

ARTICLE

Open Access

# Long noncoding RNA *Sox2ot* and transcription factor YY1 co-regulate the differentiation of cortical neural progenitors by repressing *Sox2*

Jennifer L. Knauss<sup>1</sup>, Nan Miao<sup>2</sup>, Seung-Nam Kim<sup>1,3</sup>, Yanzhen Nie<sup>4</sup>, Yuelin Shi<sup>4</sup>, Tao Wu<sup>5,6</sup>, Hugo Borges Pinto<sup>5,6</sup>, Mary E. Donohoe<sup>5,6</sup> and Tao Sun<sup>1,2</sup>

## Abstract

Long noncoding RNAs (lncRNAs) are emerging as key regulators of crucial cellular processes. However, the molecular mechanisms of many lncRNA functions remain uncharacterized. *Sox2ot* is an evolutionarily conserved lncRNA that transcriptionally overlaps the pluripotency gene *Sox2*, which maintains the stemness of embryonic stem cells and tissue-specific stem cells. Here, we show that *Sox2ot* is expressed in the developing mouse cerebral cortex, where it represses neural progenitor (NP) proliferation and promotes neuronal differentiation. *Sox2ot* negatively regulates self-renewal of neural stem cells, and is predominately expressed in the nucleus and inhibits *Sox2* levels. *Sox2ot* forms a physical interaction with a multifunctional transcriptional regulator YY1, which binds several CpG islands in the *Sox2* locus in a *Sox2ot*-dependent manner. Similar to *Sox2ot*, YY1 represses NP expansion in vivo. These results demonstrate a regulatory role of *Sox2ot* in promoting cortical neurogenesis, possibly by repressing *Sox2* expression in NPs, through interacting with YY1.

## Introduction

Production of distinct types of neural progenitors and neurons is regulated by coding and noncoding RNAs in the mammalian cerebral cortex<sup>1–5</sup>. Long noncoding RNAs (lncRNAs), generally considered as transcripts of hundreds of nucleotides (nt) in length with little or no protein-coding potential, comprise a large proportion of the mammalian genome<sup>6–8</sup>. Though thousands of lncRNAs have been reported, relatively few have been mechanistically characterized. Those that have been characterized appear to work through a wide variety of mechanisms, including control of chromatin structure,

transcription, mRNA processing, and translation<sup>9,10</sup>. Emerging studies have shown a specific and dynamic expression of lncRNAs in embryonic stem (ES) cells, tissue-specific stem cells, and progenitors in various species. However, the molecular mechanisms of lncRNA function in specific cells and tissues and their relationship to diseases remain unclear<sup>11–13</sup>.

*Sox2* is a pluripotency gene that maintains the stemness of human and rodent ES cells<sup>14–16</sup>. However, the molecular regulation of *Sox2* expression is incompletely understood. Several CpG islands are harbored in the *Sox2* locus, which cover and flank the *Sox2* gene. This suggests a possible role for CpG islands in *Sox2* regulation by attracting transcription factors for transcriptional initiation, or propagating transcriptional silencing via DNA methylation<sup>17</sup>. For instance, the *Sox2* promoter in ES cells is marked by permissive H3K4me3 histone marks, while the surrounding CpG islands are bivalently marked,

Correspondence: Tao Sun (taosun@hqu.edu.cn)

<sup>1</sup>Department of Cell and Developmental Biology, Cornell University Weill Medical College, 1300 York Avenue, Box 60, New York, NY 10065, USA

<sup>2</sup>Center for Precision Medicine, School of Medicine and School of Biomedical Sciences, Huaqiao University, Xiamen, China 361021

Full list of author information is available at the end of the article.

Edited by G. Raschella

© The Author(s) 2018



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

indicating that the locus may also be poised for repression<sup>18</sup>. *Sox2* is also highly expressed in neural stem cells (NSCs) and neural progenitors (NPs) in the mouse embryonic cerebral cortex, maintains the populations of NSCs and NPs, and promotes proliferation<sup>19,20</sup>. *Sox2* expression is downregulated in postmitotic, differentiated neurons; however, the regulatory mechanisms governing this transition are incompletely understood<sup>16,21–24</sup>.

Interestingly, the lncRNA *Sox2* overlapping transcript (*Sox2ot*) is highly conserved among species and overlaps the *Sox2* gene in the genome<sup>25–27</sup>. *Sox2ot* appears to be transcribed from several transcription start sites and can be differentially spliced, thus generating several *Sox2ot* isoforms<sup>25,28</sup>. While the function of *Sox2* is defined, the roles of *Sox2ot* in development, stem cell expansion, and differentiation are unknown.

In this study, we demonstrate that *Sox2ot* negatively regulates neural progenitor proliferation by interacting with the epigenetic regulator YY1 in the developing mouse cerebral cortex. Overexpression of *Sox2ot* causes a decrease in NPs with a concomitant increase in neurons, while *Sox2ot* knockdown results in increased NPs and a loss of differentiated neurons. There is a strong interaction between *Sox2ot* and YY1 in the nucleus of neuroectodermal cells. Furthermore, we found that YY1 binds to CpG islands at the *Sox2* locus and represses NP proliferation, and that this binding is dependent on *Sox2ot*. Thus, we propose that *Sox2ot* exerts its effects in NP development with its RNA-binding partner YY1 through negative regulation of *Sox2* in the developing cortex.

## Results

### *Sox2ot* and *Sox2* are coexpressed in the ventricular zone in the developing mouse cortex

The main isoform of *Sox2ot* (GenBank Accession No. BC057611) contains five exons spanning >100 kilobases (kb) of the mouse genome, though the final mRNA is spliced to about 3 kb. *Sox2*, a single-exon transcript of about 2.5 kb, is located within one of the large introns of *Sox2ot* (Fig. 1a). None of the known isoforms of *Sox2ot* contain any overlapping exonic sequence with *Sox2*<sup>25</sup>. Given the overlapping genomic orientation of *Sox2* and *Sox2ot*, we performed in situ hybridization on embryonic mouse cerebral cortices to examine the expression patterns of each of these genes. Both *Sox2* and *Sox2ot* were expressed in the ventricular zone (VZ) and subventricular zone (SVZ) of cortices at embryonic day 13.5 (E13.5) and E15.5, which correspond to the active stages of neural progenitor expansion (Fig. 1b–e). The expression level of *Sox2ot* was not as robust as *Sox2*. Expression of *Sox2* and *Sox2ot* was reduced in postnatal day 0 (P0) cortices, corresponding with a reduction in the NP population (Supplementary Figure S1A). Moreover, sense probes for

*Sox2* and *Sox2ot* lacked specific staining in E13.5 cortices (Supplementary Figure S1B).

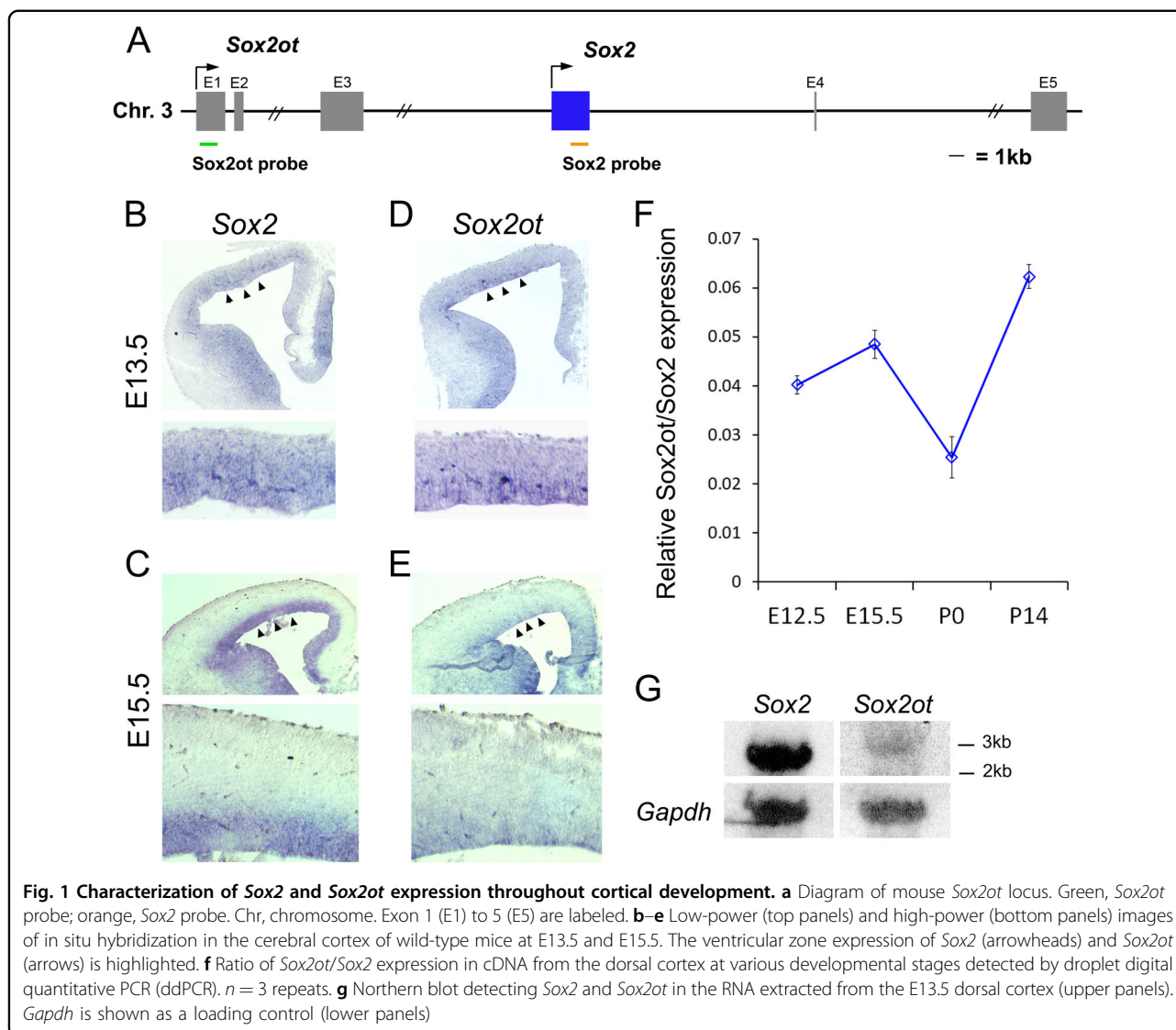
To establish a more quantitative picture of *Sox2* and *Sox2ot* expression levels, we performed droplet digital quantitative PCR (ddPCR) to test their mRNA copy numbers in the dorsal cortex. From E12.5 to P14, the mRNA copy numbers of *Sox2* and *Sox2ot* showed a reduced trend (Supplementary Figure S1C and D). Moreover, *Sox2* was expressed at a higher level than *Sox2ot* at all developmental stages tested, and the ratio of *Sox2ot/Sox2* expression was slightly increased, suggesting an elevated expression of *Sox2ot* through development (Fig. 1f).

To further verify the expression levels and sizes of *Sox2* and *Sox2ot* transcripts, we performed Northern blot analyses on RNA extracted from E13.5 mouse dorsal cortices (Fig. 1g). As expected, *Sox2* was expressed at a higher level than *Sox2ot*, which reflected the expression observed in the analyses of ddPCR (Fig. 1b–f). Both *Sox2* and *Sox2ot* transcripts were detected at the expected sizes, 2.5 kb and 3 kb, respectively (Fig. 1g). Taken together, these experiments establish that *Sox2ot* is expressed in the VZ, where NPs reside in the developing cortex.

### *Sox2ot* overexpression causes a reduction of neural progenitors

To explore the function of *Sox2ot* in NPs, we first overexpressed *Sox2ot* by in utero electroporation (IUE) at E13.5 and collected the brains at E14.5 for analysis. A bromodeoxyuridine (BrdU) pulse was given 1 h before tissue collection to label the dividing cells in the S-phase in a cell cycle. Electroporation in the mouse embryonic cortex will first label neural progenitors in the VZ. To avoid biased analyses, all quantification in this study was based on calculating the ratio of marker<sup>+</sup>/GFP<sup>+</sup> versus GFP<sup>+</sup>. Measures of the general number of NPs with BrdU incorporation (BrdU<sup>+</sup>/GFP<sup>+</sup> cells versus GFP<sup>+</sup> cells) and *Sox2* staining (*Sox2*<sup>+</sup>/GFP<sup>+</sup> cells versus *Sox2*<sup>+</sup> cells) were both decreased in *Sox2ot* overexpression (*OE*) cortices compared to controls (*Ctrl*) (Fig. 2a–d). The reduction in the numbers of *Sox2*<sup>+</sup> cells in GFP<sup>+</sup> cells suggests the potential for direct negative regulation of *Sox2* by *Sox2ot*. Two specific NP populations, radial glia cells (RGCs) and intermediate progenitors (IPs) that can be labeled by Pax6 and Tbr2, respectively, also showed a decrease in the *Sox2ot* *OE* cortices compared to controls, by calculating the ratios of Pax6<sup>+</sup>/GFP<sup>+</sup> and Tbr2<sup>+</sup>/GFP<sup>+</sup> cells versus GFP<sup>+</sup> cells (Supplementary Figure S2A–D).

To confirm these results quantitatively, we dissected the GFP-positive area of the cortex 24 h after E13.5 electroporation and extracted RNA for real-time quantitative reverse transcription PCR (qRT-PCR) (Supplementary Figure S3). As expected, *Sox2ot* was upregulated in *Sox2ot* *OE* brains, while *Sox2* was downregulated, reflecting the



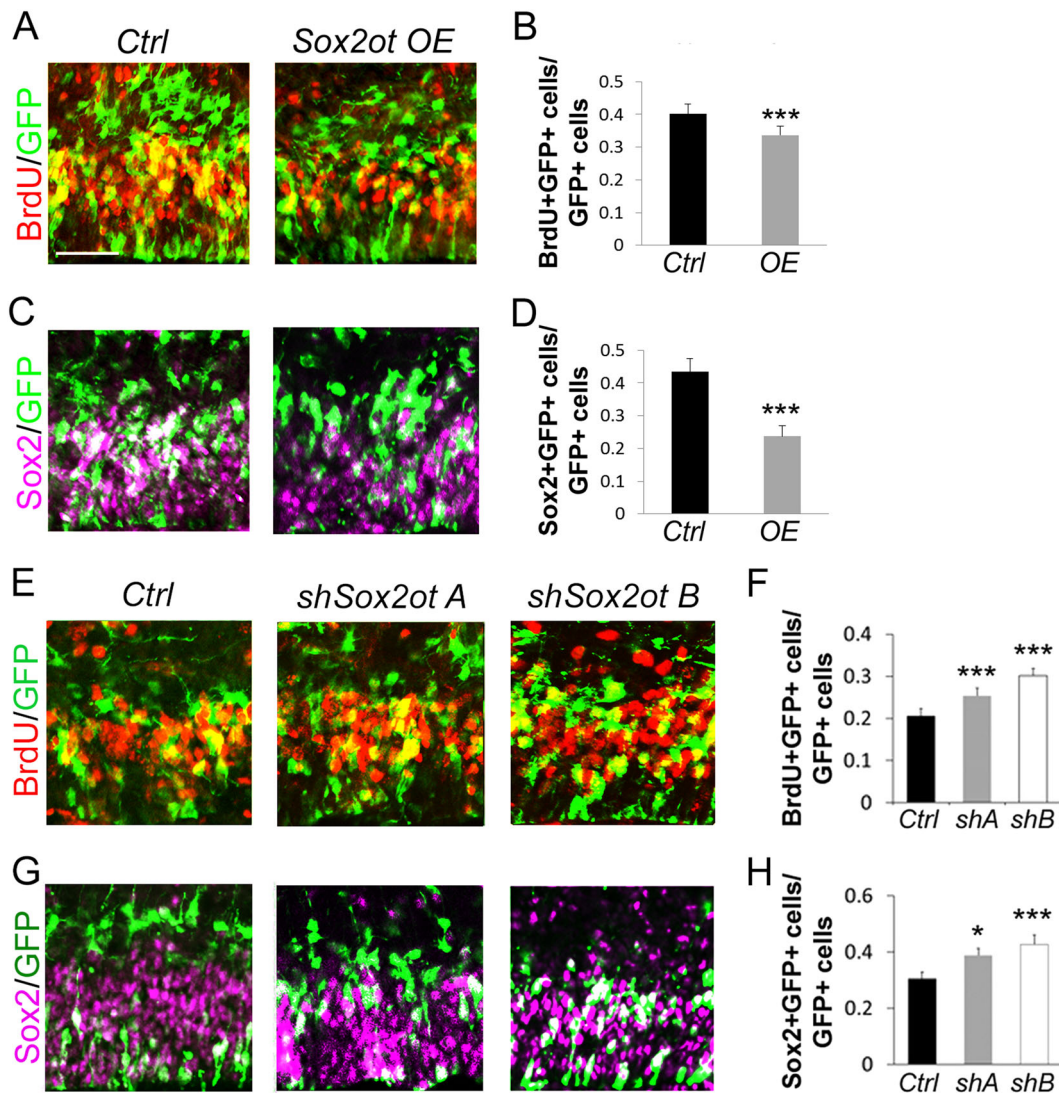
immunostaining results (Supplementary Figure S3A). *Pax6* and *Tbr2* RNA levels also mirror protein-level changes upon *Sox2ot* OE (Supplementary Figure S3B).

Moreover, we hypothesized that the loss of NPs upon *Sox2ot* OE could be explained by cell death or by differentiation. We then examined cell death using activated Caspase3 as a marker. In E14.5 brains, activated Caspase3 expression was not significantly altered due to *Sox2ot* overexpression, indicating that cell death does not explain the loss of NPs (Supplementary Figure S2E and F). These results demonstrate that *Sox2ot* OE represses NP expansion in both RGCs and IPs, and that this function is not carried out through cell death.

#### Sox2ot knockdown increases neural progenitors

To further examine the effect of *Sox2ot* on NP expansion, we applied a knockdown approach using short

hairpin RNAs (shRNAs) for *Sox2ot*, named as *shSox2otA* (*shA*) and *shSox2otB* (*shB*) (Supplementary Figure S4A). In cell culture testing, both *shA* and *shB* significantly reduced *Sox2ot* expression, though *shB* had a slightly stronger effect (Supplementary Figure S4B). We again used the IUE technique to knock down *Sox2ot* in the developing cortex. Both BrdU incorporation and Sox2 staining increased in *shSox2ot* brains compared to control brains (Fig. 2e–h). The specific progenitor population markers for RGCs and IPs also showed increases upon *Sox2ot* knockdown (Supplementary Figure S5). In each case, the increase in NPs is slightly larger for *shB* than *shA*, though both are significant, reflecting the strength of knockdown observed in cell culture experiments (Fig. 2; Supplementary Figure S4B). Our qRT-PCR of electroporated brains also confirmed these results, showing *Sox2ot* knockdown, *Sox2* increase, and



**Fig. 2 Manipulation of *Sox2ot* affects the development of neural progenitors.** a–d Electroporation of ectopic *Sox2ot* (OE) at E13.5 for analysis at E14.5 significantly decreased the number of BrdU-incorporating or Sox2<sup>+</sup> cells co-labeled with GFP in the cortex, compared to the control (Ctrl). e–h Electroporation of shRNAs against *Sox2ot* (*shSox2ot A* and *shSox2ot B*) at E13.5 for analysis at E14.5 significantly increased the number of BrdU-incorporating or Sox2<sup>+</sup> cells co-labeled with GFP in the cortex. Yellow and white cells indicate co-labeled cells. Data are presented as mean±SD;  $n \geq 5$  sections from at least four different brains;  $p$  values in relation to the empty vector control (\* $p < 0.05$ , \*\*\* $p < 0.001$ ). Scale bar=50  $\mu$ m

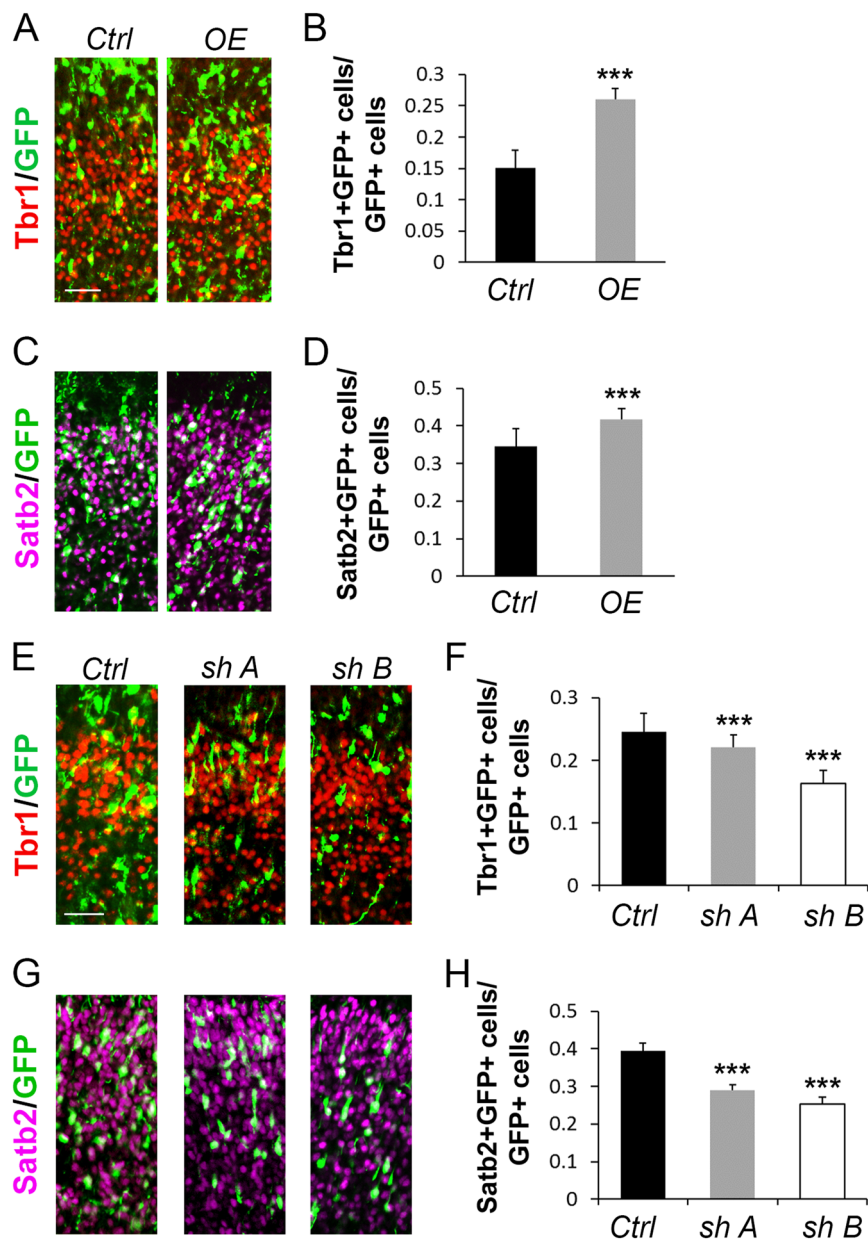
upregulation of both *Pax6* and *Tbr2* upon electroporation of *shSox2otA* (Supplementary Figure S4C and D). These observations further support the functionality of *Sox2ot* as a negative regulator in NP expansion in the developing cortex.

#### **Sox2ot affects neural differentiation**

To determine the effects of *Sox2ot* OE on the differentiation of NPs, we electroporated the *Sox2ot* OE construct at E13.5 and collected brain tissues 4 days after electroporation at E17.5, thus allowing more time for *Sox2ot* to exert its effects on neurogenesis. Two markers

for differentiated neurons, *Tbr1* and *Satb2*, which label early-born and late-born neurons, respectively, each showed an increase with *Sox2ot* OE (Fig. 3a–d). This expansion of differentiated neurons accounts for the loss of NPs observed in 1-day electroporations, suggesting an early differentiation (Fig. 2a–d). These results were confirmed at the RNA level for *Tbr1* by qRT-PCR of RNA extracted from electroporated cells in the cortex (Supplementary Figure S3B).

We next examined the effect of *Sox2ot* knockdown on neuron populations and found that both *Tbr1* and *Satb2* expression was decreased, which is opposite to the effect

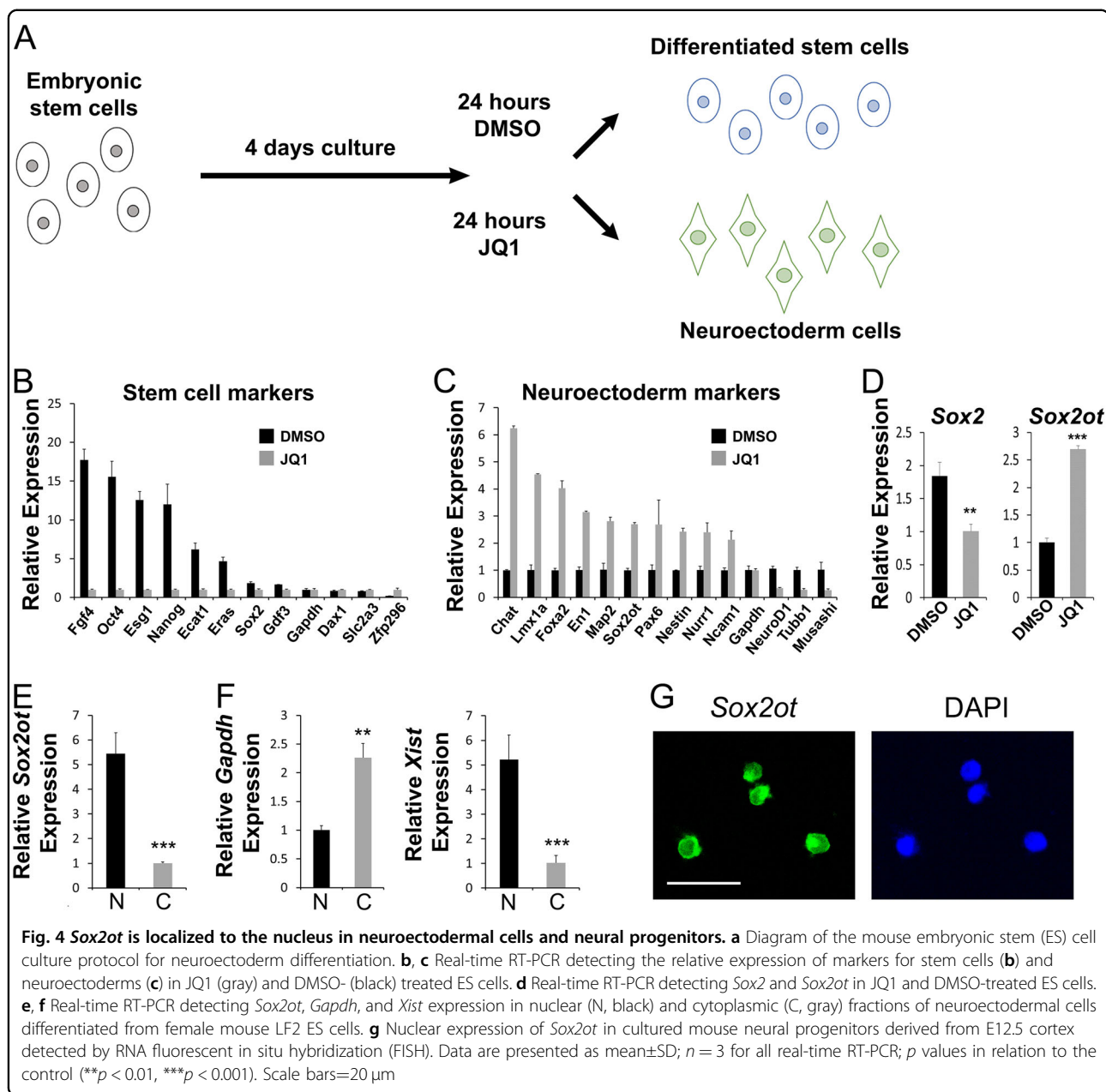


**Fig. 3 Loss of neural progenitors caused by altered *Sox2ot* is due to early differentiation.** **a–d** Electroporation of ectopic *Sox2ot* (OE) at E13.5 for analysis at E17.5 significantly increased the number of Tbr1<sup>+</sup> or Satb2<sup>+</sup> cells co-labeled with GFP in the cortex, compared to the control (Ctrl). **e–h** Electroporation of shRNAs against *Sox2ot* (*shSox2ot A* and *shSox2ot B*) at E13.5 for analysis at E17.5 significantly decreased the number of Tbr1<sup>+</sup> or Satb2<sup>+</sup> cells co-labeled with GFP in the cortex. Yellow and white cells indicate co-labeled cells. Data are presented as mean  $\pm$  SD;  $n \geq 5$  sections from at least four different brains;  $p$  values in relation to the empty vector control (\*\*\*)  $p < 0.001$ ). Scale bars: 50  $\mu$ m

of *Sox2ot* OE (Fig. 3e–h). Decreased neurogenesis corresponds to the increase in the neural progenitor pool upon *Sox2ot* knockdown (Fig. 2e–h). The observed decrease in Tbr1 was confirmed by qRT-PCR on RNA purified from electroporated brain cells (Supplementary Figure S4D). Collectively, these results establish a role for *Sox2ot* in the NP choice between proliferation and differentiation.

#### ***Sox2ot* is localized to the nucleus in neuroectodermal cells**

Having established that *Sox2ot* plays a role in regulating cortical neurogenesis, we sought to characterize how *Sox2ot* exerts its function. Because there are a wide variety of cell types present in the developing cortex, we turned to cultured and relatively homogeneous mouse ES cells (mESCs) and their derived neuroectodermal cells. mESCs were cultured for 4 days and then treated for 24 h with

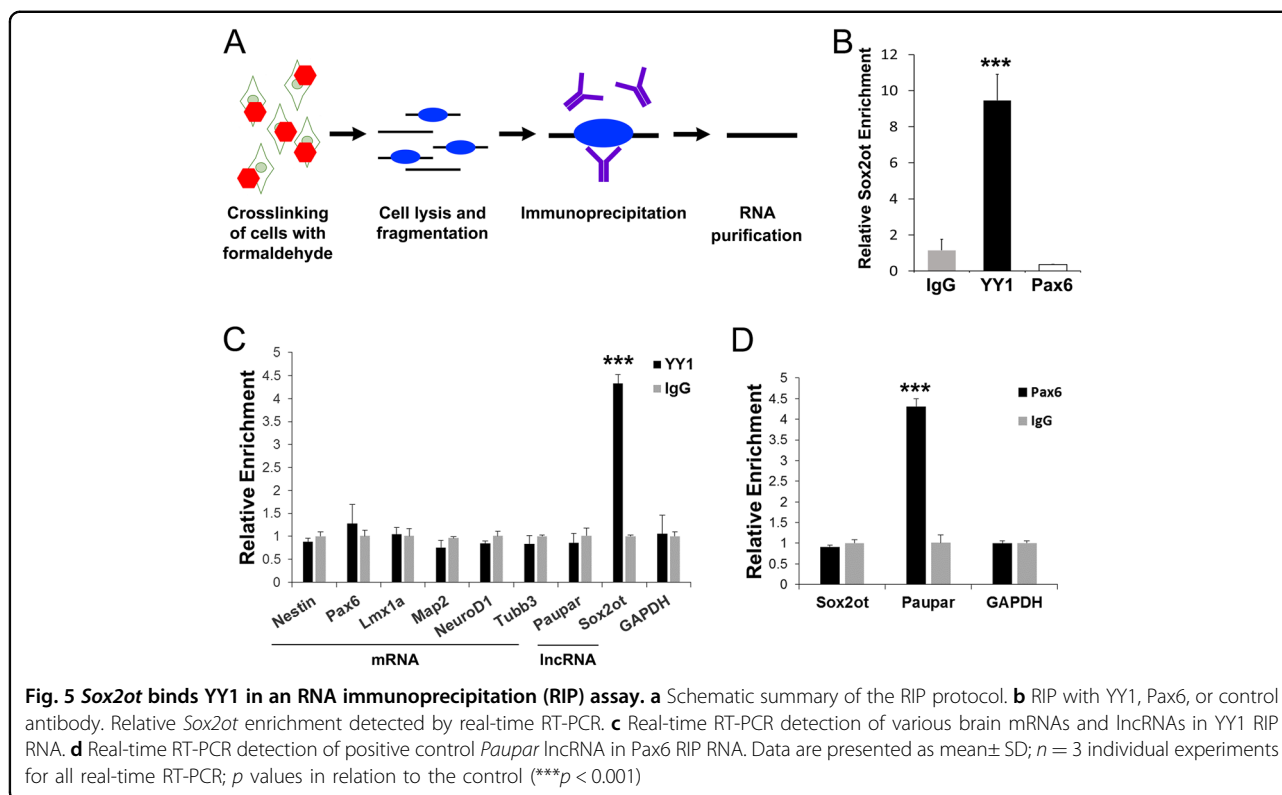


**Fig. 4** *Sox2ot* is localized to the nucleus in neuroectodermal cells and neural progenitors. **a** Diagram of the mouse embryonic stem (ES) cell culture protocol for neuroectoderm differentiation. **b, c** Real-time RT-PCR detecting the relative expression of markers for stem cells (**b**) and neuroectoderms (**c**) in JQ1 (gray) and DMSO- (black) treated ES cells. **d** Real-time RT-PCR detecting *Sox2* and *Sox2ot* in JQ1 and DMSO-treated ES cells. **e, f** Real-time RT-PCR detecting *Sox2ot*, *Gapdh*, and *Xist* expression in nuclear (N, black) and cytoplasmic (C, gray) fractions of neuroectodermal cells differentiated from female mouse LF2 ES cells. **g** Nuclear expression of *Sox2ot* in cultured mouse neural progenitors derived from E12.5 cortex detected by RNA fluorescent in situ hybridization (FISH). Data are presented as mean $\pm$ SD;  $n = 3$  for all real-time RT-PCR;  $p$  values in relation to the control (\*\* $p < 0.01$ , \*\*\* $p < 0.001$ ). Scale bars=20  $\mu$ m

either DMSO vehicle or the small molecule JQ1, a BET protein inhibitor shown to cause differentiation to a neuroectodermal fate<sup>29</sup> (Fig. 4a). To confirm the identity of mESCs after treatment, we performed qRT-PCR for various stem cell and neuronal markers. mESCs treated only with DMSO showed enrichment for stem cell markers, including *Fgf4*, *Oct4*, and *Nanog*, while mESCs treated with JQ1 expressed higher levels of neuroectodermal markers, such as *Nestin*, *Pax6*, and *Foxa2* (Fig. 4b, c). Notably, markers of more mature neurons, *NeuroD1* and *Tubb1*, were not enriched in the neuroectodermal cells, further validating their NP identity (Fig. 4c). Moreover, *Sox2* and *Sox2ot* were differentially

expressed between differentiated stem cells and neuroectodermal cells, with *Sox2* decreased in neuroectodermal cells while *Sox2ot* increased, reflecting the changes in *Sox2ot* and *Sox2* expression observed in vivo (Fig. 1f), and indicating that *Sox2ot* may be specific to the differentiation of the neural lineage (Fig. 4c, d).

As lncRNAs can function in both the nucleus and the cytoplasm, we next determined the subcellular localization of *Sox2ot*. Lysis of cells followed by low-speed centrifugation pellets the relatively large nuclei intact, while leaving the cytoplasm as the supernatant, allowing reliable separation of nuclear and cytoplasmic RNAs. Using RNA extracted from neuroectodermal cells differentiated from



female ES cells, we performed qRT-PCR for *Sox2ot* and determined that it is significantly enriched in the nuclear fraction (Fig. 4e). We also measured *Gapdh* and *Xist* as cytoplasmic and nuclear controls, respectively, and found that each of them was highly expressed in the expected fraction (Fig. 4f). Moreover, we examined *Sox2ot* expression in cultured mouse neural progenitors derived from E12.5 cortex using RNA fluorescent in situ hybridization (FISH). *Sox2ot* was mostly expressed in the nucleus (Fig. 4g).

These results demonstrate *Sox2ot* nuclear localization, suggesting that *Sox2ot* may play roles in transcriptional and epigenetic regulation of the switch in NP fate choice.

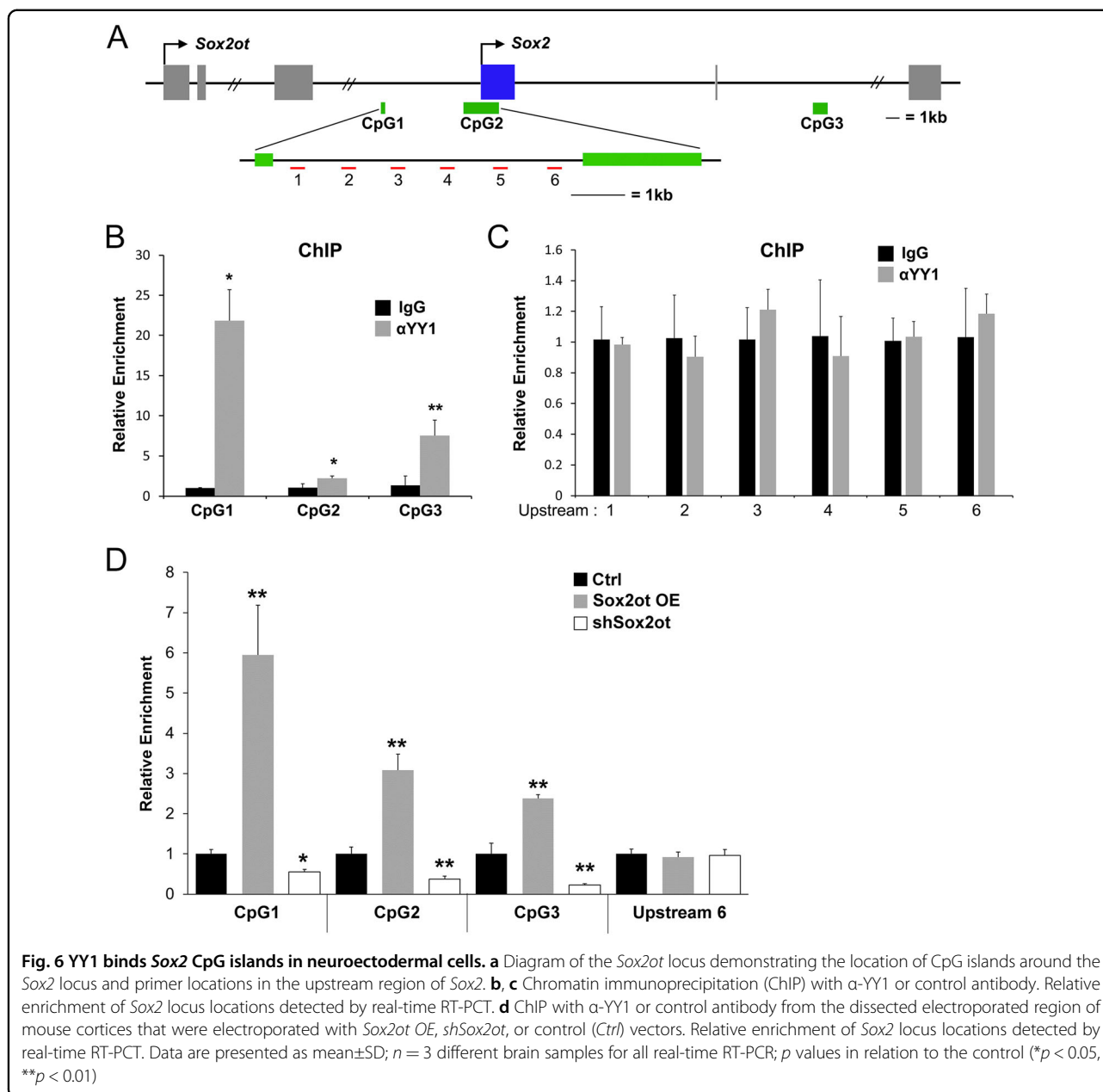
### Sox2ot binds YY1 in neuroectodermal cells

In some cases, the simple act of transcription constitutes lncRNA functionality, while in other cases, the lncRNA interacts with protein partners to exert its function<sup>30,31</sup>. After finding *Sox2ot* in the nucleus, we asked whether *Sox2ot* might interact with any proteins to repress *Sox2* expression. Using a candidate approach with several proteins previously shown to bind lncRNAs and to regulate target gene expression, we performed RNA immunoprecipitation (RIP) on neuroectodermal cells (Fig. 5a). The candidates included YY1, which is important for tethering *Xist* lncRNA to the X chromosome during X inactivation<sup>32</sup>; CTCF interacts with lncRNAs during imprinting and is a protein-interacting partner of

YY1<sup>33,34</sup>; and Pax6 interacts with lncRNA *Paupar* in neuroblastomal cells<sup>35</sup>. We found that *Sox2ot* was highly enriched by YY1 pulldown, while CTCF and Pax6 showed no significant enrichment over IgG control (Fig. 5b, data not shown). To ensure that the RIP experiment was specific, we tested the pulldown of several neural mRNAs such as *NeuroD1* and other lncRNAs (e.g., *Paupar*) by YY1 and found significant enrichment only of *Sox2ot* (Fig. 5c). As a positive control, we also tested the pull-down of the known Pax6-bound lncRNA *Paupar*, and found its significant enrichment with Pax6 (Fig. 5d)<sup>35</sup>. These results demonstrate a *Sox2ot* binding partner and give insight into the possible mechanisms of *Sox2ot* action through the multiple known functions of YY1.

### YY1 binds CpG islands in the *Sox2* locus

A subset of lncRNAs function by the *cis*-regulation of target genes near the same genomic locus<sup>36</sup>. Due to the overlapping genomic organization of the *Sox2/Sox2ot* locus (Fig. 1a), the similar expression patterns of *Sox2* and *Sox2ot* (Fig. 1b–e), the negative effects of *Sox2ot* on *Sox2* expression (Fig. 2c, d, Supplementary Figure S3A), and the presence of favorable CpG island binding sites in the *Sox2* locus (Fig. 6a), we hypothesized that *Sox2ot* and its binding partner YY1 might function through repression of *Sox2*. Thus, we performed chromatin immunoprecipitation (ChIP) on chromatin in neuroectodermal cells to determine if YY1 could be found at the *Sox2* locus



(Fig. 6a). Indeed, YY1 was found to bind at each of the three CpG islands in the *Sox2* locus (Fig. 6b). YY1 was not detected in the non-CpG island upstream region of the *Sox2* transcription start site (Fig. 6c).

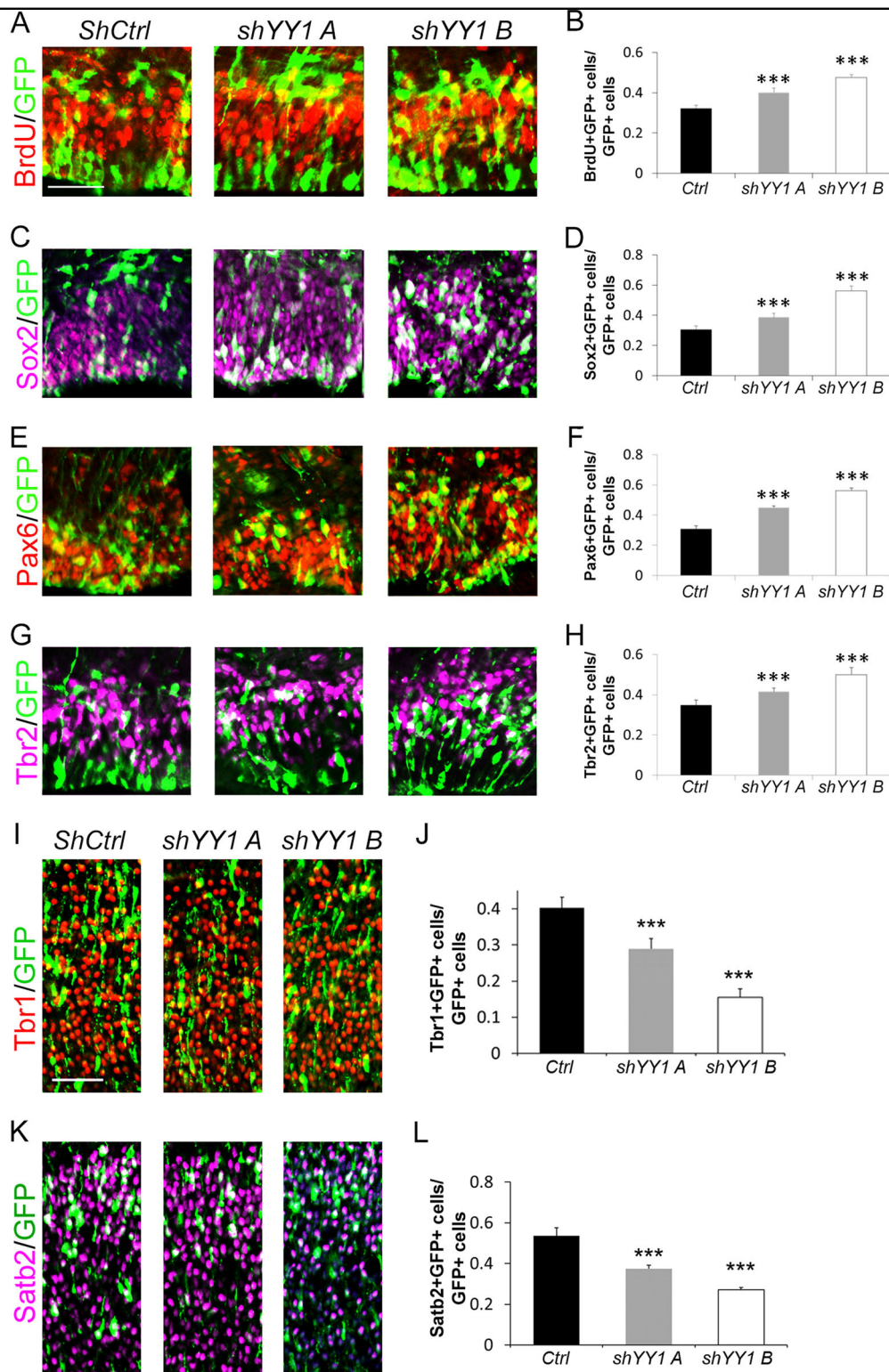
To address the dependence of YY1 binding on *Sox2ot*, we performed ChIP of YY1 binding at each *Sox2* CpG island under *Sox2ot* overexpression or knockdown conditions. Upon overexpression of *Sox2ot*, YY1 binding at each CpG island was enriched, while knockdown of *Sox2ot* depleted YY1 binding (Fig. 6d and Supplementary Figure S6). As a control, we also tested YY1 binding at upstream region #6, and found no effect from *Sox2ot* manipulation (Fig. 6d). These results support a direct

regulatory role for *Sox2ot* and YY1 in repression of *Sox2* in NPs.

#### YY1 knockdown causes an increase in neural progenitors

To further validate the functional relationship between *Sox2ot* and YY1, we performed IUE to knock down YY1 in developing cortices using two different shRNAs, *shYY1A* and *shYY1B* (Fig. 7; Supplementary Figure S3C). We examined NPs at E14.5, 1 day after IUE, and found that the general progenitor cell markers BrdU and *Sox2*, both increased upon YY1 knockdown (Fig. 7a–d). The specific RGC and IP markers Pax6 and Tbr2 also showed increased expression in *shYY1* cortices (Fig. 7e–h). We





**Fig. 7 YY1 knockdown increases neural progenitors. a–h** Electroporation of shRNAs against *YY1* (*shYY1 A* and *shYY1 B*) at E13.5 for analysis at E14.5 significantly increased the number of BrdU-incorporating or Sox2<sup>+</sup>, Pax6<sup>+</sup>, or Tbr2<sup>+</sup> cells co-labeled with GFP in the cortex, compared to the scrambled shRNA control (*ShCtrl*). **i–l** Electroporation of shRNAs against *YY1* at E13.5 for analysis at E17.5 significantly decreased the number of Tbr1<sup>+</sup> or Satb2<sup>+</sup> cells co-labeled with GFP in the cortex. Yellow and white cells indicate co-labeled cells. Data are presented as mean±SD; *n* ≥ 5 sections from at least four different brains for all electroporations; *p* values in relation to the scrambled shRNA control (\*\*\*) *p* < 0.001). Scale bar=50 μm

next examined the changes in cortical neurogenesis due to *YY1* knockdown using differentiated neuronal markers *Tbr1* and *Satb2*. We found that both early- and late-born neurons were decreased (Fig. 7i–l).

Furthermore, changes in both the neural progenitor and mature neuronal populations caused by *YY1* knockdown were confirmed by RT-PCR of RNA from electroporated brains (Supplementary Figure S4C and D). These results are very similar to those observed upon *Sox2ot* knockdown in vivo (Figs. 2 and 3, Supplementary Figure S4C and D). We also performed RT-PCR for *YY1* in RNA extracted from brains electroporated with *Sox2ot* OE and *shSox2ot* to ensure that the observed phenotypes are not due to changes in *YY1* expression upon *Sox2ot* manipulation (Supplementary Figure S3C). Altogether, these results further support the functional relationship between *Sox2ot* and *YY1*.

## Discussion

As lncRNAs are continually annotated and validated, it has become clear that they are potent regulators of many cell processes. However, the mechanistic explorations of lncRNA functions are lacking. Here, we characterized the function of lncRNA *Sox2ot* in the developing mouse cerebral cortex through in vivo methods, determining that it promotes the differentiation of NPs into neurons, and suppresses the expansion of NPs. Using a mouse ES cell culture system, we also investigated the mechanism of *Sox2ot* regulation in neuroectodermal cells and found that it is active in the nucleus and interacts with the multifunctional transcription factor *YY1*. We demonstrated *YY1* binding to CpG islands in the *Sox2* locus and provided evidence that this binding is mediated by *Sox2ot* and provides direct *cis*-regulation of *Sox2* expression in NPs (Supplementary Figure S7). Our results characterize *Sox2ot* function and give insight into the mechanism of a specific lncRNA, while providing support for the general archetype of lncRNA function as a scaffold for protein recruitment.

In the developing cerebral cortex, the balance between progenitor pool maintenance and neurogenesis is a highly complex and tightly regulated process<sup>37</sup>. RGCs are the main neural progenitor pool, have the ability to self-renew, generate a secondary IP pool, or directly generate neurons. IPs also proliferate transiently to expand the progenitor pool before generating neurons. As neurons are generated, they migrate outward into the cortex to form layers in an inside-out manner, such that the earliest-born neurons lie close to the inner surface of the cortex, while later-born neurons migrate past to populate the outer layers<sup>38</sup>. Errors in this intricate process have severe implications for proper cortex formation and function<sup>39–41</sup>.

In this study, we demonstrated *Sox2ot* expression in NPs, and identified a novel role for *Sox2ot* in repression of NP expansion. Our result is distinct from studies of *Sox2ot* in cancers, where it is found to positively regulate *Sox2* expression<sup>42,43</sup>. Though the cancer studies did not explore the mechanism of *Sox2ot*, one possible explanation for the opposing effects on *Sox2* is that *Sox2* must be reactivated, often by amplification, in cancers, whereas *Sox2* is already highly expressed in NPs in the cortex<sup>44</sup>. Thus, the transcriptional and epigenetic landscape surrounding the *Sox2* locus may look very different and requires different regulation in cancer cells versus NPs.

lncRNAs have been shown to work through a variety of mechanisms, including as scaffolds for ribonucleoprotein complexes, activators of distinct transcriptional programs, and recruiters of proteins to specific loci<sup>45</sup>. Our results support *Sox2ot* action through epigenetic regulation by recruiting the binding partner *YY1* to the *Sox2* locus. *YY1* is a ubiquitously expressed, multifunctional polycomb group transcription factor, can be activating or repressive depending on the context, and often works by binding CpG islands and recruiting cofactors to perform epigenetic modifications<sup>46</sup>. Studies of *YY1* in early nervous system development have demonstrated its importance for neurulation and proper patterning, but its role in cortical development has not been characterized<sup>33,47</sup>. Here, we demonstrate a new prodifferentiative role of *YY1* in NPs, as its knockdown results in increased NPs and decreased neurons.

We propose that the function of the *Sox2ot*–*YY1* complex is to modulate *Sox2* expression in NPs, helping to balance progenitor pool maintenance with neurogenesis, such that low levels of *Sox2ot* allow high *Sox2* expression, thus promoting proliferation (Supplementary Figure S7A), while higher *Sox2ot* expression represses *Sox2*, resulting in differentiation of neurons (Supplementary Figure S7B). There are several proposed models of *YY1*-mediated repression, including passively blocking activation by covering DNA recognition sites or interfering with activation factor binding, or actively repressing transcription through co-repressor recruitment<sup>48</sup>. Mechanistically, we propose that low *Sox2ot* expression allows transcription from the *Sox2* locus (Supplementary Figure S7C), while higher expression of *Sox2ot* allows recruitment of *YY1*, likely with cofactors, to the CpG islands at the *Sox2* locus (Supplementary Figure S7D). Since *YY1* binding sites are pervasive throughout mammalian genomes, *Sox2ot* is potentially important for preferentially recruiting *YY1* to the *Sox2* locus through *cis*-regulation of *Sox2ot* to its overlapping gene *Sox2*<sup>49,50</sup>. Moreover, the *Sox2ot*–*YY1* complex perhaps also binds to promoters of other genes beside of *Sox2*. Identification of other *YY1* cofactors in the lncRNA–*YY1* system would

give further insight into the method of YY1-mediated repression occurring in brain development.

In recent years, lncRNAs as a group have been shown to be expressed in specific and regulated patterns in the brain, important for neural fate specification, and critical for the regulation of neurogenesis<sup>9,51</sup>. Specific lncRNAs have also demonstrated functions in neural development, including *RMST*, which is required for Sox2 binding to neurogenic transcription factors in neural stem cells, thus promoting neurogenesis, and *TUNA*, which is required for pluripotency and neural lineage commitment of mouse ES cells<sup>52,53</sup>. In the embryonic mouse cortex, the lncRNA *Pinky* interacts with RNA-splicing protein PTBP1 to regulate neurogenesis<sup>54</sup>. Finally, *Paupar* induces neural differentiation through both Pax6-dependent and Pax6-independent processes<sup>35</sup>.

In this study, we have characterized the expression of *Sox2ot* in NPs in the developing mouse cortex and established its role in repression of NP expansion. We also demonstrated a novel interaction of *Sox2ot* with the multifunctional transcription factor YY1 and showed that YY1 binds to CpG islands at the *Sox2* locus in neuroectodermal cells. Through this work, we have provided a further example of one lncRNA important in neural development, gained insight into the modes of NP regulation, and expanded the understanding of lncRNA regulatory mechanisms.

## Materials and methods

### In situ hybridization

In situ hybridization was performed according to previously published methods<sup>55</sup>. Briefly, brain sections were hybridized with DIG-labeled RNA probes at 65 °C overnight. After washing with preheated wash solution (1×SSC, 50% formamide, 0.1% Tween-20) and MABT, sections were blocked with blocking buffer (1×MABT, 2% blocking reagent, 20% heat-inactivated sheep serum) and incubated with anti-DIG antibody (1:1,500, Roche) at 4 °C overnight. Brain sections were washed with 1×MABT and staining buffer (0.1 M NaCl, 50 mM MgCl<sub>2</sub>, and 0.1 M Tris-HCl, pH 9.5), and stained with BM Purple (Roche) at room temperature until ideal intensity. The probes were labeled with DIG-ddUTP using the DIG RNA labeling kit (Roche). The images of in situ hybridization were collected using a Leica digital camera under a dissection microscope (Leica, MZ16F).

### In utero electroporation

In utero electroporation was performed as described previously<sup>56</sup>. Briefly, electroporation was conducted at E13.5, and the brain tissues were collected either 24 h later at E14.5 or 96 h later at E17.5. Bromodeoxyuridine (BrdU, 50 µg/g body weight) was administered intraperitoneally as a single pulse 23 h after electroporation, 1 h

before sacrifice. Plasmid DNA was prepared using the EndoFree Plasmid Maxi Kit (Qiagen) and diluted to 2 µg/µL. DNA solution was injected into the lateral ventricle of the cerebral cortex and electroporated with five 50-ms pulses at 35 V using an ECM830 electro-squarator (BTX).

### Tissue preparation and immunohistochemistry

Mouse brains were fixed in 4% paraformaldehyde (PFA) in phosphate-buffered saline (PBS) overnight, incubated in 30% sucrose in PBS, embedded in OCT, and stored at -80 °C until use. Brains were sectioned (14 µm) using a cryostat. For antigen recovery, sections were incubated in heated (95–100 °C) antigen recovery solution (1 mM EDTA and 5 mM Tris, pH 8.0) for 20 min and cooled for 30 min. Before applying antibodies, sections were blocked in normal goat serum (NGS) in PBS with 0.1% Tween-20 for 1 h at room temperature. Sections were incubated with primary antibodies overnight at 4 °C and visualized using goat anti-rabbit IgG-Alexa-Fluor-488 or goat anti-chicken IgG-Alexa-Fluor-488 and goat anti-rabbit IgG Alexa-Fluor-546 or goat anti-mouse IgG-Alexa-Fluor-546 (1:1,000, Molecular Probes) for 1 h at room temperature. For sections stained with Sox2 primary antibody, normal donkey serum was substituted for NGS and donkey anti-rabbit IgG-Alexa-Fluor-488 and donkey anti-goat IgG-Alexa-Fluor-546 were used for visualization (1:1000, Molecular Probes).

Primary antibodies against the following antigens were used: Sox2 (1:200, Santa Cruz), Pax6 (1:500, Covance), Tbr2 (1:500, Abcam), BrdU (1:50, DSHB), Caspase3 (1:1000, R&D Systems), Tbr1 (1:500, Abcam), Satb2 (1:500, Abcam), GFP (1:1000, Rockland, rabbit), GFP (1:1000, Abcam, chicken), and Yy1 (1:200, Santa Cruz).

Images were captured using a Zeiss confocal microscope.

### Quantitation of immunostained tissue

Coronal sections were collected in the medial cortical region. At least four sections from each brain and three electroporated brains from the same litter were chosen for antibody labeling. For neural progenitor markers (Sox2, BrdU, Pax6, and Tbr2), positive cells were quantified in fixed areas of 100 × 100 µm<sup>2</sup>, starting at the ventricular surface. For mature neuronal markers (Tbr1 and Satb2), cells were quantified in columns with a width of 200 µm and height from the lowest point of marker staining to the pial surface.

Cell counting in the mouse brain sections was performed on a fixed width (200-µm bin) of a representative column in the cortical wall. All sections analyzed were selected from a similar medial point on the anterior–posterior axis. Cell counting was performed in minimal three chosen areas in each brain, and at least

three brains were analyzed in each group. Cell counting in each chosen area was repeated at least three times and a mean was obtained. All data are presented as mean±SEM. *P* values were calculated using an unpaired Student's *t* test.

### Embryonic stem cell culture

Stem cell culture and JQ1 treatment were performed, as described previously<sup>29</sup>. Briefly, R1 and LF2 ES cells were maintained for 4 days in DME and 15% FCS. Cells were treated for 24 h with JQ1 (BPS Biosciences) diluted to 2 μM in DMSO.

### RNA immunoprecipitation

RNA immunoprecipitation was performed following the Abcam RIP protocol. Briefly, cells were cross-linked with a final concentration of 0.75% formaldehyde, harvested by trypsinization, and resuspended in PBS. The nuclei were pelleted by centrifugation and resuspended in RIP buffer (150 mM KCl, 25 mM Tris, pH 7.4, 5 mM EDTA, 0.5 mM DTT, 0.5% NP40, and 100 U/mL RNase inhibitor). Chromatin was sheared by sonication, and nuclear membranes were pelleted by centrifugation. For each IP, 4 μg of antibody (YY1, Santa Cruz; CTCF, Santa Cruz; Pax6, Covance) was added, and the samples were incubated overnight at 4 °C with rotation. Protein A/G PLUS-agarose beads (Santa Cruz Biotechnology) were added, and the samples were incubated at 4 °C for 2 h with rotation. The beads were pelleted by centrifugation, resuspended, and washed in RIP buffer. RNA was eluted in RIP elution buffer (50 mM Tris, pH 8, 10 mM EDTA, and 1% SDS) supplemented with proteinase K and RNase inhibitor overnight at 65 °C. Samples were treated with DNase and precipitated overnight.

### Statistics

For immunostaining, at least four sections from each brain and at least three different brains from the same litter were chosen for antibody labeling and quantification. For qRT-PCR, triplicated samples were tested. Statistical comparisons were made by analysis of variance (unpaired Student's *t* test).

### Acknowledgements

We thank Dr. Haijun Zhang for assistance in experimental troubleshooting. This work was supported by the NIH Institutional T32 Training Grant HD 060600 (J.L.K.), an 1R01MH090267 (M.E.D.), the Burke Foundation (M.E.D.), the Hirschl/Weill-Caulier Trust (T.S.), an R01-MH083680 grant from the NIH/NIMH (T.S.), China Postdoctoral Science Foundation (2017M622053), and the National Natural Science Foundation of China (81471152, 31771141, and 81701132).

### Author details

<sup>1</sup>Department of Cell and Developmental Biology, Cornell University Weill Medical College, 1300 York Avenue, Box 60, New York, NY 10065, USA. <sup>2</sup>Center for Precision Medicine, School of Medicine and School of Biomedical Sciences, Huaqiao University, Xiamen, China 361021. <sup>3</sup>College of Korean Medicine, Dongguk University, Ilsandonggu, Goyangsi 10326 Gyeonggi-do, Korea. <sup>4</sup>School

of Life Sciences and Technology, Shanghai Jiao Tong University, Shanghai, China 200240. <sup>5</sup>Burke Medical Research Institute, 785 Mamaroneck Avenue, White Plains, NY 10605, USA. <sup>6</sup>Department of Neuroscience, Department of Cell and Developmental Biology, Cornell University Weill Medical College, 1300 York Avenue, New York, NY 10065, USA

### Author contributions

J.L.K. and T.S. conceived the study and designed the experiments; J.L.K. performed all immunohistochemistry; J.L.K., N.M., H.B.P., and M.E.D. designed and performed fractionation, RIP and ChIP experiments in mouse ES cells, and ddPCR and FISH experiments; S.K. performed in utero electroporations; Y.S. performed bioinformatic analyses; Y.N. extracted RNA samples and cell cultures; T.W. and M.E.D. designed and validated shRNA constructs against YY1; J.L.K. and T.S. wrote the paper; and T.S. supervised the project.

### Conflict of interest

The authors declare that they have no conflict of interest.

### Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Supplementary Information** accompanies this paper at (<https://doi.org/10.1038/s41419-018-0840-2>).

Received: 6 February 2018 Revised: 4 June 2018 Accepted: 18 June 2018  
Published online: 23 July 2018

### References

- Leone, D. P. et al. *Satb2* regulates the differentiation of both callosal and subcortical projection neurons in the developing cerebral cortex. *Cereb. Cortex* **25**, 3406–3419 (2015).
- McKenna, W. L. et al. Mutual regulation between *Satb2* and *Fezf2* promotes subcortical projection neuron identity in the developing cerebral cortex. *Proc. Natl. Acad. Sci. USA* **112**, 11702–11707 (2015).
- Alcamo, E. A. et al. *Satb2* regulates callosal projection neuron identity in the developing cerebral cortex. *Neuron* **57**, 364–377 (2008).
- Zhang, L. et al. *Satb2* is required for dendritic arborization and soma spacing in mouse cerebral cortex. *Cereb. Cortex* **22**, 1510–1519 (2012).
- Huang, Y. et al. Expression of transcription factor *Satb2* in adult mouse brain. *Anat. Rec.* **296**, 452–461 (2013).
- Derrien, T. et al. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res.* **22**, 1775–1789 (2012).
- Djebali, S. et al. Landscape of transcription in human cells. *Nature* **489**, 101–108 (2012).
- Caminci, P. et al. The transcriptional landscape of the mammalian genome. *Science* **309**, 1559–1563 (2005).
- Mercer, T. R., Dinger, M. E. & Mattick, J. S. Long non-coding RNAs: insights into functions. *Nat. Rev. Genet.* **10**, 155–159 (2009).
- Wang, K. C. & Chang, H. Y. Molecular mechanisms of long noncoding RNAs. *Mol. Cell* **43**, 904–914 (2011).
- Wapinski, O. & Chang, H. Y. Long noncoding RNAs and human disease. *Trends Cell Biol.* **21**, 354–361 (2011).
- Niland, C. N., Mery, C. R. & Khalil, A. M. Emerging roles for long non-coding RNAs in cancer and neurological disorders. *Front. Genet.* **3**, 25 (2012).
- Gibb, E. A., Brown, C. J. & Lam, W. L. The functional role of long non-coding RNA in human carcinomas. *Mol. Cancer* **10**, 38 (2011).
- Fong, H., Hohenstein, K. A. & Donovan, P. J. Regulation of self-renewal and pluripotency by *Sox2* in human embryonic stem cells. *Stem Cells* **26**, 1931–1938 (2008).
- Avilion, A. A. et al. Multipotent cell lineages in early mouse development depend on *SOX2* function. *Genes Dev.* **17**, 126–140 (2003).
- Zappone, M. V. et al. *Sox2* regulatory sequences direct expression of a (beta)-geo transgene to telencephalic neural stem cells and precursors of the mouse embryo, revealing regionalization of gene expression in CNS stem cells. *Development* **127**, 2367–2382 (2000).

17. Deaton, A. M. & Bird, A. CpG islands and the regulation of transcription. *Genes Dev.* **25**, 1010–1022 (2011).
18. Mikkelsen, T. S. et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**, 553–560 (2007).
19. Hutton, S. R. & Pevny, L. H. SOX2 expression levels distinguish between neural progenitor populations of the developing dorsal telencephalon. *Dev. Biol.* **352**, 40–47 (2011).
20. Miyagi, S. et al. Consequence of the loss of Sox2 in the developing brain of the mouse. *FEBS Lett.* **582**, 2811–2815 (2008).
21. Miyagi, S. et al. The Sox2 regulatory region 2 functions as a neural stem cell-specific enhancer in the telencephalon. *J. Biol. Chem.* **281**, 13374–13381 (2006).
22. Tomioka, M. et al. Identification of Sox-2 regulatory region which is under the control of Oct-3/4-Sox-2 complex. *Nucleic Acids Res.* **30**, 3202–3213 (2002).
23. Julian, L. M. et al. Opposing regulation of Sox2 by cell-cycle effectors E2f3a and E2f3b in neural stem cells. *Cell Stem Cell* **12**, 440–452 (2013).
24. Marques-Torres, M. A. et al. Cyclin-dependent kinase inhibitor p21 controls adult neural stem cell expansion by regulating Sox2 gene expression. *Cell Stem Cell* **12**, 88–100 (2013).
25. Amaral, P. P. et al. Complex architecture and regulated expression of the Sox2ot locus during vertebrate development. *RNA* **15**, 2013–2027 (2009).
26. Fantès, J. et al. Mutations in SOX2 cause anophthalmia. *Nat. Genet.* **33**, 461–463 (2003).
27. Williamson, K. A. et al. Mutations in SOX2 cause anophthalmia-esophageal-genital (AEG) syndrome. *Hum. Mol. Genet.* **15**, 1413–1422 (2006).
28. Shahryari, A. et al. Two novel splice variants of SOX2OT, SOX2OT-S1, and SOX2OT-S2 are coexpressed with SOX2 and OCT4 in esophageal squamous cell carcinoma. *Stem Cells* **32**, 126–134 (2014).
29. Wu, T., Pinto, H. B., Kamikawa, Y. F. & Donohoe, M. E. The BET family member BRD4 interacts with OCT4 and regulates pluripotency gene expression. *Stem Cell Rep.* **4**, 390–403 (2015).
30. Kornienko, A. E., Guenzl, P. M., Barlow, D. P. & Pauler, F. M. Gene regulation by the act of long non-coding RNA transcription. *BMC Biol.* **11**, 59 (2013).
31. Rinn, J. L. & Chang, H. Y. Genome regulation by long noncoding RNAs. *Annu. Rev. Biochem.* **81**, 145–166 (2012).
32. Jeon, Y. & Lee, J. T. YY1 tethers Xist RNA to the inactive X nucleation center. *Cell* **146**, 119–133 (2011).
33. Donohoe, M. E. et al. Targeted disruption of mouse Yin Yang 1 transcription factor results in peri-implantation lethality. *Mol. Cell Biol.* **19**, 7237–7244 (1999).
34. Zhang, H. et al. Long noncoding RNA-mediated intrachromosomal interactions promote imprinting at the Kcnq1 locus. *J. Cell Biol.* **204**, 61–75 (2014).
35. Vance, K. W. et al. The long non-coding RNA Paupar regulates the expression of both local and distal genes. *EMBO J.* **33**, 296–311 (2014).
36. Guil, S. & Esteller, M. Cis-acting noncoding RNAs: friends and foes. *Nat. Struct. Mol. Biol.* **19**, 1068–1075 (2012).
37. Noctor, S. C., Martinez-Cerdeno, V., Ivic, L. & Kriegstein, A. R. Cortical neurons arise in symmetric and asymmetric division zones and migrate through specific phases. *Nat. Neurosci.* **7**, 136–144 (2004).
38. Götz, M. & Huttner, W. B. The cell biology of neurogenesis. *Nat. Rev. Mol. Cell Biol.* **6**, 777–788 (2005).
39. Rakic, P. Evolution of the neocortex: a perspective from developmental biology. *Nat. Rev. Neurosci.* **10**, 724–735 (2009).
40. Sun, T. & Hevner, R. F. Growth and folding of the mammalian cerebral cortex: from molecules to malformations. *Nat. Rev. Neurosci.* **15**, 217–232 (2014).
41. Rubenstein, J. L. Annual research review: development of the cerebral cortex: implications for neurodevelopmental disorders. *J. Child Psychol. Psychiatry* **52**, 339–355 (2011).
42. Askarian-Amiri, M. E. et al. Emerging role of long non-coding RNA SOX2OT in SOX2 regulation in breast cancer. *PLoS ONE* **9**, e102140 (2014).
43. Hou, Z. et al. A long noncoding RNA Sox2ot regulates lung cancer cell proliferation and is a prognostic indicator of poor survival. *Int. J. Biochem. Cell Biol.* **53**, 380–388 (2014).
44. Hussenet, T. et al. SOX2 is an oncogene activated by recurrent 3q26.3 amplifications in human lung squamous cell carcinomas. *PLoS ONE* **5**, e8960 (2010).
45. Moran, V. A., Perera, R. J. & Khalil, A. M. Emerging functional and mechanistic paradigms of mammalian long non-coding RNAs. *Nucleic Acids Res.* **40**, 6391–6400 (2012).
46. Zhang, Q., Stovall, D. B., Inoue, K. & Sui, G. The oncogenic role of Yin Yang 1. *Crit. Rev. Oncog.* **16**, 163–197 (2011).
47. Morgan, M. J., Woltering, J. M., In der Rieden, P. M., Durston, A. J. & Thiery, J. P. YY1 regulates the neural crest-associated slug gene in *Xenopus laevis*. *J. Biol. Chem.* **279**, 46826–46834 (2004).
48. Thomas, M. J. & Seto, E. Unlocking the mechanisms of transcription factor YY1: are chromatin modifying enzymes the key? *Gene* **236**, 197–208 (1999).
49. Yant, S. R. et al. High affinity YY1 binding motifs: identification of two core types (ACAT and CCAT) and distribution of potential binding sites within the human beta globin cluster. *Nucleic Acids Res.* **23**, 4353–4362 (1995).
50. Becker, K. G., Jedlicka, P., Templeton, N. S., Liotta, L. & Ozato, K. Characterization of hUCRBP (YY1, NF-E1, delta): a transcription factor that binds the regulatory regions of many viral and cellular genes. *Gene* **150**, 259–266 (1994).
51. Knauss, J. L. & Sun, T. Regulatory mechanisms of long noncoding RNAs in vertebrate central nervous system development and function. *Neuroscience* **235**, 200–214 (2013).
52. Ng, S. Y., Bogu, G. K., Soh, B. S. & Stanton, L. W. The long noncoding RNA RMST interacts with SOX2 to regulate neurogenesis. *Mol. Cell* **51**, 349–359 (2013).
53. Lin, N. et al. An evolutionarily conserved long noncoding RNA TUNA controls pluripotency and neural lineage commitment. *Mol. Cell* **53**, 1005–1019 (2014).
54. Ramos, A. D. et al. The long noncoding RNA Pnky regulates neuronal differentiation of embryonic and postnatal neural stem cells. *Cell Stem Cell* **16**, 439–447 (2015).
55. Bian, S. et al. MicroRNA cluster miR-17-92 regulates neural stem cell expansion and transition to intermediate progenitors in the developing mouse neocortex. *Cell Rep.* **3**, 1398–1406 (2013).
56. Saito, T. In vivo electroporation in the embryonic mouse central nervous system. *Nat. Protoc.* **1**, 1552–1558 (2006).