired reads. Bedtools ⁴⁰ (version 2.30.0) was used to remove reads mapped to the CUT&RUN mm10 Suspect List ⁴¹ from bam files. Peaks were called for each replicate using MACS2 ⁴² with the options `-f BAMPE --keep-dup all -p 1e-2 -SPMR -bdg` against the anti-HA negative control. Output narrowPeak files were overlapped using Bedtools intersect, keeping only reproducible peak regions called in both biological replicates. Signal intensity plots were created using ngsplot (<https://doi.org/10.1186/1471-2164-15-284>, version 2.63) with options `-N 4 -GO none -SC global`, plotting signal intensity of the replicates compared to the anti-HA control. Motif analysis was done using HOMER ³⁹ (version 4.11) findMotifsGenome to find motifs in the mm10 genome using `-size given`, and peak region annotation was done with HOMER annotatePeaks on default settings. Peak set gene annotation was done using GREAT ¹⁹ (version 4.0.4) with default parameters, and gene names were used to compare CUT&RUN results with the RNA-seq analysis. Gene ontology was performed for GO biological processes using ShinyGO ⁴³ (<https://doi.org/10.1093/bioinformatics/btz931>), results were cutoff at FDR of 0.05 and the top 20 results were graphed. For neural stem cell CUT&RUN data, datasets were downloaded from ⁴⁴ and processed as described above for the dLGN datasets. The two peak sets were overlapped with Bedtools intersect. STRING ⁴⁵ was used to graph protein-protein interactions between the set of 79 genes that were found to be dLGN SOX2 targets coregulated by SOX2 and NR2F1, disconnected nodes were removed. Bedgraphs were visualized in IGV ⁴⁶.

Accession to genomic data

RNA-seq data are accessible through Gene Expression Omnibus (GEO) Accession Number GSE233131. CUT&RUN data are accessible through ArrayExpress Accession Number E-MTAB-13004.

Acknowledgements

We thank Dr. D. O'Leary for providing the ROR α -Cre mouse, Dr. Carolina Frassoni and her laboratory for contributing to the immunofluorescence experiments in Fig. 1A, and Dr. Aixa Morales for supplying the anti-SOX5 antibody. Work in the Nicolis laboratory is supported by EU ERANET-NEURON grant Brain4Sight and by Fondazione Telethon – Fondazione Cariplo Alliance GJC21176 grant to S.K.N. Work in the Cantù lab is supported by Cancerfonden (CAN 2018/542 and 21 1572 Pj), the Swedish Research Council, Vetenskapsrådet (2021–03075), Linköping University support and LiU-Cancer. C.C. is a Wallenberg Molecular Medicine (WCMM) fellow and receives generous financial support from the Knut and Alice Wallenberg Foundation. The computations and data handling for CUT&RUN analyses were enabled by resources provided by the National Supercomputer Centre (NSC), funded by Linköping University. Peter Mürger at the National Supercomputer Centre is acknowledged for assistance concerning technical and implementational aspects in making the codes run on the Sigma resource.

References

- 1 Williamson, K. A. & FitzPatrick, D. R. The genetic architecture of microphthalmia, anophthalmia and coloboma. *European journal of medical genetics* **57**, 369-380, doi:10.1016/j.ejmg.2014.05.002 (2014).
- 2 Cardozo, M. J., Sanchez-Bustamante, E. & Bovolenta, P. Optic cup morphogenesis across species and related inborn human eye defects. *Development* **150**, doi:10.1242/dev.200399 (2023).
- 3 Graw, J. The genetic and molecular basis of congenital eye defects. *Nat Rev Genet* **4**, 876-888, doi:10.1038/nrg1202 (2003).
- 4 Schaaf, C., Yu-Wai-Man, P. & Valentin, I. in *GeneReviews*((R)) (eds M. P. Adam *et al.*) (1993).
- 5 Williamson, K. A., Yates, T. M. & FitzPatrick, D. R. in *GeneReviews*((R)) (eds M. P. Adam *et al.*) (1993).
- 6 Armentano, M. *et al.* COUP-TFI regulates the balance of cortical patterning between frontal/motor and sensory areas. *Nat Neurosci* **10**, 1277-1286, doi:10.1038/nn1958 (2007).
- 7 Chou, S. J. *et al.* Genuculocortical input drives genetic distinctions between primary and higher-order visual areas. *Science* **340**, 1239-1242, doi:10.1126/science.1232806 (2013).
- 8 Mercurio, S., Serra, L., Nicolis, S.K. More than just Stem Cells: Functional Roles of the Transcription Factor Sox2 in Differentiated Glia and Neurons. *Int J Mol Sci* **20**, doi:10.3390/ijms20184540 (2019).
- 9 Mercurio, S. *et al.* Sox2 Acts in Thalamic Neurons to Control the Development of Retina-Thalamus-Cortex Connectivity. *iScience* **15**, 257-273, doi:10.1016/j.isci.2019.04.030 (2019).
- 10 Pevny, L. H. & Nicolis, S. K. Sox2 roles in neural stem cells. *Int J Biochem Cell Biol* **42**, 421-424, doi:10.1016/j.biocel.2009.08.018 (2010).
- 11 Kalish, B. T. *et al.* Single-cell transcriptomics of the developing lateral geniculate nucleus reveals insights into circuit assembly and refinement. *Proc Natl Acad Sci U S A* **115**, E1051-E1060, doi:10.1073/pnas.1717871115 (2018).
- 12 Skene, P. J. & Henikoff, S. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *eLife* **6**, doi:10.7554/eLife.21856 (2017).
- 13 Bertolini, J. A. *et al.* Mapping the Global Chromatin Connectivity Network for Sox2 Function in Neural Stem Cell Maintenance. *Cell Stem Cell* **24**, 462-476 e466, doi:10.1016/j.stem.2019.02.004 (2019).
- 14 Pagin, M. *et al.* Sox2 controls neural stem cell self-renewal through a Fos-centered gene regulatory network. *Stem Cells*, doi:10.1002/stem.3373 (2021).
- 15 Pagin, M. *et al.* FOS Rescues Neuronal Differentiation of Sox2-Deleted Neural Stem Cells by Genome-Wide Regulation of Common SOX2 and AP1(FOS-JUN) Target Genes. *Cells* **10**, doi:10.3390/cells10071757 (2021).
- 16 Sato, H. *et al.* Thalamocortical axons control the cytoarchitecture of neocortical layers by area-specific supply of VGF. *eLife* **11**, doi:10.7554/eLife.67549 (2022).
- 17 Kwan, K. Y. *et al.* SOX5 postmitotically regulates migration, postmigratory differentiation, and projections of subplate and deep-layer neocortical neurons. *Proc Natl Acad Sci U S A* **105**, 16021-16026, doi:10.1073/pnas.0806791105 (2008).
- 18 Lai, T. *et al.* SOX5 controls the sequential generation of distinct corticofugal neuron subtypes. *Neuron* **57**, 232-247, doi:10.1016/j.neuron.2007.12.023 (2008).
- 19 McLean, C. Y. *et al.* GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol* **28**, 495-501, doi:10.1038/nbt.1630 (2010).

- 20 Szklarczyk, D. *et al.* The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res* **49**, D605-D612, doi:10.1093/nar/gkaa1074 (2021).
- 21 Xu, D. & Esko, J. D. Demystifying heparan sulfate-protein interactions. *Annu Rev Biochem* **83**, 129-157, doi:10.1146/annurev-biochem-060713-035314 (2014).
- 22 Kreuger, J. & Kjellen, L. Heparan sulfate biosynthesis: regulation and variability. *J Histochem Cytochem* **60**, 898-907, doi:10.1369/0022155412464972 (2012).
- 23 Habuchi, H. & Kimata, K. Mice deficient in heparan sulfate 6-O-sulfotransferase-1. *Prog Mol Biol Transl Sci* **93**, 79-111, doi:10.1016/S1877-1173(10)93005-6 (2010).
- 24 Pratt, T., Conway, C. D., Tian, N. M., Price, D. J. & Mason, J. O. 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, doi:10.1523/JNEUROSCI.0505-06.2006 (2006).
- 25 Tillo, M. *et al.* 2- and 6-O-sulfated proteoglycans have distinct and complementary roles in cranial axon guidance and motor neuron migration. *Development* **143**, 1907-1913, doi:10.1242/dev.126854 (2016).
- 26 Hamilton, B. A. *et al.* Disruption of the nuclear hormone receptor RORalpha in staggerer mice. *Nature* **379**, 736-739, doi:10.1038/379736a0 (1996).
- 27 Steinmayr, M. *et al.* staggerer phenotype in retinoid-related orphan receptor alpha-deficient mice. *Proc Natl Acad Sci U S A* **95**, 3960-3965, doi:10.1073/pnas.95.7.3960 (1998).
- 28 Franzen, O., Gan, L. M. & Bjorkegren, J. L. M. PanglaoDB: a web server for exploration of mouse and human single-cell RNA sequencing data. *Database (Oxford)* **2019**, doi:10.1093/database/baz046 (2019).
- 29 Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21, doi:10.1093/bioinformatics/bts635 (2013).
- 30 Navarro Gonzalez, J. *et al.* The UCSC Genome Browser database: 2021 update. *Nucleic Acids Res* **49**, D1046-D1057, doi:10.1093/nar/gkaa1070 (2021).
- 31 Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323, doi:10.1186/1471-2105-12-323 (2011).
- 32 Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140, doi:10.1093/bioinformatics/btp616 (2010).
- 33 Wang, X., Park, J., Susztak, K., Zhang, N. R. & Li, M. Bulk tissue cell type deconvolution with multi-subject single-cell expression reference. *Nat Commun* **10**, 380, doi:10.1038/s41467-018-08023-x (2019).
- 34 Li, L. *et al.* SoxD genes are required for adult neural stem cell activation. *Cell reports* **38**, 110313, doi:10.1016/j.celrep.2022.110313 (2022).
- 35 Nakagawa, Y. & O'Leary, D. D. Dynamic patterned expression of orphan nuclear receptor genes RORalpha and RORbeta in developing mouse forebrain. *Dev Neurosci* **25**, 234-244, doi:10.1159/000072271 (2003).
- 36 Skene, P. J., Henikoff, J. G. & Henikoff, S. Targeted in situ genome-wide profiling with high efficiency for low cell numbers. *Nat Protoc* **13**, 1006-1019, doi:10.1038/nprot.2018.015 (2018).
- 37 Bushnell, B., Rood, J. & Singer, E. BBMerge - Accurate paired shotgun read merging via overlap. *PLoS one* **12**, e0185056, doi:10.1371/journal.pone.0185056 (2017).

- 38 Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**, R25, doi:10.1186/gb-2009-10-3-r25 (2009).
- 39 Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079, doi:10.1093/bioinformatics/btp352 (2009).
- 40 Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841-842, doi:10.1093/bioinformatics/btq033 (2010).
- 41 Nordin, A., Zambanini, G., Pagella, P. & Cantu, C. The CUT&RUN suspect list of problematic regions of the genome. *Genome Biol* **24**, 185, doi:10.1186/s13059-023-03027-3 (2023).
- 42 Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137, doi:10.1186/gb-2008-9-9-r137 (2008).
- 43 Ge, S. X., Jung, D. & Yao, R. ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics* **36**, 2628-2629, doi:10.1093/bioinformatics/btz931 (2020).
- 44 Pagin, M. *et al.* Sox2 controls neural stem cell self-renewal through a Fos-centered gene regulatory network. *Stem Cells* **39**, 1107-1119, doi:10.1002/stem.3373 (2021).
- 45 Szklarczyk, D. *et al.* The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res* **51**, D638-D646, doi:10.1093/nar/gkac1000 (2023).
- 46 Robinson, J. T. *et al.* Integrative genomics viewer. *Nat Biotechnol* **29**, 24-26, doi:10.1038/nbt.1754 (2011).

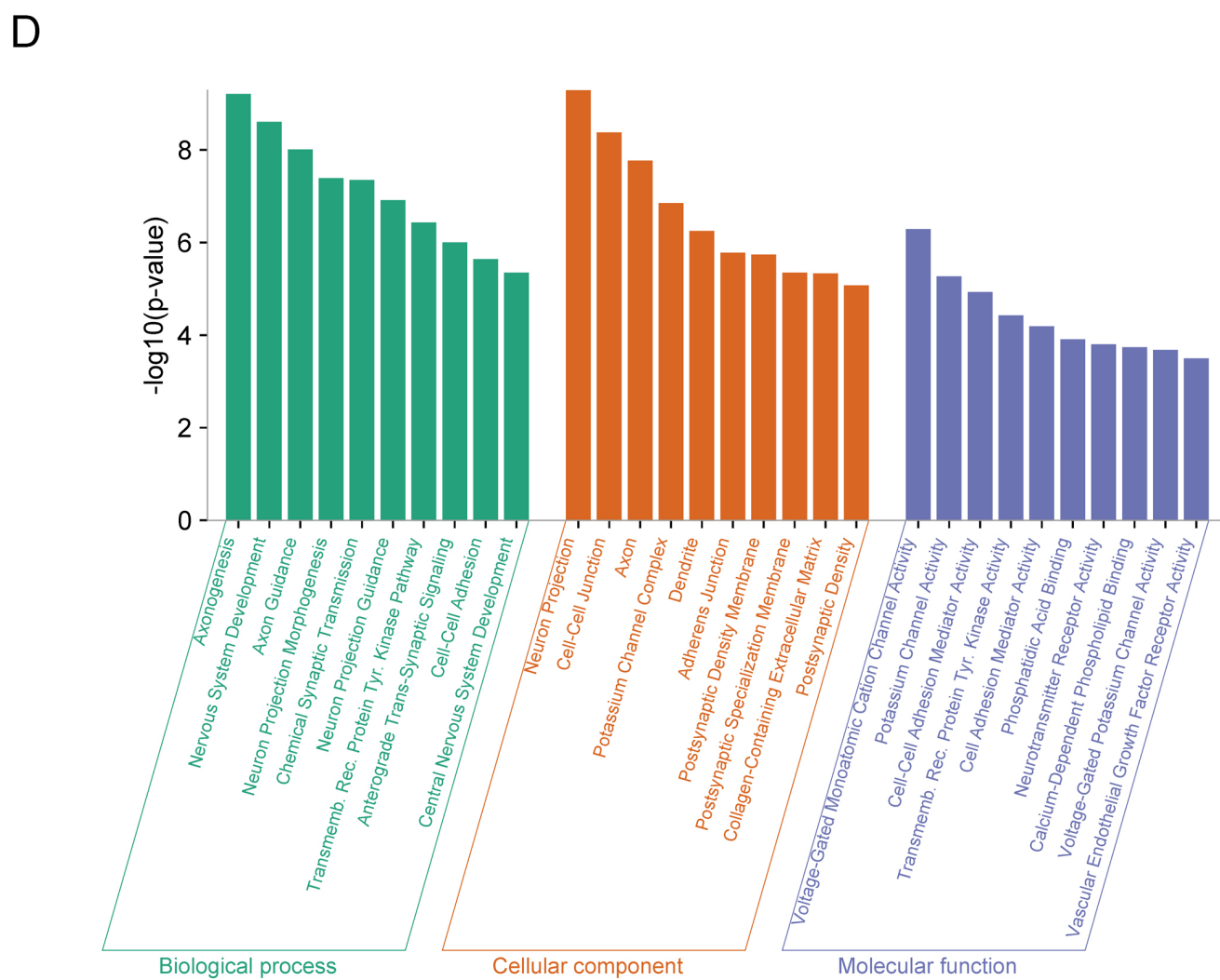
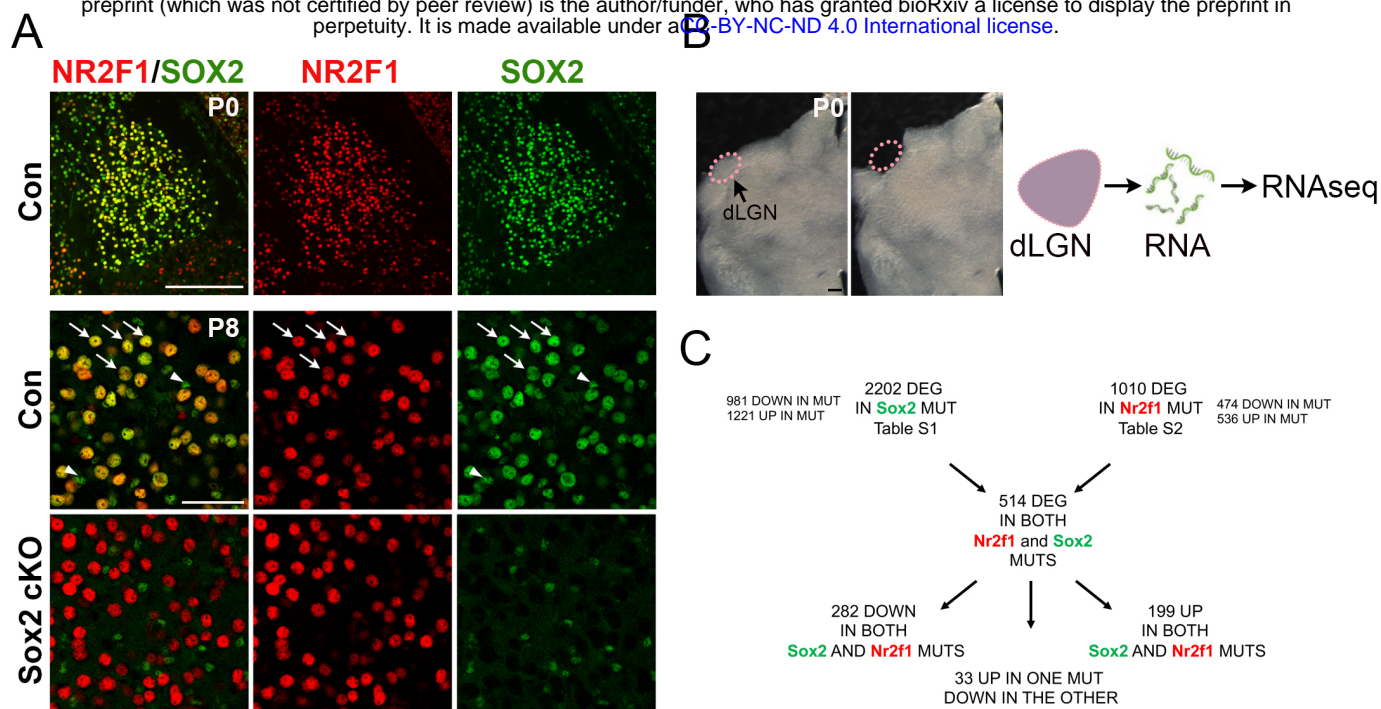
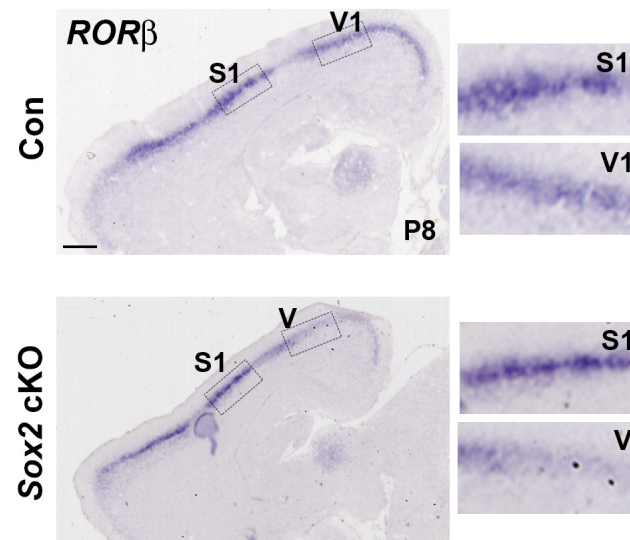


FIGURE 1

A



B

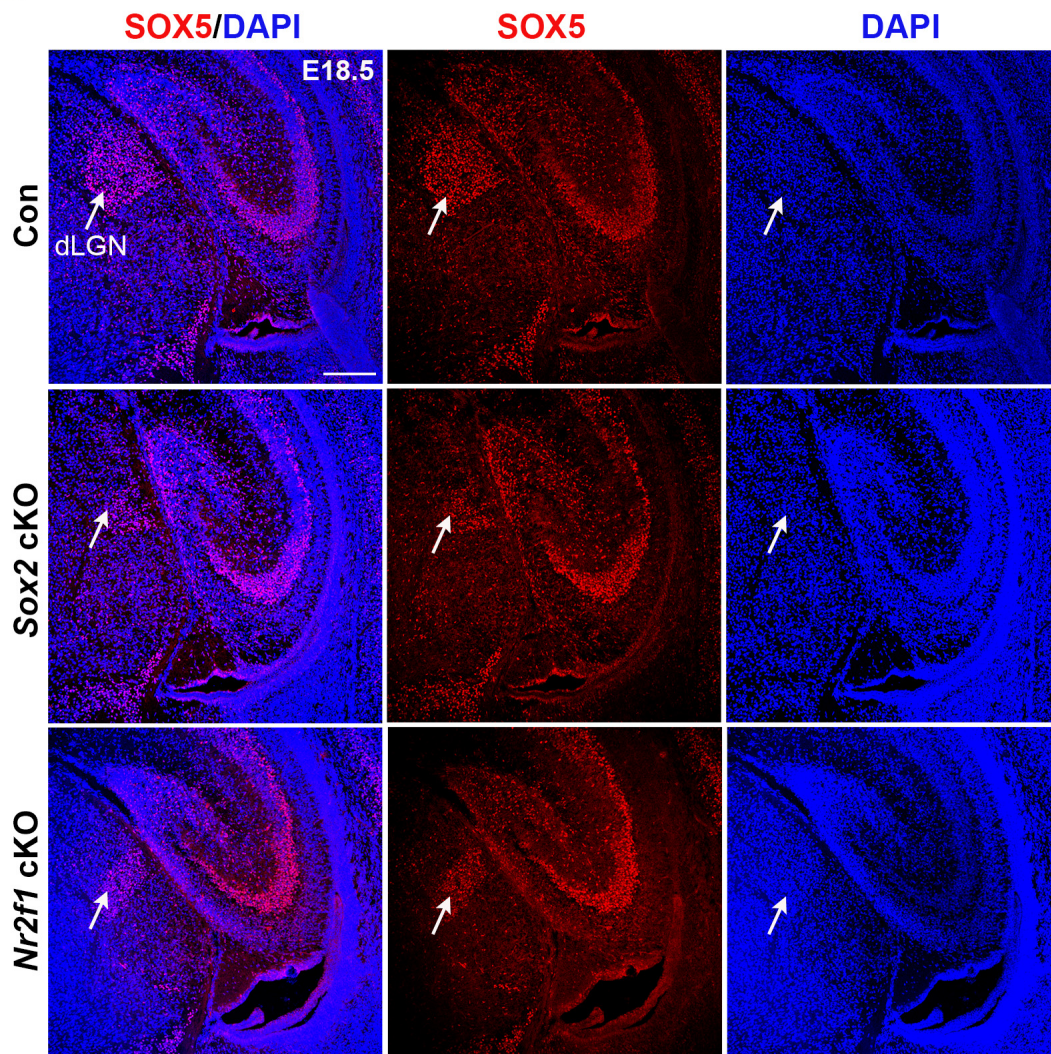


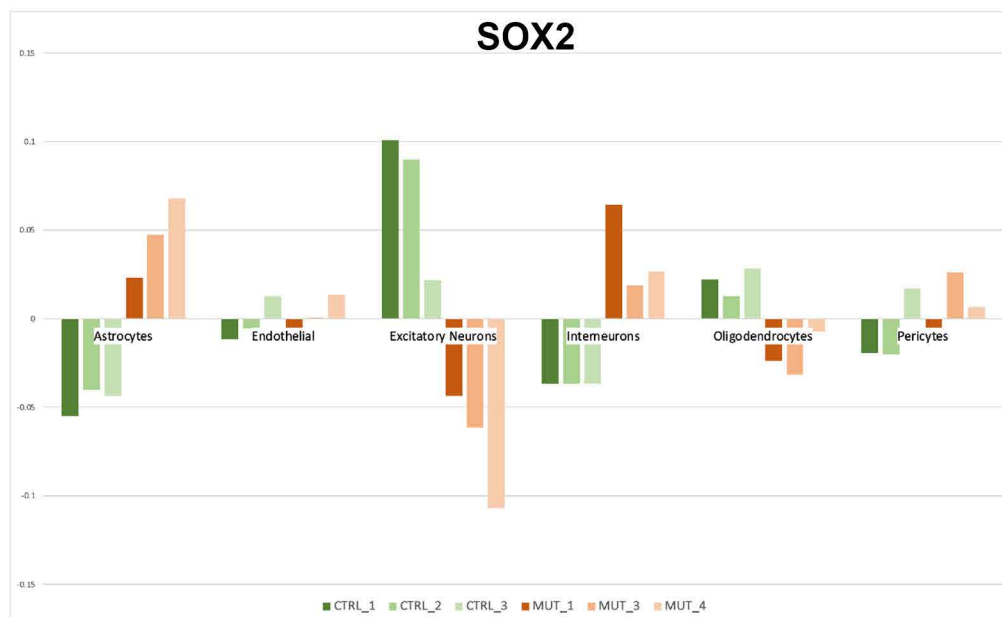
FIGURE 2

A

SOX2	Astrocytes	Endothelial	Excitatory Neurons	Interneurons	Macrophages	Microglia	Oligodendrocytes	Pericytes
CTRL_1	0.213081271	0.022375771	0.526474487	0	0	0	0.191123656	0.046944814
CTRL_2	0.227754199	0.028452389	0.515665914	0	0	0	0.181831305	0.046296193
CTRL_3	0.224258948	0.046979395	0.447569773	0	0	0	0.197510047	0.083681837
MUT_1	0.291031764	0.024707894	0.38233879	0.101241455	0	0	0.145002937	0.055677159
MUT_3	0.315334175	0.034585251	0.364612072	0.055663882	0	0	0.137209494	0.092595126
MUT_4	0.335715419	0.047484279	0.318867819	0.063333098	0	0	0.161753642	0.072845743
AVG_CTRL	0.221698139	0.032602518	0.496570058	0	0	0	0.190155003	0.058974281
AVG_MUT	0.314027119	0.035592475	0.355272894	0.073412812	0	0	0.147988691	0.073706009
DELTA (MUT-CTRL)	0.09232898	0.002989956	-0.141297165	0.073412812	0	0	-0.042166311	0.014731728

NR2F1	Astrocytes	Endothelial	Excitatory Neurons	Interneurons	Macrophages	Microglia	Oligodendrocytes	Pericytes
CTRL_B	0.184913058	0.02164381	0.585684132	0	0	0	0.169820051	0.03793895
CTRL_D	0.179982364	0.030975147	0.553928848	0	0	0	0.182119058	0.052994583
CTRL_E	0.198041351	0.024037446	0.574918539	0	0	0	0.150452908	0.052549756
MUT_A	0.175299218	0.040911253	0.481892139	0.055084815	0	0	0.169646459	0.077166116
MUT_B	0.204066367	0.029008107	0.49729486	0.023911725	0	0	0.195928516	0.049790426
MUT_C	0.177804144	0.041713906	0.469749899	0.069657652	0	0	0.16239597	0.078678429
AVG_CTRL	0.187645591	0.025552134	0.571510506	0	0	0	0.167464005	0.047827763
AVG_MUT	0.185723243	0.037211088	0.482978966	0.049551397	0	0	0.175990315	0.06854499
DELTA (CTRL-MUT)	-0.00192235	0.011658954	-0.08853154	0.049551397	0	0	0.00852631	0.020717227

B



C

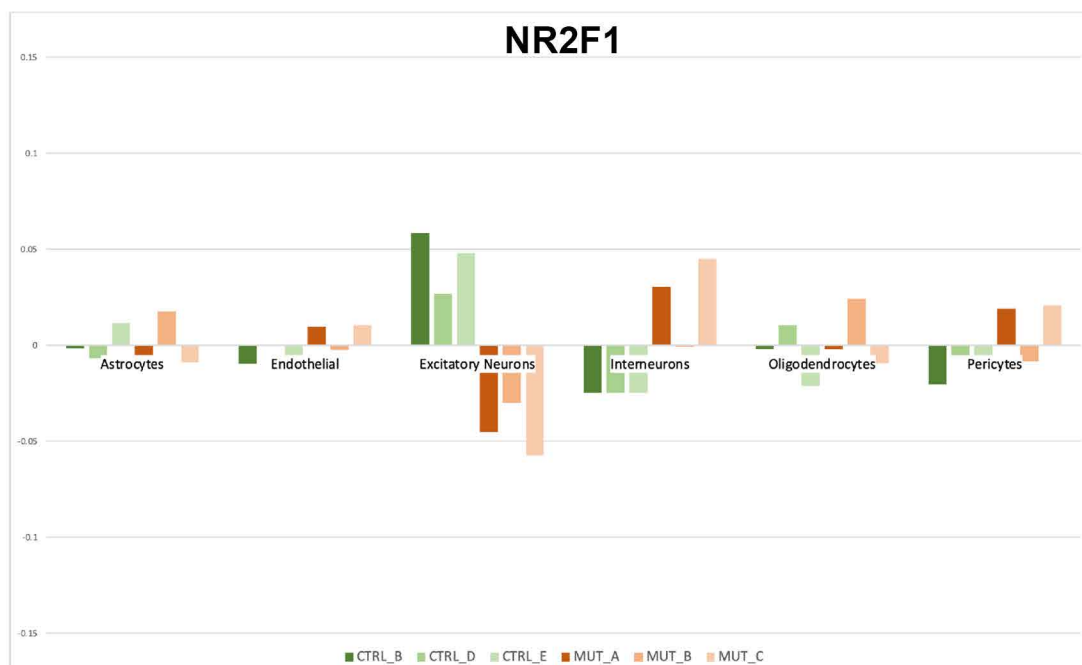


FIGURE 3

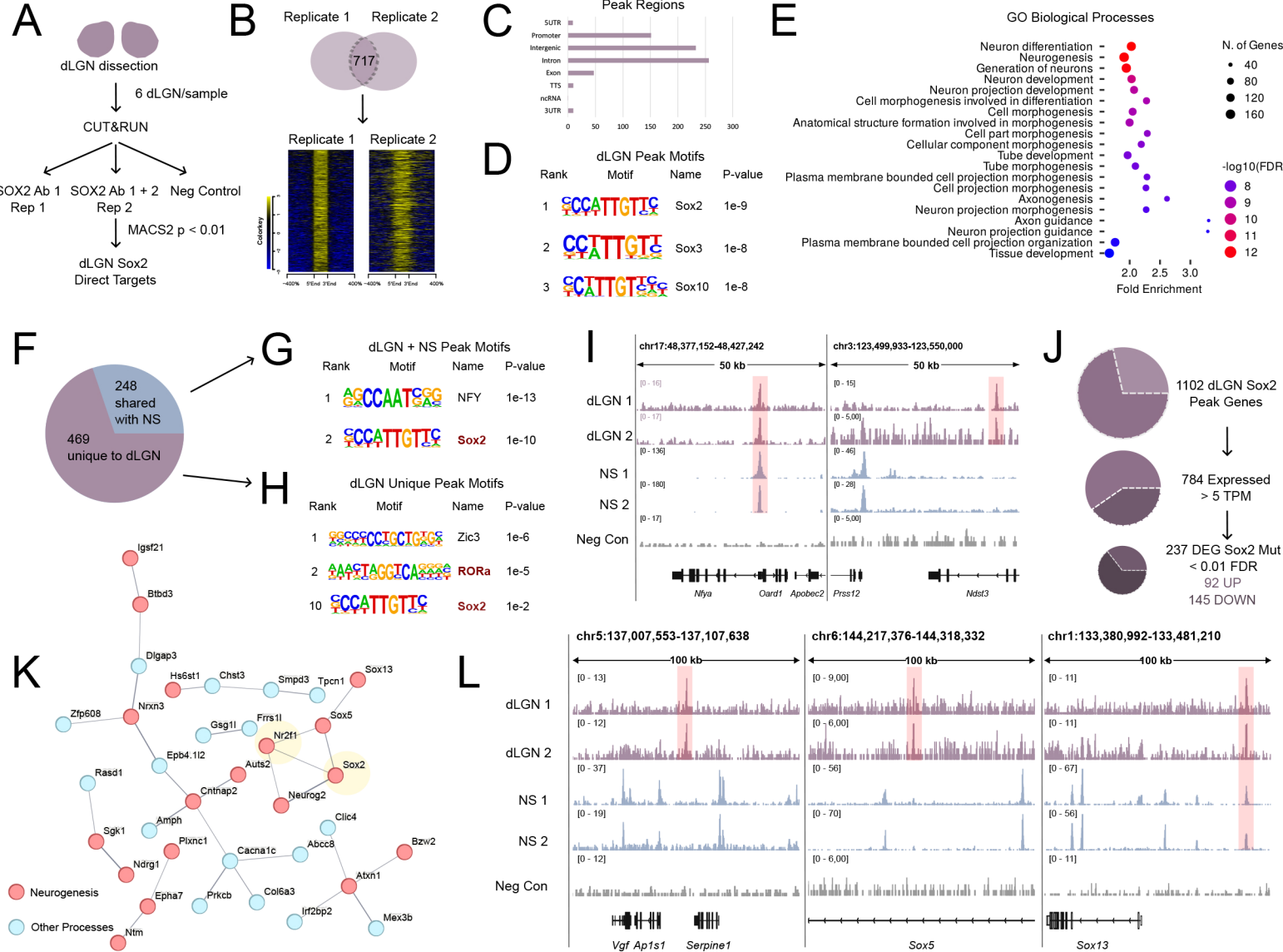


FIGURE 4