Characterization of Enhancers Active in the Mouse Embryonic Cerebral Cortex Suggests Sox/Pou cis-Regulatory Logics and Heterogeneity of Cortical Progenitors

Amandine Bery, Ben Martynoga, François Guillemot, Jean-Stéphane Joly, Sylvie Rétaux

To cite this version:
Amandine Bery, Ben Martynoga, François Guillemot, Jean-Stéphane Joly, Sylvie Rétaux. Characterization of Enhancers Active in the Mouse Embryonic Cerebral Cortex Suggests Sox/Pou cis-Regulatory Logics and Heterogeneity of Cortical Progenitors. Cerebral Cortex, Oxford University Press (OUP), 2014, 24 (11), pp.2822-2834. 10.1093/cercor/bht126. hal-00850761

HAL Id: hal-00850761
https://hal.archives-ouvertes.fr/hal-00850761
Submitted on 22 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Characterization of Enhancers Active in the Mouse Embryonic Cerebral Cortex Suggests Sox/Pou cis-Regulatory Logics and Heterogeneity of Cortical Progenitors

Amandine Berry¹, Ben Martynoga³, François Guillemot³, Jean-Stéphane Joly² and Sylvie Rétaux¹

¹Equipe Development and Evolution of the Forebrain, ²Equipe Morphogenesis of the Chordate Nervous System, UPR3294 N&D, Institut de Neurobiologie Alfred Fessard, CNRS, Gif-sur-Yvette, France and ³MRC-NIMR, London, UK

Address correspondence to S. Rétaux, UPR 3294-NeD, Institut de Neurobiologie Alfred Fessard, CNRS, Avenue de la Terrasse, 91198 Gif-sur-Yvette cedex, France. Email: retaux@inaf.cnrs-gif.fr

We aimed to identify cis-regulatory elements that control gene expression in progenitors of the cerebral cortex. A list of 975 putative enhancers were retrieved from a ChiP-Seq experiment performed in NS5 mouse stem cells with antibodies to Sox2, Brm2/Pou3f2, or Brm1/Pou3f3. Through a selection pipeline including gene ontology and expression pattern, we reduced the number of candidate enhancer sequences to 20. Ex vivo electroporation of green fluorescent pProtein (GFP) reporter constructs in the telencephalon of mouse embryos showed that 35% of the 20 selected candidate sequences displayed enhancer activity in the developing cortex at E13.5. In silico transcription factor binding site (TFBS) searches and mutagenesis experiments showed that enhancer activity is related to the presence of Sox/Pou TFBS pairs in the sequence. Comparative genomic analyses showed that enhancer activity is not related to the evolutionary conservation of the sequence. Finally, the combination of in utero electroporation of GFP reporter constructs with immunostaining for Tbr2 (basal progenitor marker) and phospho-histoneH3 (mitotic activity marker) demonstrated that each enhancer is specifically active in precise subpopulations of progenitors in the cortical germinal zone, highlighting the heterogeneity of these progenitors in terms of cis-regulation.

Keywords: cis-regulation, ChiP-Seq, electroporation, Pou, Sox

Introduction

The mammalian cortex is a finely organized structure that is thought to have facilitated the emergence of higher cognitive functions during evolution. The development of such a structure depends on tight regulation of the properties and the behavior of cortical progenitors during embryogenesis, through the activation of gene regulatory networks and the combinatorial activity of multiple transcription factors (TFs). Such factors involved in the expansion and differentiation of neural progenitors have been identified. But how they act in space and time is poorly understood because the distant cis-regulatory sequences that they bind to regulate gene expression remain mostly unknown. In this study, we set out to identify cis-regulatory elements that control gene expression, hence cellular behavior, of cortical progenitors.

During corticogenesis, neuroepithelial cells initially undergo exclusively proliferative, symmetric divisions, expanding the progenitor pool. They then switch to neurogenic divisions that give rise to the earliest born cortical neurons and to other progenitor populations distributing in layered subdivisions of the neuroepithelium (Boulder committee 1970; Bystron et al. 2008). Classically, 2 types of progenitors are recognized: (1) The apical progenitors (APs) including radial glial cells and short neuronal progenitors residing in the ventricular zone (VZ), and (2) the basal progenitors (BPs) that originate as daughters of APs and are located in the subventricular zone (SVZ; Malatesta et al. 2000; Miyata et al. 2001; Noctor et al. 2001, 2004; Haubensak et al. 2004; Gal et al. 2006; Stancik et al. 2010; reviewed in Mérot et al. 2009). APs and BPs are distinguished on the basis of molecular markers such as Pax6 or Tbr2, respectively (Hartfuss et al. 2001; Tarabykin et al. 2001; Heins et al. 2002; Nieto et al. 2004; Zimmer et al. 2004; Englund et al. 2005; Gal et al. 2006). APs and BPs also differ in their mitotic positions and patterns (reviewed in Sidman and Rakic 1973; Noctor et al. 2008; Taverna and Huttner 2010). Of note, based on the varied behaviors and molecular properties exhibited by cells in the germinal zone, some authors have proposed that cortical progenitors are more heterogeneous than initially thought, even in lissencephalic or nonprimate cortices (Gal et al. 2006; Kowalczyk et al. 2009; Stancik et al. 2010; Shitamukai et al. 2011; Wang et al. 2011; Kelava et al. 2012; Reillo and Borrell 2012).

Transcription factors expressed in cortical progenitors critically influence their behavior. Among them, Sox2 is involved in the acquisition and maintenance of neuronal progenitor identity (Bylund et al. 2003; Graham et al. 2003; Ellis et al. 2004). Together with Class III POU (Brn) TFs, Sox2 binds the Nestin neural enhancer and regulates Nestin expression (Tanaka et al. 2004; Jin et al. 2009) in a cell cycle-dependent manner, thereby orchestrating gene regulation that correlates with the 3-dimensional morphological changes in neural progenitors in vivo (Sunabori et al. 2008). Another example is given by Pax6, which expression level controls the balance between neural stem cell self-renewal and neurogenesis (Sansom et al. 2009). Thus, these TFs are major effectors steering the dynamic progression of regulatory states during the generation of cell types.

Only few enhancers, active in neural stem cells, have been identified. The Sox2 enhancer, SRR2, is active specifically in neural stem/progenitor cells and recruits Sox and POU TFs in telencephalic progenitors (Miyagi et al. 2006). More recently, Visel et al. (2013) have released an atlas of 329 enhancers active in the developing mouse telencephalon, some being active in the cortical germinal zone. From this dataset, they deduced cis-regulatory logics that confers regional identity in the telencephalic neuroepithelium in its tangential dimension (e.g., pallium vs. subpallium). Here, and in a similar manner, we sought to describe cis-regulatory logics underlying the behavior and possible heterogeneity of a radial compartment of the developing cortex: The progenitors of the germinal zone.

To address this question, we used ChiP-seq followed by functional assay in the embryonic mouse cortex to identify enhancers on which 3 specific cortical progenitor-expressed TFs bind. We report the identification and functional characterization of 7 sequences that behave as active enhancers in distinct categories of cortical progenitors in vivo.
Materials and Methods

Chip-Seq
NS5 cells were kindly provided by Austin Smith and cultured under standard conditions (Conti et al. 2005). Cells were dual fixed with di-usscinimidyl glutarate (2 nM) and 1% formaldehyde and then lysed and sonicated as previously described (Castro et al. 2011). Chromatin immunoprecipitation was performed according to the standard Uptase/Millipore protocol using specific antibodies to Sox2 (Santa Cruz sc-17320), Brm2/Pou3f2 (Santa Cruz sc-6029), or Brm1/Pou3f3 (kindly provided by Dr Dies Meijer). Immunoprecipitated material was sequenced according to Illumina’s standard protocols. Reads were mapped with Bowtie to the mm9 genome assembly and significant peaks called with MACs v1.3 (Zhang et al. 2008). We defined 2892 regions significantly bound by Sox2 compared with input chromatin and 7176 regions bound by Brm1 and/or Brm2 (these were pooled due to predicted functional redundancy of these factors). Intersection of these lists of ChIP-seq peak coordinates revealed 925 genomic regions bound by Sox2 and Brm1 and/or Brm2. Of note, both Sox2 and Brm TFs are expressed homogeneously in NS5 cells, in a manner that is analogous to their expression throughout the cortical VZ in vivo (immunohistochemistry, data not shown). These data will be described more thoroughly in a manuscript currently in preparation in the Guillemot lab.

In Silico Analyses of Candidate Enhancer Sequences

Evolutionary Conservation
The conservation of the potential candidate enhancers was studied through a multiple alignment program for comparative genomics, MLAGAN (Brudno et al. 2003), and the results were visualized via the VISTA server (Mayor et al. 2000). The different vertebrate species used in the alignment were mouse (Mus musculus, baseline), human (Homo sapiens), dog (Canis familiaris), platypus (Ornithorhynchus anatinus), gibbon (Nomascus leucogenys), chicken (Gallus gallus), xenopus (Xenopus laevis), medaka (Oryzias latipes), stickleback (Gasterosteus aculeatus), and fugu (Takifugu rubripes).

TFBS Search
p500 binding sites were retrieved from Visel et al. (2009) and compared with chromosomal positions of candidate enhancer sequences. Sox and Pou TF binding motifs were searched on the sequences as follows: [A/G][CACTA/T]A and [CA/TTGTGTA/JG/T/A] for Sox (Harley et al. 1994; Tomioka et al. 2002); ATTA/TGCGAT for POU (Schöler et al. 1989; Petryniak et al. 1990; Tomioka et al. 2002; Castro et al. 2006; Cook and Sturm 2008; Ryan and Rosenfeld 2011). Each sequence was scanned with both the forward and the reverse sequences of each motif, the distribution was visualized using the Regulatory Sequence Analysis Tool (van Helden 2003), and frequencies of occurrence were calculated.

PWM Calculation
The position weight matrix (PWM) for Sox and Pou transcription factor binding sites (TFBSs) present in the 7 active enhancers was calculated from the alignment of the different consensus motifs found in the enhancers (http://genome.crg.es/courses/cshl2007/T2/MakeProfile.html). Then, both PWM were visualized and identified through the JASPAR database (Wasserman and Sandelin 2004).

Cloning of Target Enhancer Regions and Vector Constructs
Candidate enhancer sequences were amplified as attB3/attB4 PCR fragment using proper primers, Accuprime Taq (Invitrogen) and mouse genomic DNA as template. Each amplified sequence was inserted into a plasmid to get a library of putative enhancers. This library was generated by Gateway cloning using a 3–4 donor vector and a customized destination vector (p34globGFP34HR34), where green fluorescent protein (GFP) reporter expression is under control of a β-globin minimal promoter. DNA preps for electroporation were endotoxin-free (Qiagen).

For mutagenesis experiments, the TFBSs of Pou (del1) or Sox (del2), or both (del3), were deleted by using the GENEART site-directed mutagenesis system and specific primers (sequences available in Supplementary Table 1).

Electroporations
Time-pregnant Swiss mice were used, with the midpoint of the dark period as embryonic day 0 (E0). Animals were treated according to the regulations and laws of the European Union (86/609/EEC) for care and handling of animals in research, using procedure ref10/version1 approved by the Ethics Committee CEE-Part1. SR’s authorization for animal experimentation is 91-329.

Ex Vivo Electroporation and Cortical Slice Culture
E13.5 embryos were obtained by cesarian opening after cervical dislocation of the pregnant mother. One to 1.5 µl of plasmid DNA (1 µg/µl) mixed with Fast green was injected into the lateral ventricle, using a beveled and calibrated glass micropipette (pulled on Harvard apparatus) and a microinjector (Femtojet, Eppendorf). For electroporation, 2 × 35 ms pulses of 35 mV with a 1-ms interval were delivered with two 5-mm electrode paddles positioned on either side of the head (BTX, ECM830 Harvard apparatus). Throughout procedure, the embryos were bathed in an ice-cold 1× Krebs solution (for 10×: 1.26 mM NaCl, 25 mM KCl, 12 mM NaH2PO4, 12 mM MgCl2, 25 mM CaCl2) plus 1 g glucose and 1.05 g NaHCO3.

Immediately after ex vivo electroporation, the brains of embryos were dissected out in ice-cold 1× Krebs solution, embedded in 2% low-melting agarose, and sectioned at 250 µm using a vibratome (VT1000S Leica Microsystems). Brain slices were transferred onto a slice culture insert (Millipore) in a plastic tissue culture dish with culture medium containing: 47.5 mM Neurobasal (Invitrogen/Gibco), 1 mL B27 (Invitrogen/Gibco), 0.5 mL 50% glucose, and 0.5 mL 100× glutamine. Cultures were maintained in a humidified incubator at 37°C with 5%CO2 supply. Pictures were taken with an Olympus binocular microscope.

In Utero Electroporation
In utero electroporation was performed as described by Shimogori and Ogawa (2008) including few modifications. A time-pregnant Swiss mouse was anesthetized with a solution of Ketamine (Merial) and Xylazine (Bayer). The uterine horns containing E13.5 embryos were exposed. One to 1.5 µl of plasmid DNA (1 µg/µl) mixed with Fast green was manually microinjected through the uterine wall into the lateral ventricle, using a beveled and calibrated glass micropipette. Thirty-five-millisecond pulses of 35 mV with a 1-ms interval were delivered across the uterus with two 5-mm electrodes paddles positioned on either side of the head (BTX, ECM830 Harvard apparatus). Throughout the surgical procedure, the animal was placed on a heater block at 37°C and the uterus was bathed with warm phosphate buffer saline (PBS) (pH 7.4). After the procedure, the uterus was placed back in the abdominal cavity and the wound was surgically sutured. Twenty hours after electroporation the embryonic brains (E14.5) were dissected and fixed in 4% paraformaldehyde in PBS (pH 7.4) for subsequent analyses.

Immunohistochemistry
Fixed E14.5 brains were embedded in 2% agarose (made in MilliQ H2O) and sectioned at 80 µm using a vibratome (VT1000S Leica Microsystems). The sections were washed several times in PBS/0.5%Triton X-100/0.5% bovine serum albumine. Incubation with primary and secondary antibodies was performed at 4°C overnight. The primary antibodies utilized were as follows: Rabbit antiphospho-histone H3 (pH3; 1:500, Upstate Biotechnology), rabbit anti-TPR2 (1:500, Abcam), and mouse anti-GFP monoclonal antibody (1:500, Roche). Before incubation with rabbit anti-Thr2, the sections were blocked in PBS/2% Tween/2% Triton X-100 with 10% of blocking reagent (Roche) and 10% of normal goat serum (Sigma-Aldrich). Secondary antibodies were goat antirabbit IgG (H + L) Alexa 594 and goat anti-mouse IgG1 Alexa 488 (Molecular Probes). Finally, sections were washed and mounted in
ProLong Gold antifade reagent with 4′,6-diamidino-2-phenylindole (Invitrogen). Photographs were taken with an ApoTome (Zeiss).

Probes and In Situ Hybridization
cDNA probes for the 7 active enhancers were cloned by RT-PCR from E13.5 brain total RNA. The resulting DNA fragments were inserted in pCR™-TOPO (Invitrogen). After linearization, digoxigenin-labeled antisense RNA probes were generated by T3 or T7 RNA polymerase, and in situ hybridization (ISH) was performed as previously described (Abellan et al. 2010).

Statistics
The statistical analyses were performed on Statview4.57 (Abacus Concept, Berkeley, CA, USA). The nonparametric Mann–Whitney test was used to compare the 2 populations of enhancers (GFP+/active and GFP−/inactive in cortical progenitors).

Results
Selection of Potential Candidate Enhancers Active in Cortical Progenitors From ChIP-Seq Data
A ChIP-seq experiment on mouse NS5 cells with antibodies specific for Sox2, Brn1, or Brn2 TFs provided a list of 925 potential candidate enhancer sequences, which bind these TFs in this neural stem cell line (Fig. 1). To shorten this list and select sequences to be tested for enhancer activity in the cortical germinal zone, 2 filters were applied. We reasoned that the closest coding gene to the candidate enhancer (found using the NCBI database) was most likely to be the one regulated by that enhancer. We therefore selected sequences that closest genes were:

1. Annotated with a gene ontology (GO) that was related to nervous system development and/or proliferation control. Selected GO terms for flanking genes were: Nervous system, forebrain, neurons, synapses, and axon guidance (n = 36); cell proliferation, differentiation, and apoptosis (n = 35); cell differentiation and nervous system development (n = 30); cell proliferation and nervous system development (n = 9); cell cycle (n = 8); apoptosis of neurons only (n = 2); and cell cycle and nervous system development (n = 1). This first “GO” filter shortened down the list to 120 candidate enhancer sequences.

2. Expressed in the developing cortex, according to the following databases: GenePaint, Genesat, and Allen Brain Atlas. This filter further reduced the number of potential candidate enhancer sequences to 20, which are listed in Table 1.

Features and Evolutionary Conservation of the 20 Candidate Enhancer Sequences
Although the selected 20 ChIP-seq-generated sequences were relatively evenly distributed around the closest neighboring genes, a slight bias toward an intronic localization was observed. Eight were intronic, 6 in the 5′ region, and 6 in the 3′ region (Table 1). Consistent with potential enhancers, selected sequences were located within 1–200 kb of the nearest gene. For those that are not intronic, their distance to the first exon of closest gene ranges from 2189 to 160 948 bp, with an average of 52 810 bp. The sequences sizes ranged from 232 to 1074 bp, with an average of 504 bp.

The evolutionary conservation of the 20 mouse candidate enhancer sequences was tested on genomic alignments from 10 vertebrate species, including mammals, birds, amphibians, and fishes, with the mouse genomic sequence as baseline (Fig. 2 and ninth column of Table 1). Only 1 of the 20 sequences, enh18(Pou2f1), showed conservation among all species including fish (Fig. 2A). The other sequences were either conserved between mammals only (such as enh20(Sox2), Fig. 2B) or else conserved between tetrapods only (such as enh06(Fubp3); Fig. 2C). Other, specific, conservation patterns were also found, as well as no conservation at all (ex: enh07(Fzd9) or enh12(Sox11); Table 1).

Most of these 20 putative enhancer sequences are novel and not characterized. Indeed, a search in the online database “VISTA Enhancer Browser” (Visel et al. 2007) retrieved none of the sequences listed in Table 1. Although this database contains enhancers for about half of the flanking genes presented in Table 1 (Big1, Cdh4, Gli3, Nr2f1, Sox11, Tie3, Sall3, Pou2f1, and Sox2), none of them corresponds to sequences found in our study. Only 2 of our 20 potential enhancer sequences have already been reported as confirmed active enhancers: enh19(Nes) corresponds to the nestin enhancer called Nes30 (Tanaka et al. 2004); and enh20(Sox2) is the Sox2 enhancer called SRR2 (Miyagi et al. 2006), and both are described as active in cortical progenitors (Table 1).

Seven of 20 Candidate Enhancers Are Active in the Developing Cortex
To test the activity of the 18 novel putative enhancer sequences in cortical progenitors in vivo, these sequences were amplified from genomic DNA to generate a library of enhancers and were subcloned into an enhancer-reporter vector containing a minimal β-globin promoter (Fig. 1).

Figure 1. Pipeline for the selection, analysis, and testing of enhancers. The strategy used to narrow down the list of 925 putative enhancer sequences to a “testable” number of 20 is summarized, as well as the types of functional analyses performed to test and characterize these 20 sequences.
First, to validate the enhancer-reporter vector, 2 control ex vivo electroporation experiments were performed in the E13.5 mouse embryonic cortex. (1) Coelectroporation of pCMV-mCherry and the reporter vector without enhancer showed mCherry expression in the embryonic cortex, validating the electroporation and cortical slice culture procedures; but did not show GFP expression, validating that the minimal β-globin promoter had no basal activity without enhancer (Fig. 3A,B). (2) Coelectroporation of pCMV-mCherry and the reporter vector containing a known Rnd2 enhancer (Heng et al. 2008) showed both mCherry and GFP expression, thus validating the use of the chosen vector for testing enhancer function in the cortex (data not shown).

When tested in the ex vivo coelectroporation assay with pCMV-mCherry, 5 of the 18 putative enhancers cloned in the reporter vector (27%) were active in the embryonic cortex between E13.5 and E14.5 (Fig. 3A,C, see also last column in Table 1, and Supplementary Figure 1 for inactive sequences electroporation tests). The 5 sequences that act as active enhancers in the mouse cortex are: enh14(Cdh4), enh15(Sall3), enh16(Gsh1), enh17(Pptrz1), and enh18(Pou2f1). Two of them are intronic, 2 are located in the 5’ of the closest gene and 1 in the 3’ of the closest gene. We also noticed that enh18 (Pou2f1), the only enhancer exhibiting pan-vertebrate conservation in the candidate list, is active and drives strong GFP expression (Table 1).

TFBS Composition and Mutagenesis in Enhancers with Activity in the Developing Cortex

We then checked whether particular combinations of TFBSs correlate with active enhancers in the cortex.

First, the consensus TFBSs corresponding to Sox and Pou factors used in the ChiP-seq experiment were searched for in the 20 candidate sequences of Table 1 (see Materials and Methods), and their distributions were compared between active (n = 7, Fig. 4A) and inactive (n = 13, Fig. 4B) sequences. This analysis highlighted several differences between the 2 categories of enhancers. First, Pou-type TFBSs (but not Sox) were more represented in active enhancer sequences (normalized to sequence length; Fig. 4C). Secondly, the average distance between TFBSs in Pou/Sox pairs were twice shorter in active enhancers (16 ± 6 bp) than in inactive sequences (40 ± 8 bp; P = 0.0446). Thirdly, we also checked for the presence of p300 transcriptional coactivator binding sites; 71% of active enhancers sequences (i.e., 5 of 7 sequences) had a p300 binding site (source: Visel et al. 2009; forebrain dataset), while only 30% of inactive sequences showed such p300 peak (Fig. 4D).

In addition, the occurrences of Pou and Sox consensus TFBS in active enhancers were compared with a random segment of the genome. Sox TFBSs were 2.1× more frequent in our 7 active enhancers sequences (16 ± 6 bp) than in inactive sequences (40 ± 8 bp; P = 0.0446). Thirdly, we also checked for the presence of p300 transcriptional coactivator binding sites; 71% of active enhancers sequences (i.e., 5 of 7 sequences) had a p300 binding site (source: Visel et al. 2009; forebrain dataset), while only 30% of inactive sequences showed such p300 peak (Fig. 4D).

Finally, the PWMs for Sox and Pou TFBSs were 14× more frequent in our 7 active enhancers than in inactive sequences (i.e., 5 of 7 sequences) had a p300 binding site (source: Visel et al. 2009; forebrain dataset), while only 30% of inactive sequences showed such p300 peak (Fig. 4D).

To test the hypothesis that a combination of Sox and Pou TFBSs is indeed important for the enhancer activity, we next performed enhancer mutagenesis by the deletion of TFBS motifs on 2 selected enhancers. We chose enh14 (Gdh4) because it harbors only 2 TFBS of interest, 1 Sox and 1 Pou; and we chose enh18 (Pou2f1) because it is on the contrary the most TFBS-rich among active sequences, containing 7 Sox and 5 Pou binding sites.

When tested in the ex vivo electroporation assay, enh14 (Gdh4) Pou-deleted (del1) or Sox-deleted (del2) were both totally inactive in the embryonic cortex (Fig. 5A), suggesting that the combination and interaction of the 2 binding sites are required to confer enhancer activity.
We also used enh18 (Pou2f1) to challenge our above observation that the distance between Sox and Pou TFBSs may matter to confer enhancer activity. We therefore generated constructs with single or double deletions of the 2 closest Sox and Pou binding sites in enh18 (Fig. 5B, schema). Both one-Pou-deleted (del1) and one-Sox-deleted (del2) versions of enh18 were still active in the embryonic cortex. However, enh18del1 showed much weaker activity than the intact enh18 sequence [compare 75% of GFP-positive slices after intact enh18 electroporation with 31% with enh18(del1); Fig. 5B]. This suggests that adjacent TFBS may partly compensate for the deleted Pou site. Finally, the deletion of both TFBSs (del3) further reduced enh18 activity: We observed 27% of GFP-positive slices after enh18(del3) electroporation; and in all cases, the GFP-positive zone was very reduced when compared with the electroporated area visualized by mCherry (Fig. 5B). Thus, in enh18 (Pou2f1), the deletion of the 2 closest Sox and Pou binding sites strongly reduces the activity of the enhancer.

**Enhancer Activity in Cortical Progenitors**

We next used in utero electroporation to further characterize the activity of selected enhancers in cortical progenitors at the cellular level and with a good anatomical resolution. We selected the 3 elements with strongest activities ex vivo, enh14 (Cdh4), enh15(Sall3), and enh18(Pou2f1) to perform reporter assays at E13.5. One day after in utero electroporation, embryos were removed and systematic double-labeling for GFP and either pH3 (a marker for mitotic cells) or Tbr2 (a specific marker for cortical BPs) was performed in order to characterize the proliferative activity and the type of progenitor cells in which the enhancers were active, respectively (Fig. 6). pCMVCAGGS-GFP electroporation was taken as a control, as the strong and ubiquitous promoter drives GFP expression in all electroporated cells and their progeny, without cell type selectivity.

First, the position of GFP+ cells (i.e. the electroporated cells themselves and/or their progeny) in the cortex was studied (Figs 6 and 7A,B). Twenty-four hours after electroporation of control pCMVCAGGS-GFP (gray bars in Fig. 7), some of the GFP+ cells still resided in the VZ (14.3%, most probably corresponding to APs), while a large majority had migrated to the SVZ (62.2%) and some had already reached the cortical plate (CP, 22.6%). The distribution of GFP+ cells was markedly different from the pCMVCAGGS-GFP control for the 3 tested enhancers, highlighting the specific activity of each enhancer in given cell populations. Enh14(Cdh4) and enh15(Sall3) activities were almost in mirror image of enh18(Pou2f1) activity. Indeed, enh18(Pou2f1) was mostly active in the VZ-SVZ, while enh15(Sall3) and enh14(Cdh4) were active in the SVZ and CP, but not (or weakly) in the VZ. These significantly different activity patterns suggest that enh14(Cdh4) and enh15 (Sall3) are activated specifically in cells only when they are engaged in the differentiation pathway, whereas cells in which enh18(Pou2f1) is active are likely to be germinall zone...
active cells. The percentages of pH3+ cells were not statistically different from the pCMVCAGGS-GFP control value for enh14 (Cdh4) and enh18(Pou2f1) electroporations. For these 2 latter enhancers, among the GFP+/pH3+ cells, we further distinguished between those lining the ventricular surface and those located above or far from the ventricular surface, up to the SVZ (see schema in Fig. 7A). In pCMVCAGGS-GFP electroporated cells, the majority of the mitoses occurred along the ventricle in a very apical position, strongly suggesting that they occur in APs, while a few GFP+ cells also underwent mitoses higher in the cortical depth. The same pattern was observed for cells with enh14(Cdh4) activity. In contrast, enh18(Pou2f1) was only active in mitotic cells that divided along the ventricular surface (Fig. 7D).

In addition, we sought to characterize the activity of the 3 enhancers in SVZ progenitors. More specifically, we assessed the BP identity of GFP+ cells through Tbr2 immunofluorescence staining (Figs 6B and 7E). In control pCMVCAGGS-GFP electroporated embryos, about half of the SVZ GFP+ cells were Tbr2 immunopositive. In contrast, after electroporation of the 3 enhancer-reporter constructs, only approximately 15% of the GFP+ cells located in the SVZ were Tbr2-positive, therefore suggesting that the population of Tbr2+ cells is heterogeneous.

**mRNA Expression of Enhancer’s Nearest Flanking Genes**

To further refine the study of enh14(Cdh4), enh15(Sall3), and enh18(Pou2f1) patterns of activity, we asked whether these 3 enhancers reproduce entirely, partially, or not at all, the mRNA expression patterns of their flanking genes (of note, there is no direct evidence that these enhancers actually regulate the expression of the nearby genes). To this end, we performed ISH experiments at E11.5, E12.5, and E13.5 (Fig. 8A).

Along these 3 stages, Cdh4 mRNA was expressed dynamically, first in the upper part of the VZ at E11.5, then in the intermediate zone (IZ) and at lower levels in the VZ/SVZ at E12.5, and finally in only the upper part of the IZ at E13.5 (Fig. 8A). Sall3 and Pou2f1 mRNAs showed more prominent expression in the germinal zones of the cortex at the 3 studied stages. They were both strongly expressed in the VZ/SVZ, with Pou2f1 (but not Sall3) also showing expression in the IZ at E12.5–E13.5 (Fig. 8A2,A3). In summary, the mRNAs for the flanking genes showed globally wider expression than did the activity of their respective putative tested enhancers.

**Discussion**

We have selected 20 sequences with potential enhancer activity in cortical progenitor cells from a ChIP-seq dataset, using a selection pipeline taking into account function and expression of the closest flanking genes. Seven of these 20 sequences (35%) are active enhancers, and this was correlated to a specific signature in terms of TFBS composition, but not evolutionary conservation. Further characterization of cell types in which 3 of these enhancers are active highlights a possible regulatory basis for the heterogeneity of cortical progenitors.
Predictive Value of the Selection Pipeline on the ChIP-seq Dataset

The ChIP-seq experiment that served as a basis of this study was performed on NS5 cells, a mouse cell line derived from embryonic stem cells (Conti et al. 2005). To select sequences with potential enhancer activity in cortical progenitors, we applied 2 filters, not on the sequences themselves, but on the closest flanking gene. The first filter was a functional annotation (GO term) including neuron/nervous system or proliferation, and the second one was an expression pattern in the developing cortex, thereby conferring (1) neuronal progenitor and (2) regional brain specificity to the selected sequences to be tested. Of these 20 resulting sequences, about a third (7 of 20, 35%) behave as enhancers in cortical cells, thus showing a very significant enrichment in “cortical enhancers.” In comparison, using a similar approach with a selection pipeline on 3100 noncoding sequences based on human–fugu conservation and the presence of putative forebrain motifs, Pennacchio et al. (2006) showed that 17% of the tested sequences (4 of 23) are active in the mouse forebrain. Further, the sequences that

Figure 4. In silico analysis of the 20 candidate enhancer sequences. (A and B) Distribution of Sox and Pou TFBSs along the active (A) and inactive (B) sequences. Regulatory sequence analysis tools visualization shows Sox (red) and Pou (blue) TFBS as vertical lines along the sequences. Note the various lengths of the sequences (scale bar on top). (C) Histogram showing the mean frequency of nucleotides involved in Sox (red) and Pou (blue) TFBSs for active (act) enhancers (bright color) versus inactive (inact) sequences (dim color). The 2 asterisks for Pou TFBSs indicate $P < 0.01$ (Mann–Whitney). (D) Histogram showing the percentage of sequences with p300 TFBS in active (black) and inactive sequences (gray). (E) Sox and Pou binding sites in cortical progenitors. First line: consensus TFBS from the literature. Second line: PWMs elicited from this study. Third line: comparison with general Uniprobe PWM.
were then characterized with some details through in utero electroporation indeed showed enhancer activity in cortical progenitors. This suggests that our selection pipeline confers a high probability for a selected sequence to actually be an active enhancer in a chosen cell type of a specific brain region and may be valuably applied in other cases. In addition, it also points that such a strategy could be advantageously used to find specific drivers for subsequent functional and transgenesis experiments. In the case of the cortex, very few developmental enhancers specifically active in particular cell populations have been characterized. These include the E1 enhancer element of the Ngn2 gene that activity is confined to a subpopulation of progenitors predominantly in the region of the ventral and lateral pallium (Scardigli et al. 2001; Berger et al. 2004), or an intronic nestin enhancer active in the dorsal telencephalic germinal zone (Walker et al. 2010). More recently, a series of enhancers that can be used as tissue-specific reagents in different dorso-ventral domains of the telencephalon have been reported (Visel et al. 2013). Here, we provide 3 novel enhancers/drivers, with specific activities in certain types of cortical progenitors, and that may be used to study the biology of these progenitors through time-lapse imaging and gene function analyses, or to analyze their progeny through lineage studies. In a recent review on mammalian neural stem cells, Basak and Taylor (2009) indeed stressed the need for “clean lineage tracing experiments,” hence the need for specific enhancers/promoters active in neural progenitor cells.

**TFBS Signature and Regulatory Logics for Enhancers Active in the Cortical Germinal Zone**

When compared with the 13 inactive sequences, the 7 active enhancer sequences showed a higher frequency of Pou TFBS; their TFBS organized in Sox/Pou pairs were close (16 bp) on the sequence; and 71% (5 of 7) presented a p300 ChIP-seq peak. p300 is a transcriptional coactivator that is specifically recruited at enhancers, and was therefore expected as a mark for active enhancers. Accordingly, only 4 of 13 (30%) of inactive sequences had a p300 binding site. However, this criterion does not appear to be strict, and we would actually have “missed” 2 active sequences (i.e., enh14(cdh4) and enh15(Sall3)) if the criterion had been included in the selection pipeline, Visel et al. 2009). In the same line, the data from the DNase hypersensitivity sites (DHS) experiments (Sabo et al. 2004, 2006; John et al. 2011) of the ENCODE project (ENCODE Project Consortium et al. 2012) are relevant, but not a strict criterion, to be used to identify TFBSs. Indeed, DHS is observed for 6 of 7 (85%) of our active enhancers, and enh17(Ptrprz1) would have been discarded from our selection if we had followed this criterion to select the sequences. In addition, DHS is present in 54% of our inactive sequences. These elements of

![Figure 5](https://example.com/figure5.png)

Figure 5. Mutagenesis by the deletion of Sox and/or Pou TFBSs on enh14(Cdh4) and enh18(Pou2f1). (A) enh14(Cdh4) mutagenesis, (B) enh18(Pou2f1) mutagenesis. On each panel, the top left photograph shows merged GFP and mCherry fluorescence images after electroporation of the intact enhancer. The top right panel shows the Pou (blue) and Sox (red) organization of TFBS on the sequence, as well as the TFBS-deleted constructs that were generated by mutagenesis. The bottom panels show GFP, mCherry, and merged fluorescence photographs after electroporation of mutagenized constructs. Left column, low magnification showing the entire slice/hemisphere. Other columns, higher magnification on the electroporated zone. The N given at the bottom left corner of each low magnification picture gives the number of slices in which GFP could be detected in the electroporated area, among all the correctly electroporated, RFP+, slices. Scale bar: 250 µm.
discussion strengthen the validity and accuracy of the criteria (GO term and expression pattern) we used in our selection pipeline.

Sox2, known as an essential player in establishing and maintaining neuronal progenitors in vertebrates (Mizuseki et al. 1998; Bylund et al. 2003; Graham et al. 2003; Uchikawa et al. 2003), is known to pair off with specific partners to regulate gene transcription (Kamachi et al. 2000; Kondoh and Kamachi 2010). In particular, Sox/Pou TF dimerization and complex recruitment on enhancers are classically reported in the literature, including on the enh19(Nestin) and enh20(Sox2) sequences reported herein. Tanaka et al. (2004) have shown that Sox and Pou proteins work synergistically to activate the intronic Nes30 enhancer (present study enh19(Nestin)) in the VZ and SVZ of the mouse embryonic spinal cord. More recently, Walker et al. (2010) showed that the rat nestin enhancer (96% identical to Nes30) is also active in the VZ and SVZ of the developing cortex. Concerning the Sox2 enhancer, gel-shift analyses have identified a binding site for a Sox/Pou complex in SRR2 (present study enh20(Sox2); Tomioka et al. 2002). This SRR2 enhancer sequence functionally plays an important role in maintaining stem cells identity (Zappone et al. 2000) and drives transcription in neural stem/progenitor cells (Miyagi et al. 2006). Other examples of Sox/Pou complex formation exist, some in the context of stem cell biology: There is a precise correlation between the ability of Pou proteins to form a complex with Sox2 on the enhancer of undifferentiated transcription factor 1 (UTF1), which belongs to the core transcriptional network characterizing pluripotency and the ability to maintain the stem cell state in embryonic stem cells (Nishimoto et al. 1999). The same type of Pou/Sox complex binds to enhancers of Fgf4 (Reményi et al. 2003). Pou/Sox dimerization is also necessary for melanocyte development; and in this case, an interaction with p300 was also reported (Smit et al. 2000); or for Drosophila embryogenesis, in which case Sox, Pou but also bHLH, factors interact on a single enhancer of the slit gene (Ma et al. 2000). Our data are in line with these bibliographic elements. Further, we propose that a special regulatory logic may exist specifically in cortical progenitors, which includes a strong control by Pou TFs, as well as a crucial role for Pou/Sox TFs complexes. This logic is fully supported by our data:

Figure 6. Characterization of cortical progenitors in which enh14(Cdh4), enh15(Sall3), and enh18(Pou2f1) are active after in utero electroporation. (A and B) Photographs of fixed cortical slices after anti-GFP (green), and anti-ph3 (red) (A) or anti-Tbr2 (red) (B) double immunofluorescence staining, 1 day after in utero electroporation in the telencephalon of E13.5 mouse embryos. For each panel, the electroporated construct is indicated on the left. In A and B panels, the first column shows a general view of the electroporated cortical area, and the second column shows a high magnification on the zone in white square, with arrowheads indicating colocalization of GFP and the marker. Scale bar: 50 µm.
Figure 7. Quantification of pH3-positive and Tbr2-positive cortical progenitors in which enh14(Cdh4), enh15(Sall3), and enh18(Pou2f1) are active. (A) Schematic of the different layers/zones of the E14.5 developing cortex used for quantification. GFP-positive cells (green dots) were counted in the CP/IZ in the subventricular zone (SVZ) defined by dense Tbr2 immunostaining (gray), and in the ventricular zone (VZ). In the latter, the ventricular surface (corresponding to the thickness of 2 cell diameters) was distinguished from the rest of the VZ for pH3 analysis. (B) Distribution of GFP-positive cells in various cortical layers/germinal zones 1 day after in utero electroporation of GFP reporter-enhancer constructs. The color code for enhancers is illustrated in the legend. The activity of enhancers was quantified in different cell layers with a combination of pH3 and Tbr2 immunostaining. The histogram in panel C shows the percentage of pH3-positive cells among all GFP-positive cells in the VZ and SVZ. (D) Quantification of GFP-positive cells with regard to Tbr2 immunoreactivity. In all panels: *P < 0.05; **P < 0.01; ***P < 0.001; ****P < 0.0001 (Mann–Whitney).

1. enh15(Sall3) is the only active enhancer without Pou TFBS on its sequence and is also never active in the VZ where self-renewing, stem cell-like radial glia progenitors reside.
2. enh14(Cdh4) and enh18(Pou2f1) do present Pou/Sox TFBS pair(s) on their sequence, and they are both active in proliferating VZ progenitors.
3. single or double Sox/Pou site deletions on enh14 and enh18 abolishes or strongly reduces enhancer activity, indicating that an interaction between the 2 TFBSs is required for enhancer activity, and
4. in enh18(Pou2f1) that contains multiple Sox and Pou binding sites, the single deletion of one Pou (but not one Sox) binding site decreases enhancer activity, pointing the importance of Pou TFBS for transcriptional control.

We have also found a tendency for the Sox/Pou TFBS pairs in active enhancers to be spaced by about 16 bp. This is to be compared with a distance of 3 or 0 bp on Fgf4 orUtf1 enhancers, respectively (Reményi et al. 2003), or of 54 or 28 bp on the Drosophila slit enhancer (Ma et al. 2000), and is therefore in the range of described distances allowing functional interaction of the 2 types of TFs on DNA. We have tested this distance issue through mutagenesis on enh18(pou2f1), choosing the 2 closest Pou and Sox sites on the sequence (28 bp between the selected pair). The severe reduction in enhancer activity observed after the double deletion of the closest Pou/Sox pair on enh18 suggests that the distance is probably a criterion to take into account when considering the functional/physical interactions of these 2 particular TFs on DNA.

Enhancer Activity: cis-Regulatory Logics Underlying Cortical Progenitor Heterogeneity

We have characterized with some details the activity of enh14(Cdh4), enh15(Sall3), and enh18(Pou2f1). Strikingly, these 3 enhancers are active in clearly different types of progenitors, highlighting heterogeneity of progenitor cells in terms of gene cis-regulation (summarized in Fig. 8B). While enh18(Pou2f1) is active in proliferating APs and nonproliferating SVZ progenitors, enh15(Sall3) is on the contrary never active in APs—although the mRNAs of their closest flanking gene are expressed throughout the VZ/SVZ in a very similar pattern. Many recent neurodevelopment biology studies as well as evolutionary comparative analyses on cerebral cortex development converge on the idea that the AP/BP classification of cortical progenitors, although very useful and generally pertinent, does not reflect the actual heterogeneity of these cells. For example, the observation of a unique proliferative compartment in the monkey embryonic cortex, represented by a very large outer SVZ (OSVZ), first suggested a potential mechanism for primate cortical expansion (Smart et al. 2002). These OSVZ progenitors are heterogeneous and include both AP/radial glia-like stem cells (Pax6+, self-renewing, stem cell-like) and BP (Tbr2+, transit amplifying) cell types (Fietz et al. 2010; Hansen et al. 2010). Such OSVZ-like progenitors are in fact found in all mammals, both gyrencephalic (Reillo et al. 2011) and lissencephalic including the mouse (Kelava et al. 2011; Shitamukai et al. 2011; Wang et al. 2011). Ongoing studies on ferrets (nonprimate, gyrencephalic) have begun to uncover subtle differences when compared with primates (Reillo and Borrell 2012). Among these differences, some are quantitative (radial glia abundance), some are qualitative (genetic programs controlling cell cycle kinetics and dynamics of self-renewal), and some are time-dependent. Here, we bring some evidences that the heterogeneity and diversity in mouse cortical progenitor types may lie in a finely tuned differential transcriptional control, with cortical progenitor subtypes having specific regulatory signatures and activities. With this regard, it is also striking that the evolutionary conservation of active mouse enhancer sequences is relatively poor with other mammals. Enh14(Cdh4), which is active in both AP- and BP-like progenitors, is conserved between mouse and human, but not dog. Enh15(Sall3), which is active only in the SVZ, is conserved only between mouse and chick (but note that birds do not have a laminated pallium). These features suggest that fine regulation of gene expression and progenitor behavior differ between closely related mammalian species, and such differences may underlie subtle evolutionary variations in cortical architecture across mammals. Only enh18(Pou2f1), active in proliferating VZ progenitors, is conserved across all vertebrates. This suggests that this regulatory input may be shared by self-renewing, stem cell-like radial glial cells of the (dorsal) telencephalon in all vertebrate species. Such a hypothesis will have to be tested in birds and fishes in a near future.

Enhancer Activity versus Expression Pattern of the Flanking Gene

The activity of the 3 enhancers that we have described with some details does not recapitulate totally the expression...
pattern of the closest flanking gene (Fig. 8). We cannot rule out that other genes (i.e., not the closest flanking gene but a more distal one) may be regulated by these enhancers. Yet, the comparison between enhancer activity and flanking gene expression pattern is quite satisfactory: The activity of enh14 (Cdh4) in cells in the differentiation/migration to the CP process is correlated with low levels of Cdh4 mRNA in the cortical germinal zones; the activity of enh15 (Sall3), excluded from the VZ where Sall3 mRNA is expressed, suggests that other enhancer(s) are responsible for Sall3 expression in APs. Conversely, enh18 (Pou2f1) activity accounts for Pou2f1 mRNA expression in the germinal zones, but not in the lower part of the IZ/CP. In all cases, it appears that the regulation of mRNA expression is modular, with different enhancers involved at different steps of corticogenesis, underscoring the crucial importance of timing and dynamics of gene regulation in this developmental process.

Conclusion
Through the window of transcriptional cis-regulation, we have found that cortical progenitor types in the mouse embryonic cortex are heterogeneous. They nevertheless share a common cis-regulatory logic, involving Pou TFBS together with Sox TFBS, to control their gene expression. Importantly, our approach also provides novel enhancer sequences that can be used as drivers for future analyses of corticogenesis and cortical progenitor biology.

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

Funding
This work was supported by the FP7 European project CISSTEM.

Notes
The authors thank to Yan Jaszczyszyn and Aurélie Heuzé for their help with Gateway cloning and for modifying the p336bglobGFP3H vector, a gift of François Spitz. Many thanks to Yohann Mérot who established ex vivo and in utero electroporations in the group and gave valuable advices on these methods. Conflict of Interest: None declared.

References


Characterization of Enhancers Active in the Mouse Cerebral Cortex


