Title
Transcriptional Regulation of Patterning of the Cerebral Cortex

Author
Pattabiraman, Kartik

Publication Date
2014

Peer reviewed|Thesis/dissertation

Transcriptional Regulation of Patterning of the Cerebral Cortex

by

Kartik Pattabiraman

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Neuroscience

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

[redacted]

My thesis is dedicated to my parents and my sister, Kavita.

Abstract

The adult human cortex is divided into histologically and functionally distinct domains, which are specified early in development prior to neuronal differentiation. A handful of transcription factors (TFs) have been shown to be involved in pallial regionalization. Analysis of E11.5 expression data for 722 TFs has identified putative novel patterning TFs, which are expressed in gradients in pallium. From our analysis, we choose to study 28 TFs in more detail at E11.5. We investigated their expression in C*oupTFI, Emx2* and *Pax6* mouse mutants, TFs previously described to be involved in patterning.

TFs with rostrocaudal gradients of expression showed similar changes in the *Pax6* and *CoupTf1* mutant. This finding is unexpected based on their opposing roles in patterning. We provide evidence that *Pax6* and *CoupTf1* work together in ventral cortical development. Furthermore, both TFs bind together on putative regulatory regions of genes involved in forebrain development.

TFs that embryonically regulate pallial regionalization are expressed in gradients, raising the question of how discrete domains are generated. We provide evidence that small enhancer elements active in protodomains integrate broad transcriptional information. CreER^{T2} and GFP expression from 14 different enhancer elements in stable transgenic mice allowed us to define the first comprehensive regional fate map of the pallium. We explored transcriptional mechanisms that control the activity of the enhancers using informatics, *in vivo* occupancy by TFs that regulate cortical patterning (*CoupTFI, Pax6* and *Pbx1*), and analysis of enhancer activity in *Pax6* mutants. Overall, the results provide novel insights into how broadly expressed patterning TFs regulate the activity of small enhancer elements that drive gene expression in pallial protodomains that fate map to distinct cortical regions.

Table of Contents

Tables

Figures

Chapter 1: Background

The cerebral cortex is considered the most complicated structure in the human body. It consists of billions of neurons, with each neuron forming thousands of the connections with other neurons. This complexity underlies the multitude of behaviors we perform, including locomotion, vision, problem solving and displaying and interpreting emotions. Dysfunction in the cerebral cortex can lead to sensory processing deficits, and psychiatric diseases including personality disorders, depression, autism, and schizophrenia .

In the early 1900s, the cerebral cortex was divided into distinct regions based on cytoarchitectural organization of neurons visualized using nissl stain and gross morphological boundaries (Broadmann, 1909; Economo and Koskinas, 1925). These areas were further refined through study of model organisms, direct stimulation of different regions in awake humans, and more recently, neuroimaging studies (Mountcastle, 1978). The identification of anatomically and functionally distinct regions of the cerebral cortex created interest in how these areas are formed during development. The cerebral cortex arises from an initially equipotent set of progenitor cells. During development, these identical cells differentiate into distinct classes of neurons. It was initially thought that regions of the cortex were specified by input in the cortex from the thalamus (Creutzfeldt, 1977). This concept is called as the prototype hypothesis or "tabula rasa" (blank slate). Alternatively, neurons of the cortex could be specified at the time of their birth. This was termed the protomap hypothesis (Rakic, 1988).

While input into the cortex is important to maintain regionalization in the cortex, there have been several lines of evidence supporting the protomap hypothesis. First, progenitor cells in the ventricular zone (VZ) do express the same set of genes. In the developing cortex, there are different patterning centers that release growth factors shown to be important in patterning. These include the rostral patterning center (FGF8 and FGF17), dorsal hem (BMPs and Wnts), pallial-subpallial

boundary (FGF15 and Wnt antagonists), and a ventral source of Shh (Figure1; Rakic et al 2009). These morphogens diffuse through the cells of the VZ in gradients, and thus creates gradients of gene expression in the VZ. For example, FGF8 from the rostral patterning center increases of expression of several transcription factors (TFs) in the rostral cortex, and also represses levels of CoupTf1, which is important to caudal cortical fates (Storm et al. 2006) Second, early regionalization occurs in absence of thalamocortical (TCA) innervation (Catalano and Shatz, 1998). Nonetheless, innervation of the cortex is required to achieve the final properties of a specific cortical fate. Third, intra cortical transplant assays, in which regions of the somatosensory cortex were transplanted to the visual cortex, showed that the cortical VZ were specified prior to TCA innveration (Gitton et al. 1999). These experiments showed that inputs in the cortex act on a pre-specified cortex.

My thesis work works towards further understanding the mechanism of cortical patterning. The first chapter looks at the expression pattern of transcription factors in the cortex at E11.5, and by analyzing mutants of known patterning TFs, identify more patterning TFs. Finally, using TF and histone marks ChIP-seq datasets, works towards building TF regulatory networks involved in patterning. The second chapter focuses on evolutionarily conserved enhancers active in the developing cortex. By analyzing stable enhancer transgenic lines, we are able to assign adult fates to protodomains of the cortical VZ. We also use TF ChIP-seq and mutant analysis to study enhancer regulation. In the third chapter, we further analyze enhancer 660 to study a putative novel source of cortical subplate cells.

Bibliography

Brodmann, K. (1909) Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues. Johann Ambrosius Barth Verlag

Economo, C. and Koskinas, G.N. (1925) Die Cytoarchitektonik der Hirnrinde des erwachsenen Menschen. Julius Springer

Mountcastle, V. B. (1978), An Organizing Principle for Cerebral Function: The Unit Model and the Distributed System, in Gerald M. Edelman and Vernon B. Mountcastle, *The Mindful Brain*, MIT Press

Rakic, P. (1988) Specification of cerebral cortical areas. Science 241, 170–176

Creutzfeldt, O.D. (1977) Generality of the functional structure of the neocortex. Naturwissenschaften 64, 507–517

Rakic P, Ayoub AE, Breunig JJ, Dominguez MH (2009) Decision by Decision; Making a Cortical Map. Trends in Neuroscience 32(5), 291-301

Storm EE, Garel S, Borello U, Hebert JM, Martinez S, McConnell SK, Martin GR, Rubenstein JL (2006) Dose-dependent functions of Fgf8 in regulating telencephalic patterning centers. Development 133(9), 1831-44

Catalano, S.M. and Shatz, C.J. (1998) Activity-dependent cortical target selection by thalamic axons. Science 281, 559–562

Gitton Y, Cohen-Tannoudji M, Wassef M (1999) Specification of somatosensory area identity in cortical explants. Journal of Neuroscience 19(12), 4889-98

Chapter 2: Identification of Novel Transcription Factors and Transcription Factor Networks in Cerebral Cortical Patterning

Introduction

Previous analysis of cortical patterning has focused on individual transcription factors (TFs) and phenotypes seen in their respective loss of function mouse lines. This approach has identified a handful of transcription factors including Pax6, CoupTf1 (Nr2f1), and Emx2 (Samson and Livesey, 2009). However, further studies are required to understand the mechanism of cortical patterning. We are interested identifying novel TFs involved in cortical patterning based on expression patterns, specifically gradients and sharp boundaries in the ventricular zone. We are also interested in how these TFs work together to form gene networks important for regulating this process. Our approach is the study the effect on the gene expression pattern of these TFs in loss of function mutants with prominent patterning defects (Pax6, CoupTf1, and Emx2). The results of this study will be cross-referenced with ChIP-seq datasets for these patterning TF (Pax6 and CoupTf1). More general analysis of both TFs ChIP also provides insight into how these TFs work together in cortical patterning.

Results

Pallial Expression of 722 TFs at E11.5 and E13.5

ISH was performed using oligonucleotide probes for 722 TFs on sagittal sections of E11.5 and E13.5 mouse embryos (see: http://developingmouse.brain-map.org/). For instance, *Pax6* showed the well-known rostroventral-caudodorsal gradient of expression in pallial progenitor zones (Figure 1). We annotated the pallial expression of all TFs by assessing their regional and laminar expression patterns, using schemas that define pallial expression in ventrolateral, rostrodorsal, caudodorsal and medial pallial regions, progenitor (VZ and SVZ) and neuronal (MZ) layers (Figure 1). Here we

focused on progenitor zone (VZ and SVZ) expression, as we wished to concentrate on mechanisms that control regional patterning of the pallium. We assessed the level of expression based on both the intensity and the density of expression (0-5 scales)(Figure 1). Table X contains all of the annotations. The data were then computationally assorted in several ways to display TFs with similar expression domains (Figure SXXXX). For instance, Figure 1F shows TFs that are expressed at E11.5 in the different domains of VZ.

Several important results were derived from this analysis (Table 1). We will first address the E11.5 findings. We identified102 TFs that were clearly expressed in the pallium, but that were not broadly or ubiquitously expressed. There were very few TFs with regional expression boundaries in between pallial progenitor zones. We found 2 TFs that were largely restricted to the DP, 11 TFs largely restricted to the MP and 1 TFs largely restricted to the LVP. In addition, the majority of the TFs restricted to the MP were expressed in the primordia of the non-neurogenic choroid plexus and Fimbria. It was more common to find TFs that were broadly expressed across the E11.5 pallium; these genes were expressed in either a rostrocaudal (n= 58) or caudorostral (n=44) gradient. For instance, see the well-known rostroventral-caudodorsal gradient of Pax6 (Figure 1).

At E13.5 we identified roughly 263 TFs that were clearly expressed in the pallium, but that were not broadly or ubiquitously expressed. Compares to E11.5, fewer showed expression restricted to pallium subdomains, we only found 8 TFs largely restricted to the MP and none in the DP and LVP. As at E11.5, it was more common to find TFs that were broadly expressed across the E13.5 pallium; these genes were expressed in gradients. Unlike E11.5, more were expressed in caudorostral (n=73) than rostrocaudal (n= 24) gradients.

CoupTFI, Emx2 and Pax6 Regulation of Gradients of TF Expression at E11.5

From our analysis, we choose to study 28 TFs in more detail at E11.5. We based our choices on TFs that previously were not known to be expressed in the E11.5 pallium, TFs whose functions were

poorly known, and TFs with prominent E11.5 expression gradients. In addition to repeating their ISH at E11.5, we investigated their expression in C*oupTFI, Emx2* and *Pax6* mouse mutants (at E11.5)(Figures 2 and SXXX), as these TFs are known to have prominent defects in pallial patterning (Bishop et al. 2002; Mallamaci and Stoykova, 2006; Armentano et al., 2007; Faedo et al., 2008).

We organized our phenotypic descriptions according to the effect of the *CoupTFI, Emx2* and *Pax6* mutations on the expression of TFs that have rostrocaudal, caudorostral or regional expression patterns (Figures 2; SX). We annotated the expression changes based on the independent assessment of 2-3 experts using a 5-level qualitative expression scale: increased expression (green; ++ or +); no change; decreased expression (red; -- or -)(Figure 2D).

Pax6 promotes rostral identity and represses caudal identity (reference). Consistent with this, 6/8 genes with rostrocaudal gradients, such as *Etv5* (Figure 2A. S2X), showed reduced expression (*Etv5, Lmo1, Mycl, Npas3, Nr2e1* and *Meis2*) (Figure 2D. S2X). Likewise, 8/12 genes with caudorostral gradients, such as *Fezf2* (Figure 2B, S2X), showed increased expression (*AES, Bcl11a, Dach1, Dmrt3, Fezf2, Lef1, Lhx2, Nfix* and *Tcf4*) (Figure 2D, S2X).

Emx2 promotes caudal identity and represses rostral identity (reference). Consistent with this, 6/8 genes with rostrocaudal gradients, such as *Etv5* (Figure 2A. S2X), showed increased expression (*Etv5, Lmo1, Mycl, Nr2e1, Pou3f1* and *Sox9*) (Figure 2. S2X). Likewise, 8/12 genes with caudorostral gradients, such as *Fezf2* (Figure 2B), showed decreased expression (*AES, Bcl11a, Dach1, Dmrt3, Fezf2, Lef1, Nfix* and *Tcf4*) (Figure 2D; S2X).

While, generally, the effects of the *Pax6* and *Emx2* mutants were complementary, the effects of the *CoupTFI* mutant did not match that of the *Emx2* mutant, even though both are expressed in caudorostral gradients and both promote caudal identity and repress rostral identity (reference). Indeed, only 1/12 gene with a rostrocaudal gradient showed increased expression (*Pou3f1*) (Figure 2D; S2X), *a*nd only 1/12 gene with a caudorostral gradient showed decreased expression (*Nfix*).

These results made us reassess the idea that TFs that share caudorostral gradients have similar functions.

Although *CoupTFI and Emx2* are expressed in caudorostral gradients, they have opposite dorsoventral gradients. *Emx2* has a dorsoventral gradient, whereas like *Pax6*, *CoupTFI* has a ventrodorsal gradient (reference). Thus, perhaps TF expression differences between the *Emx2* and *CoupTFI* mutants relates to their differences in dorsoventral patterning functions. Indeed, in the CoupTFI mutant, the two down-regulated R-C TFs (Etv5 and Lmo1) were also down-regulated in the Pax6 mutant (Figure 2X). Thus, below (in at later section of the paper) we explored whether Pax6 and CoupTFI share functions in ventral patterning.

CoupTFI, Emx2 and Pax6 Regulation of TF Expression in the MP at E11.5

We also chose to analyze 4 TFs (*Id3, Lmx1a, Zic1* and *Zic5*) that have restricted expression in MP (Figure 2D; Figure S2X; Table 2). Both *Id3* and *Lmx1a* showed similar responses in the three loss-of-function mutant lines. Expression was strongly reduced in the *Emx2* mutant, and expression was shifted rostrally in the *Pax6* mutant. The results in the *Emx2* mutant are expected based on the importance of *Emx2* in medial pallium development (Tole et al. 2001). The *Zic* family TFs showed a different response to loss of *Pax6* and *Emx2*. Expression of Zic1 and Zic5 expanded to the DP in both mutants (Figure 2C; Figure S2X). There was no strong phenotype in the CoupTf1 mutant for any of the 4 TFs (Figure 2D).

Pax6, CoupTFI and CoupTFII regulate patterning of the ventrolateral pallium

The experiements described above suggest the TFs with ventrodorsal gradients (*Etv5, Lmo1*, and *Npas3*) are similarly regulated by *Pax6* and *CoupTf1.* Loss-of-function analyses show that Pax6 has prominent roles in promoting ventral properties to the pallium (Yun et al. 2001), whereas the CoupTFI$^{-/-}$ mutant doesn't provide evidence that it has a role in ventral patterning (Faedo et al.

2008). However, because CoupTFII is expressed in a partially overlapping pattern, it may compensate to loss of CoupTFI. Thus, we used mouse mutants to explore the relationship of Pax6, CoupTFI and CoupTFII in regulating pattering of ventral regions of the pallium.

To study the relationship of *Pax6, CoupTf1* and*, CoupTf2*, we analyzed expression of these Tfs in their respective loss of function mutant (Pax6 sey/sey and CoupTf1 -/-; CoupTf2 f/f; Emx1-cre). Expression of both *CoupTf1* and *CoupTf2* were reduced in the rostroventral pallium at E11.5 in the Pax6 mutant (Figure 3A, A', B and B'). In the CoupTf1/2 double mutant, Pax6 expression was slightly increased rostrally (Not shown). This data suggests *Pax6* is important to maintain expression of both *CoupTf1* and *CoupTf2*, while *CoupTf1/2* repress levels of *Pax6*. In addition, expression of CoupTf2 was increased ventrally in the CoupTf1 mutant, possibly explaining the lack ventral phenotype in the *CoupTf1* mutant.

Based on these finding, we propose that patterning of the ventral pallium by *Pax6*, involves *CoupTf1* and *CoupTf2*. Thus, we would expect both mutant lines to show similar rostroventral cortex phenotypes at later ages. In the rostroventral cortex at E16.5, ventral cortex markers, *Nurr1* and *Npas3*, were strongly reduced in both the Pax6 and CoupTf1/2 mutants (Figure 3X). Nurr1 expression was maintained ventrocaudally in the *Pax6* mutant, but lost in the *CoupTf1/2* mutants (Figure 3X). For other ventral markers, *Lmo3* and *Nrp2*, a reduction in expression was seen in the *Pax6* mutant and not the *CoupTf1/2* mutants (Figure 3X and not shown). This suggests that *Pax6* can also work through another pathway to specify the ventral cortex.

In order to interrogate how Pax6 and CoupTf1/2 regulate each other, we performed chromatin immunoprecipitation experiments (ChIP-seq) using Pax6 and CoupTf1 specific antibodies on E12.5 cortical tissue (Pattabiraman et al. 2014). There are Pax6 peaks intragenic to CoupTf1, and intragenic and downstream of CoupTf2. Pax6 peaks were also found on enhancer 1172, a putative CoupTf1 enhancer (Figure 3X) (Pattabiraman et al, 2014). In addition, Pax6 is able to activate

enhancer 1172 in P19 cells (Pattabiraman et al, 2014). CoupTf1 peaks were found on the promoter and intragenic to Pax6. This data provide evidence that regulation of Pax6 and CoupTf1/2 is through direct regulation on the genomic level.

Global Analysis of the E12.5 Pax6 and CoupTf1 ChIP datasets

While identified Pax6 and CoupTf1 peaks on promoters and putative enhancers regulating known and newly identified genes involved in cortical development and patterning, we were interested in general patterns of TF binding in the genome. To study this, we used the GREAT program created by the Bejarano Lab at Stanford University (http://bejerano.stanford.edu/great/public/html/index.php; McLean et al. 2010). This program associates putative genes being regulated with each ChIP peak based on proximity. Then, the program using gene annotation for numerous gene ontologies assigns an annotation for each peak. Finally, GREAT calculates statistical enrichment associations between ChIP peaks and annotations.

Analysis of both the E12.5 Pax6 and CoupTf1 ChIP showed statistical enrichment for the molecular function GO terms related to sequence-specific DNA binding activity and transcription factor activity. The E12.5 Pax6 ChIP was also associated with genes with Eph-Ephrin binding activity, while the E12.5 CoupTf1 ChIP was associated with WNT signaling activity. Both datasets were also enriched for genes involved in forebrain development and regulation of progenitor proliferation and differentiation. The Pax6 ChIP was also enriched for GO terms relevant to kidney development, while CoupTf1 for bone development.

More focused analysis of genomic regions surrounding TFs involved in cortical development showed Pax6 and CoupTf1 ChIP peaks on promoters, intragenic, and putative conserved enhancers from vista browser (enhancer.lbl.gov). Moreover, several of these relevant regions showed overlapping Pax6 and CoupTf1 peak (eg. Enhancers 112 (Dmrt3), 399 (Bcl11a), 844 (Sp8)). Approximately 20% of Pax6 peaks and 10% of CoupTf1 are overlapping with CoupTf1 and Pax6

peak respectively. GREAT analysis of the overlapping peaks showed further enrichment for the GO term transcription factor activity with over 20% of the peaks associated with related terms (compared to 4% and 2% for CoupTf1 and Pax6). The genomic regions with overlapping peaks are also enriched for GO terms related to nervous system and forebrain development (38% compared to 5% for each individually). These genomic regions are also not related to non-nervous system processes, as seen in the ChIP individually. This analysis suggests that Pax6 and CoupTf1 work together on the DNA level to regulate cortical development.

Understanding the interaction of Pax6 and CoupTf1

The previously described experiments provide evidence of interaction of Pax6 and CoupTf1 in genomic regions important for cortical development. However, peaks from ChIP-seq datasets do not prove direct binding of TFs to the DNA. To test if Pax6 and CoupTf1 are directly binding to these genomic regions, we searched the presence of Pax6 and CoupTf1 binding sites using RSAT (rsat.ulb.ac.be/). This program can analyze the sequences of genomic region from TF ChIP-seq datasets, and look for sequence motifs that are enriched. These enriched motifs are then run through a TF binding site database to predict the TF that binds to the motif. The E12.5 Pax6 ChIP is enriched for Pax6 paired domain binding sites, as well as several other TFs. The Pax6 paired domain binding site on average is located in the middle of these regions. The E12.5 CoupTf1 ChIP was enriched for nuclear factor binding sites including sites for Nr2f1 (CoupTf1). The nuclear receptor binding site on average is located in the center of the genomic region. This data further supports the specificity of the antibody for each transcription factor. RSAT analysis of the genomic regions with both Pax6 and CoupTf1 peaks did not enrich for either Pax6 paired domain or nuclear factor binding sites. Instead, these regions were enriched for homeobox binding sites. Pax6 contains both paired domain and a homeodomain. In genomic regions with both peaks, it is

possible the Pax6 binds the DNA through the homeodomain on Pax6, and CoupTf1 does not directly bind to the DNA.

Conclusion

Through our analysis of the expression patterns of TFs at both E11.5 and E13.5, we identified several TFs with interesting expression patterns. Specifically, the majority of these TFs are expressed ubiquitously or in gradients. The expression of only a small number of TFs had sharp boundaries. Analysis of gene expression changes in a set of these TFs revealed are novel TF regulatory network important for ventral cortical identity and possibly, cortical development.

Bibliography

Armentano, M., Chou, S.J., Tomassy, G.S., Leingärtner, A., O'Leary, D.D., and Studer, M. (2007). COUP-TF1 regulates the balance of cortical patterning between frontal/motor and sensory areas. Nat. Neurosci. 10,1277-1286.

Bishop, K.M., Goudreau, G., and O'Leary, D.D. (2000). Regulation of area identity in the mammalian neocortex by Emx2 and Pax6. Science. 288, 344-349.

Faedo, A., Tomassy, G.S., Ruan, Y., Teichmann, H., Krauss, S., Pleasure, S.J., Tsai, S.Y., Tsai, M.J., Studer, M., and Rubenstein, J.L. (2008). COUP-TF1 coordinates cortical patterning, neurogenesis, and laminar fate and modulates MAPK/ERK, AKT, and beta-catenin signaling. Cereb. Cortex. 18, 2117-2131.

Mallamaci, A., and Stoykova, A. (2006). Gene networks controlling early cerebral cortex arealization. Eur. J. Neurosci. 23, 847-856.

McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, Wenger AM, Bejerano G. (2010) GREAT improves functional interpretation of cis-regulatory regions**.** Nat Biotechnol. 28(5):495-501

Samson SN and Livesey FJ (2009) Gradients in the brain: the control of the development of form and function in the cerebral cortex. Cold Spring Harb Perspect Biol. 1(2):a002519

Tole S., Goudreau G., Assimacopoulos S, Grove EA. (2000) Emx2 is required for growth of the hippocampus but not for hippocampal field specification. J Neurosci. 20(7):2618-25.

Yun K, Potter S., Rubenstein JL. (2001) Gsh2 and Pax6 play complementary roles in dorsoventral patterning of the mammalian telencephalon. Development. 128(2): 193-205

Figure Legends

Figure 1. Annotation of TF expression in the E11.5 and E13.5 telencephalon. A,B: Schemas of sagittal view of the E11.5 (A) and E13.5 (B) telencephalon, hypothalamus and diencephalon. The pallium is red; regional subdivisions are indicated by straight black lines; laminar subdivisions are indicated by curved lines perpendicular to the regional boundaries. C,D: ISH analysis of Pax6 expression at E11.5 (C) and E13.5 (D). E: method of quantification based on a 0-5 scale used to describe the density and the intensity of expression. F: Expression levels of TFs with expression in the XXXXXX, showing differential expression in the XXX and YY.

Figure 2. CoupTFI, Emx2 and Pax6 Regulation of TF Expression at E11.5. ISH shows changes in Etv5, Fezf2, and Zic5 expression in E11.5 sagittal sections in CoupTFI, Emx2 and Pax6 mutants. Table annotating the effects of the CoupTFI, Emx2 and Pax6 mutations on the expression of 25 TFs. Annotation of expression changes are indicated by a 5-level qualitative expression scale: increased expression (green; ++ or +); no change; decreased expression (red; -- or -).

Figure 3. Co-regulation of ventral patterning by Pax6 and CoupTf1. Upper left panels: CoupTf1 and CoupTf2 mRNA expression at E11.5 in Pax6 control (column 1) and mutant sections (column 2). Upper right panels: E11.5 sections of Enhancer 1172 transient transgenic assay (top row). Bottom row: Pax6 ChIP peak (row 2) over enhancer 1172(black bar; row 1). Lower panels: E16.5 Pax6, CoupTf1/2 mutant analysis. Nurr1 mRNA expression in control (column 1), Pax6 mutant (column 2), CoupTf1/2 (column 3), and Npas3 mutant (rostral sections, row 1; caudal sections row 2). Lmo3 mRNA expression in control (column 1), Pax6 mutant (column 2), CoupTf1/2 (column 3), and Npas3 mutant

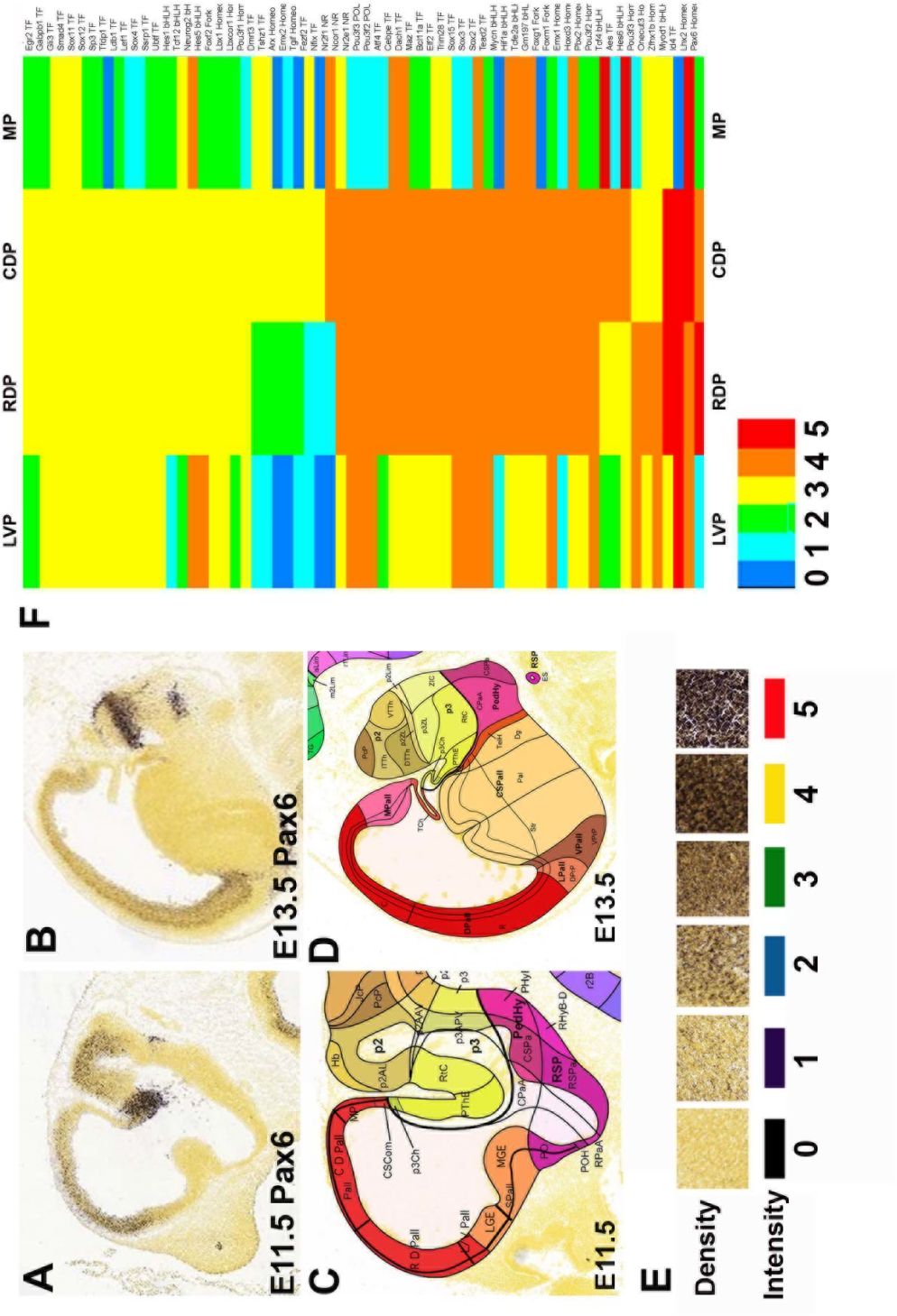
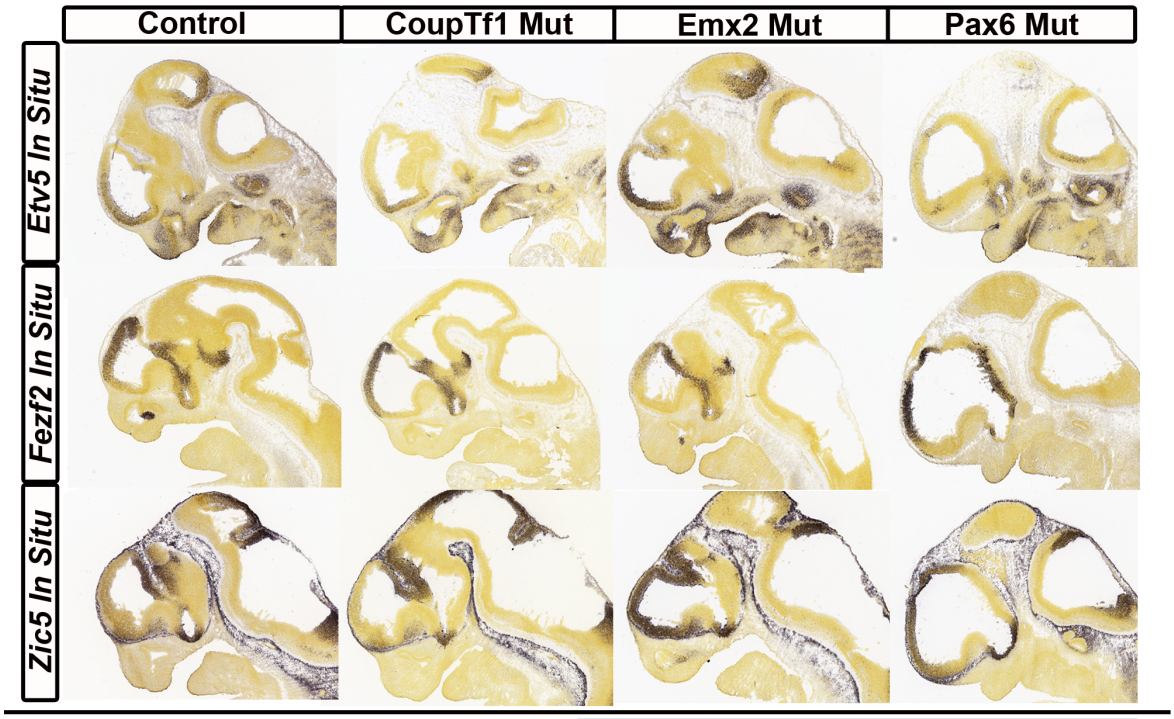

A
E11.5 Pax6
B
E13.5 Pax6
C
E11.5
D
E13.5
E
Density
Intensity
0 1 2 3 4 5
F
LVP
RDP
CDP
MP
0 1 2 3 4 5

Figure 1



	CoupTf1 Mut	Emx2 Mut	Pax6 Mut
TFs with Rostral to Caudal Gradient			
Etv5	-	++	--
Lmo1	--	+	--
Mycl		++	-
Npas3	-		--
Nr2e1		++	--
Meis2			--
Pou3f1	+	+	
Sox9		+	
TFs with Caudal to Rostral Gradient			
Aes		-	+
Bcl11a		--	++
Dach1		--	++
Dmrt3		--	--
Fezf2		--	++
Lef1		--	++
Lhx2			++
Nfix	-	--	++
Sp8	++	++	--
Tcf4	+	-	++

	CoupTf1 Mut	Emx2 Mut	Pax6 Mut
TFs with Regional Expression			
Id3		--	+
Lmx1a		--	
Nfatc4		--	
Tsh1		--	
Zic1	++		++
Zic5	+	++	++

++
+
-
--

Figure 2

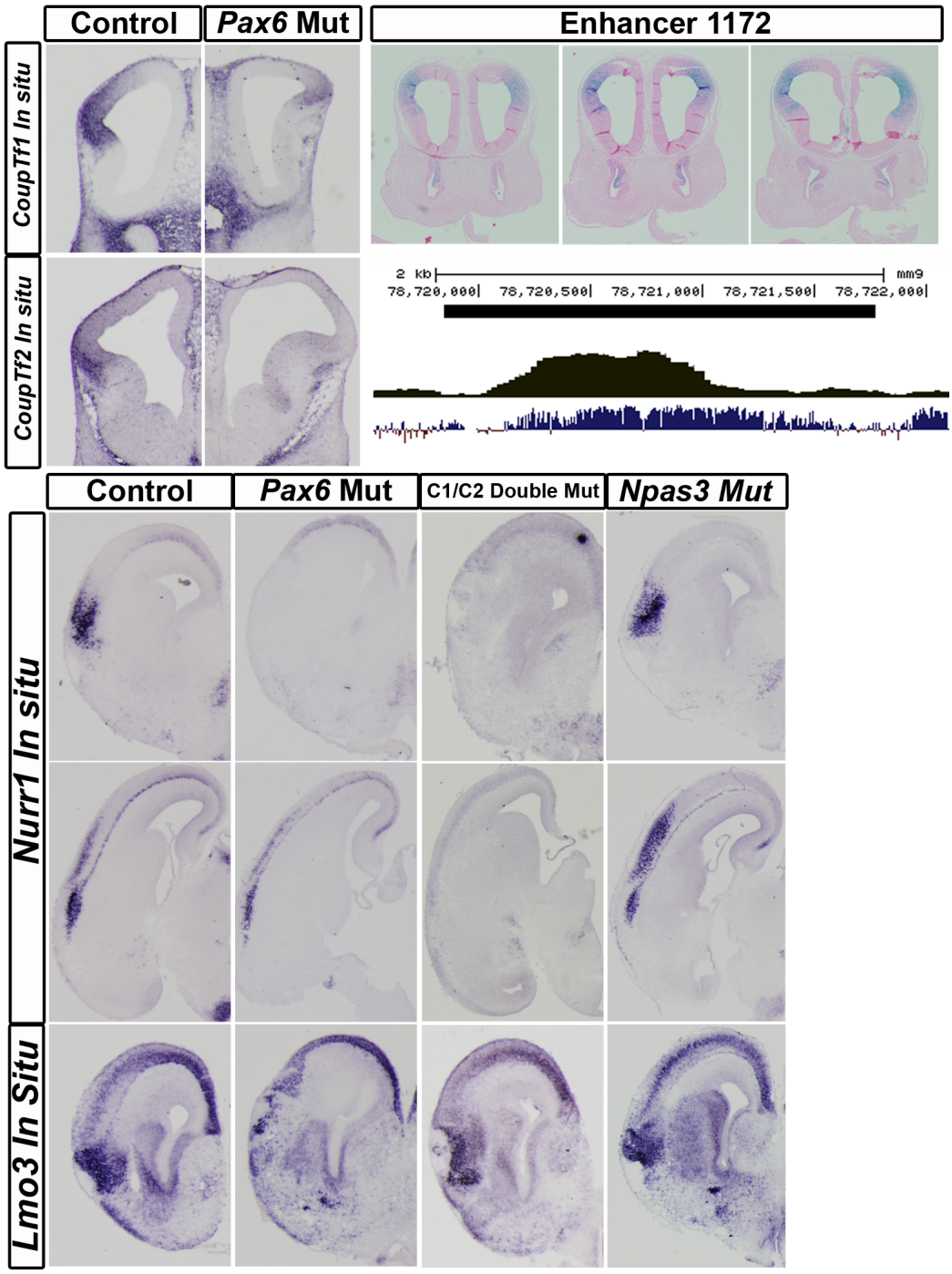

Control
Pax6 Mut
Enhancer 1172
CoupTf1 In situ
CoupTf2 In situ
2 kb
mm9
78,720,000
78,720,500
78,721,000
78,721,500
78,722,000
Control
Pax6 Mut
C1/C2 Double Mut
Npas3 Mut
Nurr1 In situ
Lmo3 In Situ

Figure 3

	E11.5	E13.5
# of Transcription Factors (TFs) analyzed	722	719
# of TFs expressed	250	263
# of TFs with rostrocaudal gradient	58	24
# of TFs with caudorostral gradient	44	73
# of TFs restricted to the medial pallium	11	8
# of TFs restricted to the dorsal pallium	2	0
# of TFs restricted to the lateral pallium	1	0

Table 1

Chapter 3: Transcriptional Regulation of Enhancers Active in Protodomains of the Developing Cerebral Cortex

Introduction

At the core of cortical development lie transcriptional programs that orchestrate a sequence of processes beginning with specification of the cortical anlage and its regional subdivisions, or the protomap (Rakic, 2009; O'Leary et al., 2013). Ongoing work has identified a set of transcription factors (TFs) that control the size and areal identities of pallial subdivisions. These include *CoupTFI, Dmrta2 (Dmrt5), Emx2, Lef1, Lhx2, Pax6,* and *Sp8* (Bishop et al., 2000; Galceran et al., 2000; Yun et al., 2000; Mallamaci and Stoykova, 2006; Armentano et al., 2007; Sahara et al., 2007; Faedo et al., 2008; Mangale et al., 2008; Chou et al., 2009; Konno et al., 2012; Borello et al., 2013; Saulnier et al., 2013). Each of these TFs is expressed in distinct gradients in progenitor cells of the pallial ventricular zone (VZ). For instance, *Pax6* is expressed in rostrocaudal and ventrodorsal gradients; *Pax6* loss-of-function in mice results in a respecification of cortical regions along both its rostrocaudal and ventrodorsal axes (Bishop et al., 2000; Yun et al., 2001). Despite the subdivision of the pallium into discrete structural/molecular units [e.g., the medial, dorsal, lateral and ventral pallium (MP, DP, LP and VP); Puelles et al., 2000], to date the TFs that are known control regional fate are expressed in gradients across these subdivisions, raising the intriguing question of how these gradients are interpreted in an integrative fashion to generate sharply delineated pallial subdivisions and later adult cortical regions.

One mechanism that could solve this conundrum would be that enhancer elements integrate TF expression to generate gene activation in distinct pallial subdivisions, much in the way that regional fate is generated in the cellular blastoderm of *Drosophila* embryos (Lagha et al., 2012). While this general paradigm had previously been supported through anecdotal reports of

individual pallial enhancers identified in gene-centric studies (Kammandel et al., 1999; Theil et al., 2002; van den Bout et al., 2002; Ahituv et al., 2007; Colasante et al., 2008), a recent more comprehensive screen for forebrain enhancers that includes spatial activity data for ~145 human enhancers that are active in the E11.5 mouse telencephalon enables a rigorous and systematic search for enhancers involved in pre-patterning of the pallium (Visel et al., 2013). Here we present evidence that enhancers integrate information from TF gradients in the embryonic day (E) 11.5 mouse pallium to generate distinct expression domains. Using a panel of 14 human enhancers carefully selected based on their *in vivo* activity patterns, we generated a set of stable mouse transgenic lines that express $CreER^{T2}$ and GFP in distinct domains within the developing pallium. Leveraging this unique set of reporter mice, we derived fate maps that elucidate the embryonic origin of pallial subdivisions. Furthermore, we used a combination of bioinformatics, ChIP-seq and *in vivo* studies to elucidate the regulation of these enhancers by major pallial transcription factors including COUPTFI, PAX6 and PBX1. Overall, we propose that the enhancers defined through this study identify protodomains of the pallial neuroepithelium, which may be fundamental units of cortical development and evolution.

Results

Pallial Protodomains Identified by Enhancer Activity Using Transient Transgenic Assay

To define enhancers potentially marking neuroepithelial subdivisions in the E11.5 pallium, we mined a previously described large collection of enhancers active in the developing telencephalon, assayed using transient transgenic mouse *LacZ* expression (Visel et al., 2013). We identified more than 40 enhancers that showed regional pallial expression, many of which showed intrapallial boundaries (Figure 1A-C and Figure S1). For instance, in the MP, several enhancer lines showed nested patterns of expression, varying between a small dorsocaudal

domain (643), a domain in the ventral caudomedial telencephalon (653), a larger domain that includes the entire caudomedial telencephalon (192), and the entire dorsomedial and caudomedial region including the primordial septum (348) (Figure 1C). Regional patterns of activity were also observed for enhancers expressed in the DP, LP and VP (Figures 1A-B). We mapped these expression limits onto a model schema of the E11.5 pallial neuroepithelium, from which we hypothesize the existence of a set of sharply delimited pallial progenitor domains or protodomains (A-I) (Figure 1D; Table S1).

Enhancer Activity of Pallial Enhancer CreER^{T2}-IRES-GFP Alleles

To test the idea that these human enhancers are active in protodomains that generate distinct pallial subdivisions, we produced stable transgenic mouse lines to characterize the properties of 14 enhancers that reproducibly exhibited boundaries in the E11.5 pallium (Figure 1A-C and Figure S1; asterisks label the enhancers used to make stable lines).

We generated stable transgenic mouse lines that express CreER^{T2}-IRES-green fluorescent protein (GFP) and downstream of each one of the 14 selected "pallial" enhancers and a minimal Hsp68 promoter. We generated 2-3 founders for 10/14 of the lines; their expression domains were reproducible (Table S2). We further analyzed the properties of one founder for each enhancer.

To characterize the activity of each enhancer, we defined the GFP expression at E11.5, and compared the enhancer activity in the stable and transient transgenic assays. The stable lines showed enhancer activity patterns that closely resembled the transient transgenic assay (Table S2). We annotated the E11.5 expression domains on a flattened topologic representation of the embryonic pallium (right hemisphere), where stippled grey color indicates GFP expression (Figure 2I and I' and Figure S2A-N). For instance, for enhancer 643, we observed progenitor GFP expression in the MP at E11.5 (Figure 2A-H). On the other hand, enhancer 1050 showed

progenitor GFP expression in the DP and MP at E11.5, but was absent in the ventrolateral pallium (VLP) (Figure 2A'-H').

Next, we examined prenatal GFP expression, at E12.5, E14.5 and E17.5 for all of the lines (Figure S2A-N and Table S2). In most cases, enhancer activity was strongest at E11.5, and was largely unchanged at E12.5 (Table S2). However, activity patterns of some of the enhancers were more dynamic. For instance, 636 was selectively active in the VLP at E10.5, but by E11.5, its activity was greatly reduced (Figure S2E). Activity of 12/14 enhancers decreased and/or became restricted to a smaller domain by E14.5 and E17.5 (Table S2). For instance, 218, 281, 653 and 1318 activity was no longer detected in the pallium by E14.5. Three of the enhancers with MP expression (348, 643 and 1006) were no longer active in the hippocampus, but maintained activity in the hippocampal fissure, choroid plexus and fimbrial area. The activity of 636, 840 and 1172 became restricted to small populations of cells in the pallium at E17.5 (Figure S2E, I, and M). Enhancer 660, which was active in the caudoventral MP at E11.5, became active in the SVZ and superficial cortical layers of the DP at E17.5 (Figure S2H).

Fate Mapping Using Pallial Enhancer CreER^{T2}-IRES-GFP Alleles

To determine the identity of the cells whose progenitors have E11.5 enhancer activity, we performed fate map analyses by introducing the *Ai14 (tdTomato)* Cre reporter allele (Madisen et al., 2010) into the enhancer CreER^{T2}-IRES-GFP lines. We administered tamoxifen at E10.5 to induce CreER^{T2} translocation to the nucleus, where it activated tdTomato expression, and then performed neuroanatomical analyses at later stages. Because of the ~24-36 hour window of tamoxifen action (Hayashi and McMahon, 2002), we assessed enhancer activity at both E11.5 and E12.5 to better interpret the results of E10.5 tamoxifen treatment (Figure S2 and Table S2). Since prenatal tamoxifen treatment frequently led to fetal death around the time of delivery, we obtained fate-mapping data at E17.5 for all enhancer lines. However, we also obtained

postnatal fate maps (P30) for a subset of the enhancers (192, 348, 636, 643, 653, and 660; Figure S2 and Table S3). We chose these enhancers because of their activity in the hippocampus; the hippocampus matures later than the neocortex; thus P30 data helped analysis of the hippocampal fate map.

We annotated the fate map domains on a flattened topological representation of the maturing/mature pallium (Figure 2S-S' and Figure S2). Here we indicated anatomical locations containing tdTomato^{+} cells using a graded rating scale of 1-4: 1 (red) high density to 4 (green) almost no tdTomato^{+} cells (Figure 2S and 2S'). For instance, 643, which showed E11.5 activity restricted to the MP, fate mapped to the rostrodorsal CA fields, dentate gyrus of the rostrodorsal hippocampus, the fimbrial area and choroid plexus (Figure 2J-R and 5B-B'''). On the other hand, 1050, which showed E11.5 activity restricted to DP and MP, was fate mapped to the neocortex and hippocampus, and only weakly labeled the LP (insular cortex) and did not label the VP (piriform cortex) (Figure 2J'-R').

Similar analyses were performed for all the enhancer lines; the data and analyses are compiled in Figures S2A-N. From these experiments, we have deciphered the embryonic origin of pallial subdivisions (see schema of pallial progenitor subdivisions in Figure 1E and Table S3 and S4); we have organized these data into Figures 3, 4 and 5 that focus on the frontal cortex, ventrolateral cortex and hippocampal structures, respectively.

Enhancers Active in Primordia of Distinct Frontal Cortex Subdivisions

The analysis of E11.5 expression and CreER^{T2} fate mapping experiments from 11 enhancer transgenic lines demonstrated which progenitor domains generated cells that populated different subdivisions of the prefrontal cortex (PFC) (Figures 3, S2; Tables S3 and S4). Fate mapping of enhancer lines (192, 348, 1056) with E11.5 activity in the rostral-most E11.5 MP, resulted in labeled cells in the medial PFC (MPFC). 192 activity generated tissue probably

representing the indusium griseum and taenia tecta, (Figures 3A-A''' and Figure S2A); 1056 generated the ventromedial PFC (including the medial orbital cortex) (Figures 3B-B''' and Figure S2L); 348 generated most regions of the MPFC (Figures 3C-C''' and Figure S2D).

Fate mapping of enhancer lines (218, 840 and 1050) with E11.5 activity in the rostral-most DP (Figures 3, S2 and Table S4), and resulted in labeling of the dorsal PFC (DPFC). For instance, 1050 only generated cells in the DPFC (Figures 3D-D''' and S2K).

Fate mapping of enhancer lines (281, 636, and 1172) that showed expression in the rostral-most E11.5 VLP (Figures 3 and S2), resulted in cells of the lateral PFC (LPFC). This included the anterior insular cortex, and the lateral orbital PFC (Figures 3 and S2). Finally, fate mapping of 1318 and 1006 activities, which were similar in the rostral telencephalic pole at E11.5, resulted in cells that populate the entire PFC (Figures 3I, S2J and N). These data are summarized in Figures 3I, S2, and Table S3 and S4.

Enhancers Active in Primordia of Distinct VP and LP Subdivisions

The analysis of E11.5 expression and CreER^{T2} fate mapping experiments from 8 enhancer transgenic lines identified progenitor domains generating the VP and LP, which contain cortical domains superficial to pallial nuclei (Puelles, 2014) (Figures 4K and Figure S2). To systematize the E17.5 fate mapping analyses, we compared the td-Tomato expression with the expression of two proteins that have boundaries in the VP and LP domains: NURR1 (NR4A2) and CTIP2 (BCL11B) (Figure 4). NURR1 was expressed dorsally in the claustrum, a nucleus lying deep to the insular cortex (LP) and more ventrally in the dorsal endopiriform nucleus, which is deep to the piriform cortex (VP). We defined boundary 1 as the dorsal limit of the claustrum, and

boundary 3 as the ventral limit of the endopiriform nuclei (Figures 4B-B’’’, E-E’’’, and H-H’’’). CTIP2 was expressed in the superficial corticoid strata of these two pallial regions; we defined boundary 2 as the limit between the insular cortex and the piriform cortex, and boundary 4 as the ventral limit of the VP with the subpallium (Figures 4C-C’’’, F-F’’’, and I-I’’’).

Fate maps from the ventrolateral enhancers (Figures 4 and S2) showed different tdTomato^{+} cell distributions in the ventrolateral cortices. Rostrally, enhancers 1050, 1006, 218, 281 and 636 showed progressively more ventral boundaries. Cells marked by 1050 activity were restricted to the DP (ending before the LP), 1006 ended roughly at the DP/LP boundary, 218 ended roughly at the LP/VP boundary, and 281 and 636 extended to the pallial-subpallial boundary (Figures 4A-F’’’ and Figure S2C, E and J). Enhancer 636 was most active in the VLP, with little activity in the DP (Figures 4A -C’’’, and Figure S2E).

Figure 4J shows enhancer fate map annotation along the dorsoventral axis in separate rostral, middle and caudal regions. Some enhancers had clear rostrocaudal differences in the dorsoventral position of their respective fate maps (281, 636, 1172 and 1318). For instance, 1172 labeled a domain that rostrally was largely ventrally restricted, whereas caudally it extended into the DP and MP (Figure S2M). Note that enhancer 1172 maps to a genomic region ~100 kb away from *CoupTFI*, which shows a very similar expression pattern (Figure S2M, Table S2; Armentano, 2007; Faedo et al., 2008).

Enhancers Active in Primordia of MP Subdivisions: Hippocampal Complex and Adjacent Structures

E11.5 expression and fate mapping experiments from 6 enhancer transgenic lines demonstrate the progenitor domains that generate different MP derivatives; these enhancers were active either in the rostrodorsal or caudoventral hippocampal fields (Figures 5, 6F-G, and Figure S2). Note that the dorso-ventral adult hippocampal topography corresponds topologically to the

embryonic rostro-caudal axis, the ventral tip next to the amygdala being caudalmost. The hippocampal region is topologically dorsal; the choroid plexus (Ch) is the dorsal-most component.

CreER^{T2} fate mapping from enhancers 192, 218, 348 and 643 generated a nested pattern of derivatives within the hippocampal complex. 192 activity was the most restricted; its derivatives contributed to the choroid plexus, fimbrial area (F or hem), and hem-originated Cajal Retzius (CR) cells, with very sparse labeling of the dentate gyrus and CA fields (Figures 5C-C’’’, and Figure S2A). 643 was active in the progenitors of the Ch, F, CR cells, dentate gyrus (DG), and CA fields (strongest in CA1) (Figure 5B-B’’’, and Figure S2F). Its activity was restricted to the rostrodorsal MP. Likewise, 218 was active in progenitors of the rostrodorsal MP, but was weak in caudoventral MP progenitors (Figure S2B). 348 was active in the entire MP, but with stronger activity in its rostrodorsal components (Figures 5A-A’’’, D-D’’, and Figure S2D).

By contrast with these rostrodorsal MP enhancers, we identified two MP enhancers that were almost exclusively restricted to the caudoventral MP: 653 and 660. These were active in progenitors that produced cells in the caudoventral DG and CA fields (Figures 5E-E’’’, and Figures S2G and H).

Enhancer 643 Marks the Formation of the Hippocampal Field.

Based on the nested activity of enhancers 192, 643 and 348 within the hippocampal primordia (GFP expression and fate maps) we studied its ontogenesis in detail by examining enhancer MP activity at E11.5 (Figure 6). We compared the expression of GFP and *Lmx1a* RNA, a marker of the F and Ch (Chizhikov et al., 2010). Histochemical analysis at E11.5 showed that GFP expression from enhancers 192 and 643 and *Lmx1a* RNA expression where nearly identical; they shared a sharp boundary (arrowhead; Figure 6A-A’’’, and B-B’’’). Likewise at E12.5 enhancer 192 activity (GFP) and *Lmx1a* RNA expression remained nearly identical

(arrowhead; Figure 6C-C''), whereas enhancer 643 GFP expression spread into the adjacent pallial neuroepithelium (between arrowhead and arrow; Figure 6D-D''). Recall that enhancer 192 fate mapping labeled very little of the hippocampus, whereas enhancer 643 fate mapping labeled the DG and the CA fields of the rostrodorsal hippocampus (Figures 5C-C''', and arrows; 6C). Thus, the hippocampal field is first detectable between E11.5 and E12.5, concomitant with the expansion of enhancer 643 activity (Figure 6E).

Computational Identification of Transcriptional Drivers of Region-Specific Enhancer Activity

To explore the molecular mechanisms controlling enhancer activity in subregions of the pallium, we compared the sequences of enhancers with activity largely restricted specific pallial domains. We were most successful when we compared MP enhancer sequences (N=9; 192, 348, 480, 611, 622, 643, 653, 660, 1006), to the sequence of enhancers active in DP, LP and VP (N=15; 22, 200, 218, 488, 595, 619, 632, 636, 671, 876, 957, 978, 987, 1025, 1050). We searched for nucleotide motifs that distinguished these groups using two models. Model 1 was trained to distinguish sequence motifs between 9 MP enhancers, 15 non-MP enhancers, and 480 random genomic sequences. Model 2 was trained to distinguish sequence motifs between 9 MP enhancers, 15 non-MP enhancers, a set of background sequences consisting of 480 random genomic sequences and 765 sequences from the VISTA Enhancer Browser that were negative for enhancer activity (see Extended Experimental Methods). This approach generated motifs that were enriched in MP enhancers compared to non-MP enhancers. The top twenty *de novo* motifs for each of sets of enhancers (MP and non-MP) were mapped to the Transfac and JASPAR database to identify TFs that have similar binding sites (Figure S3; see Table S5 and S6).

Using the list of TF binding motifs preferentially identified in MP enhancers or in non-MP enhancers, we scrutinized the E11.5 expression of the top 40 TFs using the Allen Brain Atlas (http://developingmouse.brain-map.org/). Five TFs [*Lhx5, AR (Nr4a2,* nuclear receptor family*), Lmx1a* and *Foxj1*] were only expressed in the Ch/F domain, and 5 TFs showed expression in the DP, but not in the MP, and especially Ch/F, expression was either low or not detectable (Figures S3, S4 and Table S5 and S6). Thus the method successfully selected for TFs that were either expressed within, or excluded from Ch/F. Of note, the Ch is perhaps the most distinct region of the pallium, because its derivative, the choroid plexus is a non-neural structure.

Transcriptional Mechanisms Regulating Enhancer Function: *in vivo* binding by PAX6, COUPTFI and PBX1.

Next, we directly screened the enhancers for binding sites for TFs known to regulate pallial pattering. We found binding sites for PAX6, COUPTFI (NR2F1), and PBX1 in pallial enhancers (Genomatix); each of these TFs regulates patterning of the pallial primordium (Bishop et al., 2000; Yun et al., 2000; Mallamaci and Stoykova, 2006; Armentano et al., 2007; Faedo et al., 2008; Golonzkha and Rubenstein, unpublished). We then tested whether these enhancers were bound *in vivo* by these TFs using chromatin immunoprecipitation and DNA sequencing (ChIP-Seq). We performed ChIP-Seq with antibodies to PAX6 (n=3), COUPTFI (n=1), and PBX1 (n=1) on dissected E12.5 mouse pallium; information about the quality of the sequence mapping and peak calling are reported in Table S7. We surveyed the genome for binding to the 44 enhancers assayed in Figure 1 and Figure S1. The results are organized according to the regional activity of the enhancers, MP (n=11), DP+MP (n=8), DP (n=2); LVP+DP; LVP (n=12), MP+DP+LVP (n=6) (Table S8). Then we annotated ChIP-Seq binding to each enhancer by PAX6, COUPTFI, and PBX1.

PAX6 bound to all of the enhancers that were globally expressed (MP+DP+LVP) in the pallium (6/6). As the enhancers became more restricted in their regional activity, PAX6 binding frequency reduced, particularly if LVP activity was absent. PAX6 and COUPTFI (NR2F1) bound to few MP enhancers, and PBX1 bound none (Figure 7A and Table S8). We show an example of PAX6 peaks over enhancer 840 and 636 (Figure 7C), and PAX6, COUPTFI and PBX1 binding over the other enhancers are shown in Figure S5.

We then focused on *Pax6* regulation of some of the enhancers that had *in vivo* binding sites. We used transient transfection luciferase assays to study whether *Pax6* co-transfection modulated activity of 636, 643, 840 and 1172 (n=4). In each case, we observed >5-fold activation of luciferase expression (Figure 7B). Of note, PAX6 activation declined in enhancers with MP activity (840, 643), or that were expressed in a caudorostral gradient (1172; note: *Pax6* is expressed in a rostocaudal gradient).

Finally, we tested *Pax6 in vivo* regulation of enhancer activity by introducing the 636, 643, and 840 *enhancer-CreER^{T2}-GFP* alleles into mice harboring a *Pax6* null allele (*Sey*). We generated E11.5 embryos, and found that *Pax6$^{-/-}$* mutants had reduced GFP expression from enhancers 636 and 840 in pallium (Figure 7D,D' and F,F'). On the other hand enhancer 643 continued to express GFP in the *Pax6$^{-/-}$* mutant, although the ventral boundary was less sharp (Figure S5).

Discussion

We generated stable transgenic mouse lines that express CreER^{T2} and GFP from 14 different enhancer elements with activity in distinct domains within the E11.5 pallium. These enhancer-CreER^{T2}-GFP lines have obvious broad utility for experimental manipulation of gene expression in specific domains and at specific times, including Cre-mediated gene deletion. Herein, using these unique tools, we: 1) determined the first comprehensive regional fate map of the mouse

pallium, that includes evidence for a set of progenitor domains defined by the activity of the enhancers (Figure 1E); 2) began to decipher transcriptional mechanisms that control the enhancers using: informatics, *in vivo* occupancy by TFs that regulate cortical patterning (PAX6, COUPTFI and PBX1), and analysis of enhancer activity in *Pax6* mutants. Below we elaborate on these discoveries and their implications for understanding cortical development, evolution and disorders.

Dynamic Temporal Activity of Cortical Enhancers

Previously published transient transgenic analysis of cortical enhancers (Visel et al, 2013) interrogated only one developmental time point, E11.5. Using the stable enhancer lines described in this paper, we analyzed enhancer activity at different developmental ages. The majority of the enhancers maintained similar patterns of activity between E11.5 to E13.5; however, by E14.5, the activity of most of the enhancers decreased and/or became restricted to a smaller domain (Figure S2; Table S2). . Thus, the set of enhancers we studied were primarily active during stages when regional patterning of pallium takes place (~E9.5-E12.5), and at early stages of neurogenesis (~E11-E13.5), suggesting that other enhancers have roles at later stages to drive gene expression for later developmental processes. Recently, Nord et al., 2013, provided evidence for distinct cohorts of enhancers that are active at different stages of brain development. In addition, some enhancers can be active at different stages, For example 660, as its E11.5-E13.5 activity in the caudoventral cortex wanes, expression begins at~E13.5 in the neocortical subventricular zone (Figure S2). The *DlxI12b* enhancer is active both in subpallial subventricular zone progenitor cells, and in maturing and mature GABA neurons (Ghanen et al., 2007; Vogt et al., in press).

Fate Mapping Analyses Define the Regional Derivatives From Distinct Pallial Progenitor Domains.

The transient transgenic analysis of enhancer E11.5 activity led us to hypothesize subdivisions of the pallial progenitors (Figure 1D). The stable transgenic analysis of E11.5 enhancer activity (GFP expression), and CreER^{T2} fate analyses at E17.5 and P30, supported many aspects of our initial model (Figure 1D), and importantly enabled us to describe the regional fates of each proposed progenitor domain (Figure 1E and Table S2). Among the important observations, we discovered E11.5 progenitor domains that produce distinct subdivisions of the frontral cortex, providing the first information about where these distinct regions originate (Figure 3).

Previous fate mapping of pallial regions have used transplantation (chick-quail; Garcia-Lopez et al., 2009) and Cre recombination methods, in which a constitutive Cre was driven from a gene locus. Thus, unlike our study, the previous Cre fate mapping did not obtain temporal-specific data, since generally the alleles were active over long periods of pallial development. Emx1-Cre and Foxg1 fate mapping showed that expression from these loci covers most of the pallium (Gorski et al., 2002; Hébert and McConnell, 2000). Wnt3a-Cre fate mapping labeled the cortical hem and derived Cajal-Retzius cells (Yoshida et al., 2006). Dbx1-Cre fate mapping showed that the ventral pallium is another source for Cajal Retzius cells, and that it contributes glutamatergic neurons to specific nuclei in the amygdala and ventral cortical structures (Bielle et al., 2005; Hirata et al., 2009; Teissier et al., 2010; Waclaw et al. 2010). Thus, while these studies provide important information about the fates of pallial regions that express Cre over the course development, they do not provide a comprehensive fate map from multiple pallial progenitor domains from temporally-restricted Cre activity.

The fate maps obtained using the 14 enhancer lines illuminated unexpected facets of the E11.5 expression domains. There was a rostrocaudal discontinuity in the properties of dorsomedial progenitor domains (between coronal planes 4 and 5 of schema, Figure 1E); rostrally, next to the septocommissural region, the dorsomedial progenitors generate the motor, cingulate and prefrontral cortex; caudally, next to the choroid plexus, the dorsomedial progenitors generate

the hippocampal complex and fimbrial area, (hem). Within the rostral domain we observed other rostrocaudal discontinuities, such as the restriction of the pallial septum, IL, and PrL domains within coronal planes 2 and 3 (Figure 3J).

Topographic discontinuities in caudal progenitor domains include the restriction of the Ch and F to the regions illustrated in coronal planes 4-7, the joining of the caudoventral (caudal) and rostrodorsal hippocampal domains in coronal plane 8, and the end of the hippocampal domain in coronal plane 9. Remarkably, we identified enhancers only active in the rostrodorsal (218, 348 and 643) or caudoventral (653 and 660) hippocampal primordia at E11.5.

The relative sizes of some progenitor domains were disproportionate to the size of their derived regions at E17.5 and P30, such as the size of Ch/F compared to the rest of the hippocampal complex (see sections 5-7; Figures 1E, and 6G). Thus there was not a 1:1 proportional matching of the sizes of E11.5 progenitor and mature domains, providing evidence that the timing and relative distribution of regional growth is not uniform.

We provided evidence that the hippocampal primordium begins to expand at E12.5, based on the likewise expanded activity of enhancer 643 (Figure 6). Furthermore, the activity of enhancer 1050 becomes progressively focused in the hippocampal region between E11.5-E14.5 (Figure S2K). Thus, further investigations into the transcriptional process that drives hippocampal development will be aided by understanding the transcription mechanisms that drive enhancer 643 and enhancer 1050 activities in the hippocampal primordium.

While the aim of our anatomical analyses was to derive a pallial fate map, we made some observations about regional histogenesis and cell type generation. Tamoxifen induction of recombination at E10.5 generally resulted in radial clones of cells that spanned the cortical plate, providing evidence that this set of enhancers is not lineage-restricted with regard to subsequent laminar fate. This adds evidence for the model that intrinsically produced neurons

for each cortical layer are sequentially generated from the same neuroepithelial progenitor (Leone et al., 2008; Guo et al., 2013), although it does not eliminate the possibility that enhancers will be discovered that show more restricted fate properties. Indeed, as has already been elucidated, excitatory neurons of layer 1 (Cajal Retzius cells) are generated from specific domains at the pallial perimeter (Bielle et al., 2005; Yoshida et al, 2006; Puelles, 2011); several of our enhancers (192, 348 and 643) provide additional evidence for this process (Figure S2A, D, and F). Furthermore, because the enhancers drive CreER^{T2}, tamoxifen induction of Cre activity at later time points can be used to study later stages of neuro- and gliogenesis with enhancer lines that maintain progenitor cell activity after E11.5 (Table S2).

Identification of Enhancers That Detect Pallial Subdivisions: Insights into the Transcription Networks Driving Pallial Regional Development and Evolution.

In the cellular blastoderm of *Drosophila*, enhancer activities reveal developmental domains generated by the combinatorial activity of TFs that ultimately underlie body subdivisions, as exemplified by enhancers that drive gap-gene expression (Perry et al., 2011). Enhancer activity domains can be smaller and sharper than the expression domains of the TFs that drive their expression (Perry et al., 2011).

In the pallial primordium *CoupTFI, Dmrta2 (Dmrt5), Emx2, Lhx2, Pax6, Pbx1 and Sp8,* TFs that control pallial regionalization, are expressed in broad gradients (Bishop et al., 2000; Galceran et al., 2000; Yun et al., 2000; Mallamaci and Stoykova, 2006; Armentano et al., 2007; Sahara et al., 2007; Faedo et al., 2008; Mangale et al., 2008; Chou et al., 2009; Konno et al., 2012; Borello et al., 2013; Saulnier et al., 2013). Positional information appears to lie in these TF gradients, and in their combinatorial interactions.

Here, we provide the first evidence for a mechanism that can integrate this transcriptional information to generate discrete pallial subdivisions. Many of the enhancers show patterns of

activity at E11.5 that are more discrete than the broadly expressed patterns of aforementioned patterning TFs. Thus, we suggest these enhancer activities reflect the integration of transcriptional activities that together pattern the pallium. Furthermore, it is possible that these and related enhancers are fundamental elements that have driven pallial evolution, as duplication and transposition of these distant-acting regulators have the potential to alter gene expression. Significantly, our gain-of-function transgenic assays show the ability of these enhancers to function in a variety of chromosomal locations.

Currently, we do not have definitive evidence for the gene(s) that each of these enhancers regulates. However, based on proximity, and similar expression profiles, we have some predictions for enhancer/gene pairs (Figure S2 and Table S2). For instance, the activity of enhancers 1006, 1050 and 1172, that have genomic positions close to *Wnt8b, Lef1 and CoupTFI* respectively, closely resembles the pallial expression of these genes (Figure S2J, K, and M; Visel et al., 2013). Future studies are required to test for enhancer/gene interactions using chromatin conformation methods (Clowney et al., 2012), as well as loss-of-function mutagenesis. While some enhancers are clearly required for gene expression (Shim et al., 2012), there is evidence that enhancer redundancy exists (Ahituv et al., 2007; Lagha et al., 2012).

Mechanisms That Regulate Enhancer Activity

We used informatic, biochemical and, genetic approaches to begin deciphering transcriptional mechanisms that control the activity of the enhancers. Informatic methods provide insights into candidate TFs that may regulate enhancer activities (Shim et al., 2012; Visel et al., 2013). We used a machine learning method to identify nucleotide signatures that may underlie regional differences in enhancer activities. Transcriptional binding sites that were enriched in enhancers with and without MP activity led us to identify TFs with expression either in, or excluded from,

the Ch/F part of the E11.5 MP (Figure S3 and S4). This is interesting because histogenesis of the primordium of the choroid plexus (Ch) is distinct from the rest of the pallium, as the Ch is a non-neural tissue generated from the neural tube roof plate (Puelles et al., 2014), and the fimbrial area (cortical hem) represents the border between roof and alar plate tissues. As more pallial enhancers are defined, and as the binding sites for additional TFs are identified, it is likely that informatic approaches will gain power in defining sequences that control regional expression.

Next, we used ChIP-Seq to test whether the enhancers under study were bound *in vivo* in the E12.5 mouse cortex by PAX6, COUPTFI and PBX1, three transcription factors that regulate pallial patterning (Bishop et al., 2000; Yun et al., 2000; Mallamaci and Stoykova, 2006; Armentano et al., 2007; Sahara et al., 2007; Faedo et al., 2008; Borello et al., 2013; Golonzhka and Rubenstein, unpublished). Previous analyses of PAX6 binding in the developing pallium used ChIP-promoterChIP, and thus did not examine PAX6 binding to the enhancers described herein (Sansom et al., 2009; Xie et al., 2013). Our ChIP-Seq analyses, which we will publish later in their entirety, showed that PAX6 (a general marker for the telencephalic pallium; Puelles et al., 2000) bound to all of the enhancers globally expressed in the pallium (Figure 7A and Table S8), suggesting that PAX6 may have a fundamental role coordinating pallial properties.

Enhancers with more restricted intrapallial regional activity had reduced frequencies of PAX6 binding, particularly when LVP activity was absent (Figure 7A and Table S8). In addition, COUPTFI and PBX1 bound few MP enhancers (Figure 7A and Table S8). Both *Pax6* and *CoupTFI* regulate dorsoventral patterning of the pallium; and they both promote ventral identity (Yun et al., 2000; Faedo et al., 2008). Consequently, our ChIP-Seq analysis provides evidence that these two TFs may regulate dorsoventral pallial patterning by promoting activity of enhancers specifically active in the VP, LP and DP.

This concept is consistent with that the fact that PAX6 has a potent role in patterning the ventrolateral cortex (Figure 7A; Yun et al., 2001). Furthermore, transcription assays in tissue culture showed that PAX6 strongly activated (~15-fold) 636, the enhancer with VP and LP activity (Figure 7B). On the other hand, PAX6-mediated activation was lower for enhancers with MP activity (840, 643), suggesting that they have elements that antagonize activation by PAX6.

To test hypotheses generated by the ChIP-Seq and transfection assays, we examined 636 and 840 enhancer activities in *Pax6* loss-of-function mutants. As predicted by their PAX6 ChIP-Seq peaks and their activation by PAX6 in the cell culture transcription assay, 636 and 840 pallial activity were greatly reduced in E11.5 *Pax6*$^{-/-}$ (Figure 7D,D' and 7F,F').

Implications

The identification of human enhancers with restricted spatial and temporal activities in pallial protodomains demonstrates that the genome has relatively small (0.5-3 kb) regulatory elements that can integrate transcriptional information to generate highly specific patterns of gene expression, even in ectopic genomic loci (herein, and Visel et al., 2013). Importantly, the enhancer activity patterns for the most part don't resemble the expression of single known TFs, highlighting the enhancers' roles as spatial integrators of regulatory information. This knowledge opens the door to deciphering the sequence-specific regulation of enhancer activity and how mutations alter their function and contribute to disease.

Experimental Methods

Generation and Characterization of Stable Enhancer Transgenic Mice

PCR amplified human enhancer regions were subcloned into *Hsp68-CreERT2-IRES-GFP* (Visel et al., 2013), and used to generate stable transgenic mice. Founders were screened using CreERT2 specific primers. Enhancer transgenic embryos were examined for GFP expression. For fate mapping, enhancer lines were crossed to *Ai14* Cre-reporter mice (Madisen et al., 2010). Tamoxifen was administered at E9.5 or E10.5; tdtomato was assayed at later stages. Stable transgenic mice were crossed to the *Pax6* mutant. Mice were used in accordance with National Institutes of Health and UCSF guidelines.

Histology

Immunohistochemistry was performed as in Flandin, 2010. RNA *in situ* hybridization and *in situ*/immunohistochemistry was performed as in Jeong et al., 2008.

Identification of Region Specific Motifs (see Extended Experimental Methods).

Chromatin Immunoprecipitation (ChIP)

ChIP was performed using E12.5 or E13.5 cortex and Pax6 (Millipore), CoupTf1 (R&D systems), and Pbx1/2/3 (C-20, Santa Cruz Biotechnology) antibodies (McKenna et al., 2011). Libraries were prepared using an Ovation Ultralow DR Multiplex System (Nugen) Reads from ChIP, input, and negative control (IgG) libraries were mapped to the mouse genome (mm9) using BWA and peaks were called using MACS considering both input and IgG as the control sample with filtering to remove peaks in repeat regions.

Luciferase Assay

Enhancer activity was studied in P19 cells (Farah et al., 2000) co-transfected with *pCAGGs* (empty) or *pCAGGs-Pax6/CoupTf1/Pbx1*, and Promega *pGL4.23* luciferase reporter (empty) or containing an enhancer element upstream of the luciferase gene (pGL4.23-enhancer).

ACKNOWLEDGEMENTS

Grant Support: JLRR: Allen Institute for Brain Science, Nina Ireland, Weston Havens Foundation, and NINDS R01 NS34661; OG: NARSAD; KP: NIH NIGMS MSTP T32 GMO7618; SS: T32 GM007449; RH: NIH NRSA; LP: Spanish Ministry of Science and Competitivity (BFU2008-01456), Seneca Foundation (04548/GERM/06-10891); LT: Intramural Program of NIH, NLM; ASN: NIH/NIGMS NRSA F32 GM105202; AV and LAP NIH R01NS062859A and R01HG003988, and DOE Contract DE-AC02-05CH11231.

References

Ahituv, N., Zhu, Y., Visel, A., Holt, A., Afzal, V., Pennacchio, L.A., and Rubin, E.M. (2007). Deletion of ultraconserved elements yields viable mice. PLoS Biol. 5, e234.

Armentano, M., Chou, S.J., Tomassy, G.S., Leingärtner, A., O'Leary, D.D., and Studer, M. (2007). COUP-TF1 regulates the balance of cortical patterning between frontal/motor and sensory areas. Nat. Neurosci. 10,1277-1286.

Bielle, F., Griveau, A., Narboux-Nême, N., Vigneau, S., Sigrist, M., Arber, S., Wassef, M., and Pierani, A. (2005). Multiple origins of Cajal-Retzius cells at the borders of the developing pallium. Nat. Neurosci. 8,1002-1012.

Bishop, K.M., Goudreau, G., and O'Leary, D.D. (2000). Regulation of area identity in the mammalian neocortex by Emx2 and Pax6. Science. 288, 344-349.

Borello, U., Madhavan, M., Vilinsky, I., Faedo, A., Pierani, A., Rubenstein, J., and Campbell, K. (2013). Sp8 and COUP-TF1 Reciprocally Regulate Patterning and Fgf Signaling in Cortical Progenitors. Cereb. Cortex.

Chizhikov, V.V., Lindgren, A.G., Mishima, Y., Roberts, R.W., Aldinger, K.A., Miesegaes, G.R., Currle, D.S., Monuki, E.S., and Millen, K.J. (2010). Lmx1a regulates fates and

location of cells originating from the cerebellar rhombic lip and telencephalic cortical hem. Proc. Natl. Acad. Sci. U.S.A. 107, 10725-10730.

Chou, S.J., Perez-Garcia, C.G., Kroll, T.T., and O'Leary, D.D. (2009). Lhx2 specifies regional fate in Emx1 lineage of telencephalic progenitors generating cerebral cortex. Nat. Neurosci. 12, 1381-1389.

Clowney, E.J., LeGros, M.A., Mosley, C.P., Clowney, F.G., Markenskoff-Papadimitriou, E.C., Myllys, M., Barnea, G., Larabell, C.A., and Lomvardas, S. (2012). Nuclear aggregation of olfactory receptor genes governs their monogenic expression. Cell 151, 724-737

Colasante, G., Collombat, P., Raimondi, V., Bonanomi, D., Ferrai, C., Maira, M., Yoshikawa, K., Mansouri, A., Valtorta, F., Rubenstein, J.L., and Broccoli, V. (2008). Arx is a direct target of Dlx2 and thereby contributes to the tangential migration of GABAergic interneurons. J. Neurosci. 28, 10674-10686.

Faedo, A., Tomassy, G.S., Ruan, Y., Teichmann, H., Krauss, S., Pleasure, S.J., Tsai, S.Y., Tsai, M.J., Studer, M., and Rubenstein, J.L. (2008). COUP-TF1 coordinates cortical patterning, neurogenesis, and laminar fate and modulates MAPK/ERK, AKT, and beta-catenin signaling. Cereb. Cortex. 18, 2117-2131.

Farah, M.H., Olson, J.M., Sucic, H.B, Hume, R.I., Tapscott, S.J., and Turner, D.L. (2000). Generation of neurons by transient expression of neural bHLH proteins in mammalian cells. Development. 127, 693-702.

Flandin, P., Kimura, S., and Rubenstein, J.L. (2010). The progenitor zone of the ventral medial ganglionic eminence requires Nkx2-1 to generate most of the globus pallidus but few neocortical interneurons. J. Neurosci. 30, 2812-2823.

Galceran, J., Miyashita-Lin, E.M., Devaney, E., Rubenstein, J.L., and Grosschedl, R. (2000). Hippocampus development and generation of dentate gyrus granule cells is regulated by LEF1. Development.127, 469-482.

Ghanem N, Yu M, Long J, Hatch G, Rubenstein JL, Ekker M. (2007.) Distinct cis-regulatory elements from the Dlx1/Dlx2 locus mark different progenitor cell populations in the ganglionic eminences and different subtypes of adult cortical interneurons. J Neurosci. 27, 5012-5022,

Garcia-Lopez R, Pombero A, Martinez S. (2009). Fate map of the chick embryo neural tube. Dev Growth Differ. 51, 145-165.

Gorski JA, Talley T, Qiu M, Puelles L, Rubenstein JL, Jones KR. (2002). Cortical excitatory neurons and glia, but not GABAergic neurons, are produced in the Emx1-expressing lineage. J Neurosci. 22, 6309-6314.

Guo, C., Eckler, M.J., McKenna, W.L., McKinsey, G.L., Rubenstein, J.L.R., and Chen, B. (2013). Fezf2 expression identifies a multipotent progenitor for neocortical projection neurons, astrocytes and oligodendrocytes. Neuron. 80, 1167-1174

Hayashi, S., and McMahon, A.P. (2002). Efficient recombination in diverse tissues by a tamoxifen-inducible form of Cre: a tool for temporally regulated gene activation/inactivation in the mouse. Dev. Biol. 244, 305-318.

Hébert JM, McConnell SK. (2000). Targeting of cre to the Foxg1 (BF-1) locus mediates loxP recombination in the telencephalon and other developing head structures. Dev Biol. 222(2), 296-306.

Hirata T, Li P, Lanuza GM, Cocas LA, Huntsman MM, Corbin JG. (2009). Identification of distinct telencephalic progenitor pools for neuronal diversity in the amygdala. Nat Neurosci. 12(2), 141-149.

Jeong, J., Li, X., McEvilly, R.J., Rosenfeld, M.G., Lufkin, T., and Rubenstein, J.L.

(2008). Dlx genes pattern mammalian jaw primordium by regulating both lower jaw-specific and upper jaw-specific genetic programs. Development. 135, 2905-2916.

Kammandel, B., Chowdhury, K., Stoykova, A., Aparicio, S., Brenner, S., and Gruss, P.

(1999). Distinct cis-essential modules direct the time-space pattern of the Pax6 gene activity. Dev. Biol. 205, 79-97.

Konno D, Iwashita M, Satoh Y, Momiyama A, Abe T, Kiyonari H, Matsuzaki F. (2012).

The mammalian DM domain transcription factor Dmrta2 is required for early embryonic development of the cerebral cortex. PLoS One. 7(10):e46577.

Lagha, M., Bothma, J.P., and Levine, M. (2012). Mechanisms of transcriptional precision in animal development. Trends Genet. 28, 409-416.

Leone, D.P., Srinivasan, K., Chen, B., Alcamo, E., and McConnell, S.K. (2008). The determination of projection neuron identity in the developing cerebral cortex.

Curr. Opin. Neurobiol. 18, 28-35.

Liu, Q., Dwyer, N.D., and O'Leary, D.D. (2000). Differential expression of COUP-TFI, CHL1, and two novel genes in developing neocortex identified by differential display PCR. J. Neurosci. 20, 7682-7690.

Madisen, L., Zwingman, T.A., Sunkin, S.M., Oh, S.W., Zariwala, H.A., Gu, H., Ng, L.L. Palmiter, R.D., Hawrylycz, M.J., Jones, A.R., Lein, E.S., and Zeng, H. (2010). A robust

and high-throughput Cre reporting and characterization system for the whole mouse brain. Nat. Neurosci. 13,133-140.

Mallamaci, A., and Stoykova, A. (2006). Gene networks controlling early cerebral cortex arealization. Eur. J. Neurosci. 23, 847-856.

Mangale, V.S., Hirokawa, K.E., Satyaki, P.R., Gokulchandran, N., Chikbire, S., Subramanian, L., Shetty, A.S., Martynoga, B., Paul, J., Mai, M.V., Li, Y., Flanagan, L.A., Tole, S., and Monuki, E.S. (2008). Lhx2 selector activity specifies cortical identity and suppresses hippocampal organizer fate. Science. 319, 304-309.

McKenna, W.L., Betancourt, J., Larkin, K.A., Abrams, B., Guo, C., Rubenstein, J.L., and Chen, B. (2011). Tbr1 and Fezf2 regulate alternate corticofugal neuronal identities during neocortical development. J. Neurosci. 31, 549-564.

Nord AS, Blow MJ, Attanasio C, Akiyama JA, Holt A, Hosseini R, Phouanenavong S, Plajzer-Frick I, Shoukry M, Afzal V, Rubenstein JL, Rubin EM, Pennacchio LA, Visel A. (2013). Rapid and pervasive changes in genome-wide enhancer usage during mammalian development. Cell. 155, 1521-1531.

O'Leary, D.D.M., Stocker, A.M., Zembrzycki, A. (2013). Area patterning of the mammalian cortex. in Developmental Neuroscience. Eds. Rubenstein and Rakic. . Academic Press. Vol 1, 61-86.

Perry, M.W., Boettiger, A.N., and Levine, M. (2011). Multiple enhancers ensure precision of gap gene-expression patterns in the Drosophila embryo. Proc. Natl. Acad. Sci. U.S.A. 108,13570-13575.

Puelles, L., Kuwana, E., Bulfone, A., Shimamura, K., Keleher, J., Smiga, S., Puelles, E., and Rubenstein, J.L.R. (2000). Pallial and subpallial derivatives in the embryonic chick and mouse telencephalon, traced by the expression of the Dlx-2, Emx-1, Nkx-2.1, Pax-6 and Tbr-1 genes. Journal of Compartive Neurology, 424, 409-438.

Puelles, L. (2011). Pallio-pallial tangential migrations and growth signaling: new scenario for cortical evolution? Brain Behav Evol., 78(1):108-127.

Puelles, L. (2014). Development and evolution of the claustrum. In The Claustrum: Structural, Functional, and Clinical Neuroscience. John Smythies, Lawrence Edelstein and Vilayanur S. Ramachandran (eds), Academic Press, pp.119-176.

Rakic, P. (2009). Evolution of the neocortex: a perspective from developmental biology. Nat. Rev. Neurosci. 10, 724-735.

Sahara, S., Kawakami, Y., Izpisua Belmonte, J.C., and O'Leary, D.D. (2007). Sp8 exhibits reciprocal induction with Fgf8 but has an opposing effect on anterior-posterior cortical area patterning. Neural Develop. 2,10.

Sansom, S.N., Griffiths, D.S., Faedo, A., Kleinjan, D.J., Ruan, Y., Smith, J., van Heyningen, V., Rubenstein, J.L., and Livesey, F.J. (2009). The level of the transcription factor Pax6 is essential for controlling the balance between neural stem cell self-renewal and neurogenesis. PLoS Genet. 5, e1000511.

Saulnier A, Keruzore M, De Clercq S, Bar I, Moers V, Magnani D, Walcher T, Filippis C, Kricha S, Parlier D, Viviani L, Matson CK, Nakagawa Y, Theil T, Götz M, Mallamaci A, Marine JC, Zarkower D, Bellefroid EJ. (2013). The doublesex homolog Dmrt5 is required for the development of the caudomedial cerebral cortex in mammals

Cereb Cortex. 23, 2552-2567.

Shim, S., Kwan, K.Y., Li, M., Lefebvre, V., and Sestan, N. (2012). Cis-regulatory control of corticospinal system development and evolution. Nature. 486, 74-79.

Teissier A, Griveau A, Vigier L, Piolot T, Borello U, Pierani A. (2010). A novel transient glutamatergic population migrating from the pallial-subpallial boundary contributes to neocortical development. J Neurosci. 30):10563-10574

Theil, T., Aydin, S., Koch, S., Grotewold, L., and Rüther, U. (2002). Wnt and Bmp signalling cooperatively regulate graded Emx2 expression in the dorsal telencephalon. Development. 129, 3045-3054.

van den Bout, C.J., Machon, O., Røsok, Ø., Backman, M., and Krauss, S. (2002). The mouse enhancer element D6 directs Cre recombinase activity in the neocortex and the hippocampus. Mech. Dev. 110, 179-182.

Visel, A., Taher, L., Girgis, H., May, D., Golonzhka, O., Hoch, R., McKinsey, G.L., Pattabiraman, K., Silberberg, S.N., Blow, M.J., Hansen, D.J., Nord, A.S., Akiyama, J.A., Holt, A., Hosseini, R., Phouanenavong, S., Plajzer-Frick, I., Shoukry, M., Afzal, V., Kaplan, T., Kriegstein, A.R., Rubin, E.M., Ovcharenko, I., Pennacchio, L.A., and Rubenstein, J.L.R. (2013). A High-Resolution Enhancer Atlas of the Developing Telencephalon. Cell. 152, 895-908.

Vogt, D., Hunt, R.F., Mandal, S., Sandberg, M., Silberberg, S.N., Nagasawa, T., Yang, Z., Baraban, S., Rubenstein, J.L.R. (2014) Lhx6 directly regulates Arx and CXCR7 to determine cortical interneuron fate and laminar position. *Neuron*. In press.

Waclaw RR, Ehrman LA, Pierani A, Campbell K. (2010). Developmental origin of the neuronal subtypes that comprise the amygdalar fear circuit in the mouse. J Neurosci. 30, 6944-6953.

Xie, Q., Yang, Y., Huang, J., Ninkovic, J., Walcher, T., Wolf, L., Vitenzon, A., Zheng, D., Götz, M., Beebe, D.C., Zavadil, J., and Cvekl, A. (2013). Pax6 interactions with chromatin and identification of its novel direct target genes in lens and forebrain. PLoS One. 8, e54507.

Yoshida, M., Assimacopoulos, S., Jones, K.R., and Grove, E.A. (2006). Massive loss of Cajal-Retzius cells does not disrupt neocortical layer order. Development. 133, 537-545.

Yun, K., Potter, S., and Rubenstein, J.L.R. (2001). *Gsh2* and *Pax6* play complementary roles in dorsoventral patterning of the mammalian telencephalon. Development. 128, 193-205.

Figure Legends

Figure 1. Enhancer activity assays at E11.5 of transient transgenics expressing β-galactosidase from the *LacZ* gene. *: Stable transgenic lines were made using these enhancers. Coronal sections across the rostrocaudal telencephalon studied for 15 different enhancers. A) Five enhancers with a nested pattern of *LacZ* expression in the dorsal pallium. B) Five enhancers with a nested pattern of *LacZ* expression in the latero-ventral pallium. C) Five enhancers with a nested pattern of *LacZ* expression in the medial pallium. D) Schema of coronal sections across the rostrocaudal telencephalon showing progenitor domains and boundaries deduced from analysis of enhancer-driven expression patterns. E) Schema of coronal sections across the rostrocaudal telencephalon showing progenitor domains and boundaries (A-M) deduced from analysis of enhancer activity fate mapping (see subsequent Figures). Some boundaries are specific to rostral (r), whereas other boundaries are specific to caudal (c) regions. Abbreviations: see legend to Figure 2.

Figure 2. Enhancer activity (GFP expression, E11.5 (panels A-H, A'-H') and fate mapping (Cre induced tdTomato, E17.5; panels J-R; J'-R') assays of stable transgenics encoding enhancer 643 (left) or 1050 (right). Panels C' and E'' show higher magnification view of E11.5 expression. I,I': Schemas showing approximate position of GFP expression (grey) within flatten view of E11.5 pallial progenitor zones. S,S'': Schemas showing approximate position of dtTomato expression within flatten view of E17.5 pallial subdivisions; color coded according to approximate density of $tdTomato^{+}$ cells. Abbreviations according to region:

Ventral Pallium (VPall, allopallium); AO: anterior olfactory nuclei; OB: olfactory bulb; Pir/EPir; piriform and ectopiriform; LERh: lateral entorhinal; MERh: medial entorhinal.

Lateral Pallium (LPall, mesopallium): Ins/Cl: insula/claustrum; LO: lateral orbital; PRh: perirhinal; Orb: orbitofrontal.

Dorsal Pallium (DPall; neopallium): AU (A); auditory; DPF: dorsal prefrontal; F: frontal; LPF: lateral prefrontal; M: motor; SS: somatosensory; V: visual.

Dorsomedial Pallium (DMPall): Cing (C): cingulate gyrus; IL: infralimbic (and PrL: prelimbic); MOrb: medial orbital; RSP: retrosplenial; PoRh: postrhinal.

Medial Pallium (MPall): CA1-3: CA fields 1-3; DG: dentate gyrus; fi (F); fimbria; IG: indusium griseum; Sub (S): subiculum; PaS: parasubiculm; PrS: presubiculum; TT: tenia tecta.

Dorsal Midline: bac: brachium of the anterior commissure; bcc: brachium of the corpus callosum; bhc: brachium of the hippocampal commissure; ch; choroid plexus; PSe (PS): pallial septum.

Pallial Amygdala (Pall Amygd): AA: anterior amygdala; Ahi: amygdalohippocampal area; BM: basomedial; BLA; basolateral; LA: lateral.

Subpallium: Acb: accumbens; CGE: caudal ganglionic eminence; Dg: Diagonal area; LGE: lateral ganglionic eminence; MGE: medial ganglionic eminence; Pal: pallidum; SPSe: subpallial septum; St: striatum.

Hypothalamus: hp1, 2: hypothalamic prosomere 1 and 2; PHy: peduncular; Thy: hypothalamus.

Diencephalon: Hb; habenula; p2, p3: prosomeres 2 and 3; Thy: terminal hypothalamus; PThE: prethalamic eminence; Th: thalamus.

Figure 3. Eight enhancers with activity in pallial progenitors that fate map to prefrontal cortex subdivisions: medial: 192, 1056, 348; dorsal: 1050, 840; lateral: 636, 281, 1172. Coronal

sections through prefrontal cortex are shown: left column shows GFP expression at E11.5 *in situ* or immunohistochemistry. Right columns shows fate mapping with tdTomato expression in an E17.5 rostrocaudal series. See Figure S2 for additional E11.5 and E17.5 sections. I: Annotation of fate mapping results from selected enhancers (y axis) in five regions of the frontal cortex (x axis). Different levels of density of tdTomato expression are estimated and described as high density (red), medium density (orange), low density (yellow) and negligible density (green). In some cases we note subdomain expression. J: Deduced progenitor domain organization of the rostral E11.5 pallium. Abbreviations: see legend to Figure 2, and: CR: Cajal Retzius cells; DPFC: dorsal prefrontal; DLGE: dorsal LGE; FP: Frontal pole; MPFC: medial prefrontal; SP; subpallium.

Figure 4. Four enhancers with activity in pallial progenitors that fate map to ventrolateral cortex subdivisions. Left panels show GFP protein (green fluorescence) or RNA (purple in situ) expression at E10.5 (636; arrowhead: migrating neurons), or E11.5 (281, 218). Right columns shows fate mapping with tdTomato expression at E17.5. To map the fate map boundaries, double immunofluorescnece was performed to detect tdTomato (red) and either Nurr1 or Ctip2 (green). Nurr1 or Ctip2 expression was used to define boundaries 1, 2, 3 and 4 (see Results), that distinguished the limits of the fate maps of 636, 281,and 218. J: Annotation of fate mapping results from selected enhancers (y axis) in nine regions of the ventrolateral cortex (x axis). Different levels of density of tdTomato expression are estimated and described as high density (red), medium density (orange), low density (yellow) and negligible density (green). K: Deduced progenitor domain organization of middle-to-caudal regions of the E11.5 pallium. See Figure S2 for additional E11.5 and E17.5 sections. Abbreviations: see legend to Figure 2, and: Cl: claustrum; EPir: endopirifom; OT; olfactory tubercle; Neo; Neocortex; P Amgy; Pallial Amgydala.

Figure 5. Four enhancers with activity in medial pallial (MP) progenitors that fate map to hippocampal subdivisions. Coronal sections show GFP expression at E11.5, and fate mapping with tdTomato expression at P30. Two enhancers show activity and fate map to the rostrodorsal hippocampus (348 and 643); P30 fate map pictures are shown at 2x (Panel A', B', and C'), 4x (Panel A'', B'', and C'') and 10x (Panel A''', B''' and C''') magnifications. Enhancer 192 fate-maps to the fimbria and choroid plexus. One enhancer (653) shows activity and fate maps to the caudoventral hippocampus and choroid plexus; results are compared with the rostrodorsal hippocampal enhancer (348); P30 fate map pictures are shown at 2x (Panels D' and E'), and 4x (Panels D'' and E'') magnifications. See Figure 6 for fate mapping annotation, and Figures S2 for additional E11.5 and E17.5 sections. Abbreviations: see legend to Figure 2, and: CA1, CA3: hippocampal pyramidal cell fields; DG: dentate gyrus (Do: dorsal; Ve: ventral); HC: hippocampus; Hy; hypothalamus.

Figure 6. Expansion of the hippocampal primordium at E12.5. Comparison of the activity of MP enhancers 192 and 643 at E11.5 and E12.5. Enhancer activity (GFP expression) is compared with *Lmx1a* RNA expression using double immunohistochemistry (GFP)/*in situ* hybridization (*Lmx1a*). *Lmx1a* marks the F/Ch domain; at E11.5 both 192 and 643 have nearly identical patterns, sharing a common boundary (arrowhead). However, by E12.5 enhancer 643 activity expands beyond the *Lmx1a*/enhancer 192 boundary (arrowhead), into the neuroepithelium that generates the dentate gyrus and CA fields (arrow; see Figure 5C'-C'''); note that 192 activity is present in a few scattered hippocampal progenitors (Panel C, arrows). Panel E: Schema summarizing results in panels A-D, showing the expansion of the hippocampus (HC) between E11.5 and E12.5. Panel F: Annotation of fate mapping results from selected enhancers (y axis)

in 13 regions of the medial pallium (x axis). Different levels of density of tdTomato expression are estimated and described as high density (red), medium density (orange), low density (yellow) and negligible density (green). Panel G: Deduced progenitor domain organization of E11.5 medial pallium and other regions of the caudal pallium. Abbreviations: see legend to Figure 2.

Figure 7. Transcription regulation of pallial enhancers. A: % of enhancers with ChIP-Seq peaks for PAX6, COUPTFI and PBX1 on LVP+DP+MP, LVP; LVP+ DP DP+MP and MP enhancers. B: Transcription assays in transfected P19 cells (2 days) measuring luciferase expression driven by PAX6 activation of enhancers 636, 840, 643 and 1172. C. PAX6 ChIP-seq analysis from E12.5 cortex showing a peak directly over endogenous enhancer 840 (black bar). D. GFP pallial expression driven by enhancer 840 in E11.5 cortex. D': Reduced pallial GFP expression in *Pax6*$^{-/-}$. E. PAX6 ChIP-seq analysis from E12.5 cortex showing a peak directly over endogenous enhancer 636 (black bar). F. GFP pallial expression driven by enhancer 636 in E11.5 cortex. F': Reduced pallial GFP expression in *Pax6*$^{-/-}$. (Asterisk labels migrating VP neurons) Abbreviations: see legend to Figure 2, and: CDP and CMP: caudal

Figure 1

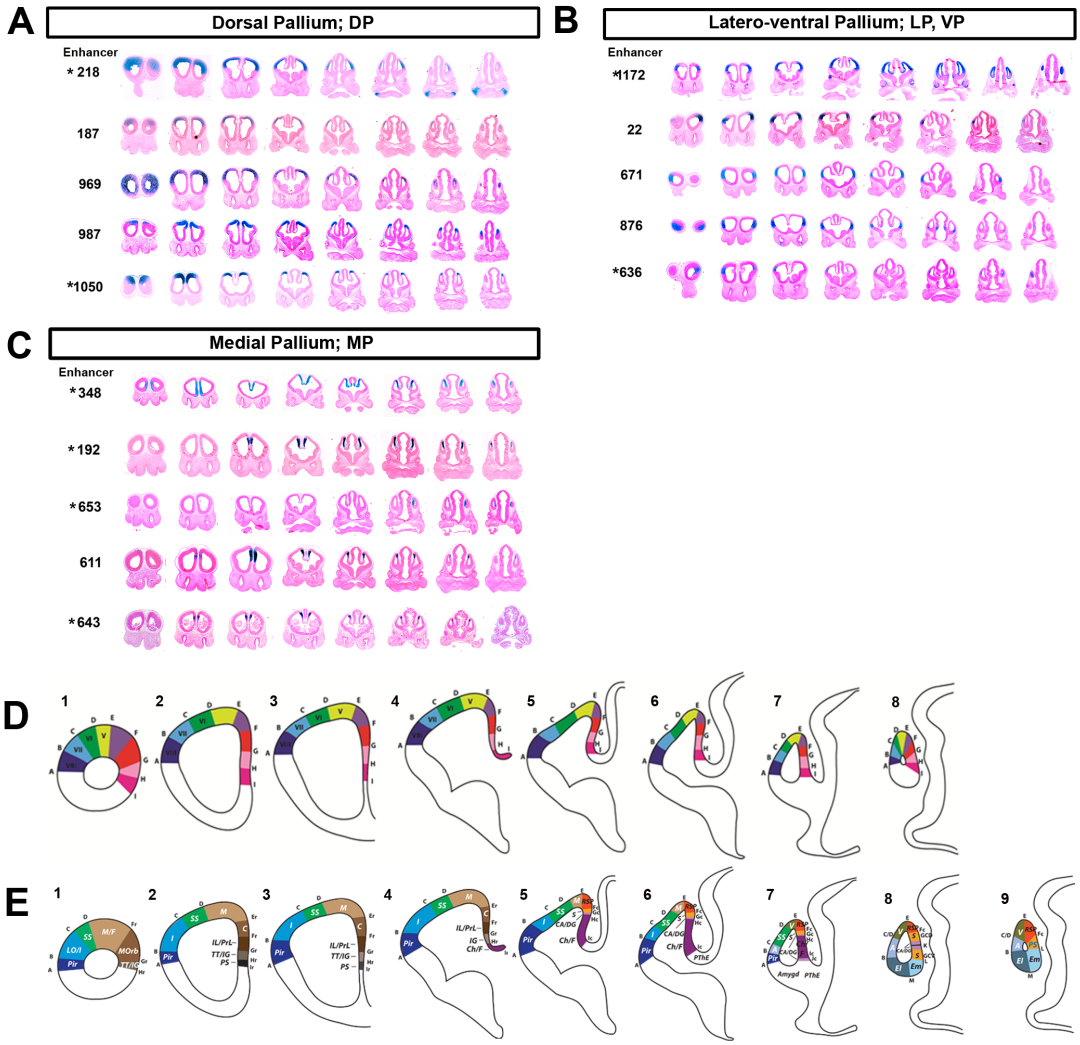

A
Dorsal Pallium; DP
Enhancer
* 218
187
969
987
*1050
B
Latero-ventral Pallium; LP, VP
Enhancer
*1172
22
671
876
*636
C
Medial Pallium; MP
Enhancer
*348
*192
*653
611
*643
D
1
2
3
4
5
6
7
8
E
1
2
3
4
5
6
7
8
9

Figure 2

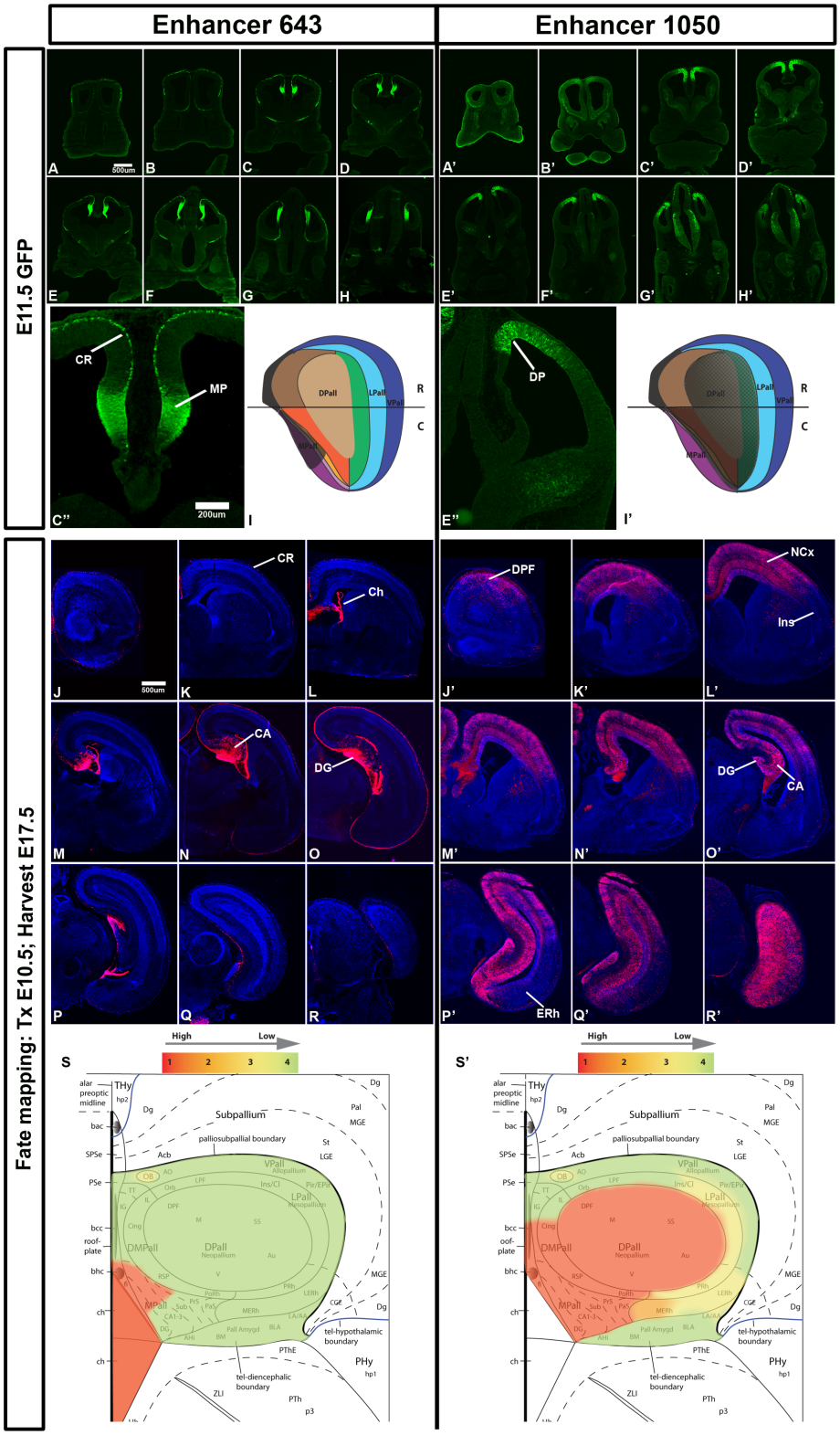

Enhancer 643
Enhancer 1050
E11.5 GFP
Fate mapping: Tx E10.5; Harvest E17.5
CR
MP
DP
DPall
LPall
VPall
MPall
R
C
200um
500um
CR
Ch
DPF
NCx
Ins
CA
DG
ERh
High
Low
Subpallium
palliosubpallial boundary
tel-hypothalamic boundary
tel-diencephalic boundary
alar preoptic midline
bac
SPSe
PSe
bcc
roof-plate
bhc
ch
THy
hp2
Dg
Pal
MGE
St
LGE
Acb
VPall
LPall
DMPall
DPall
MPall
PThE
PHy
hp1
ZLI
PTh
p3

Figure 3

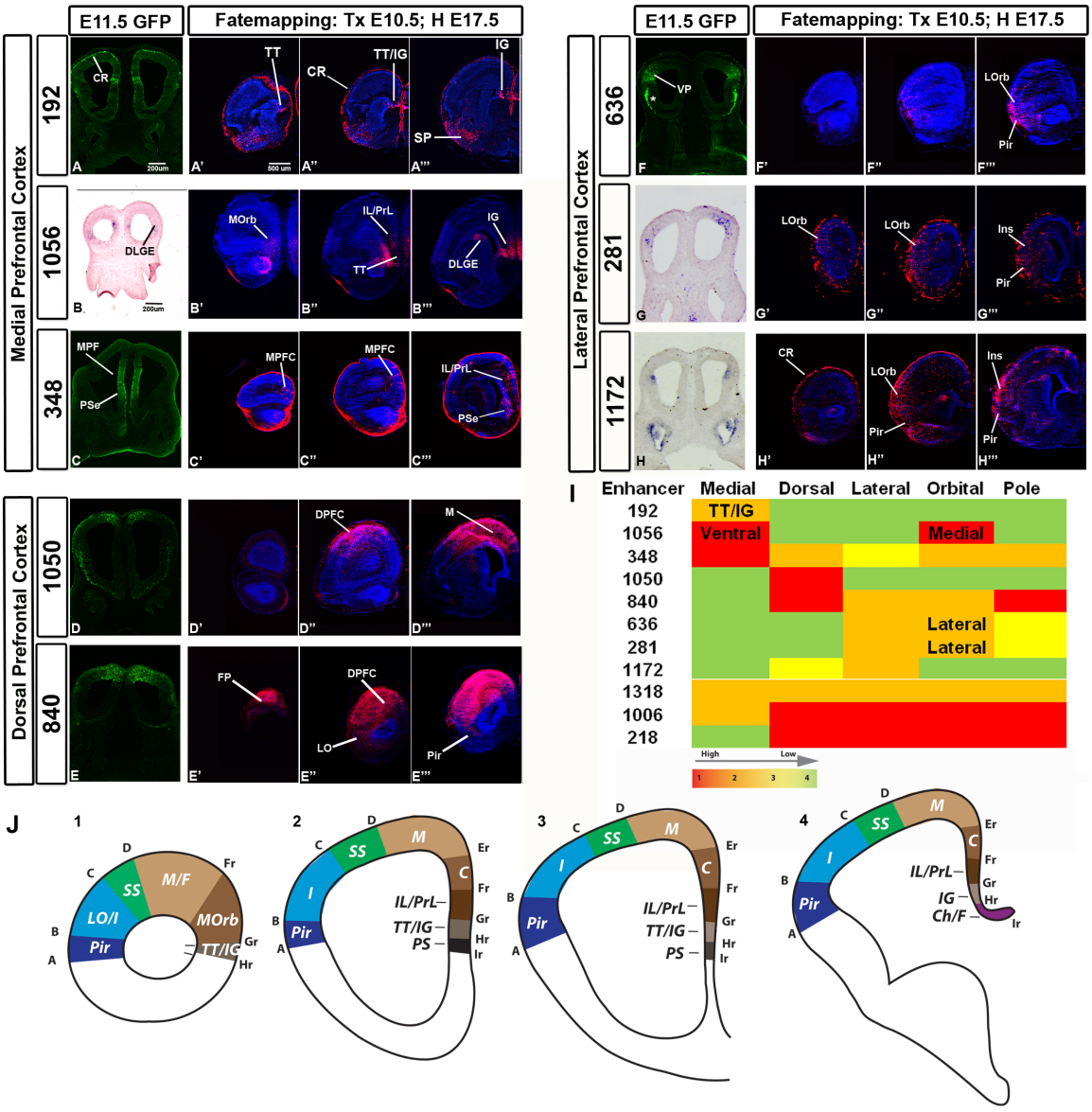
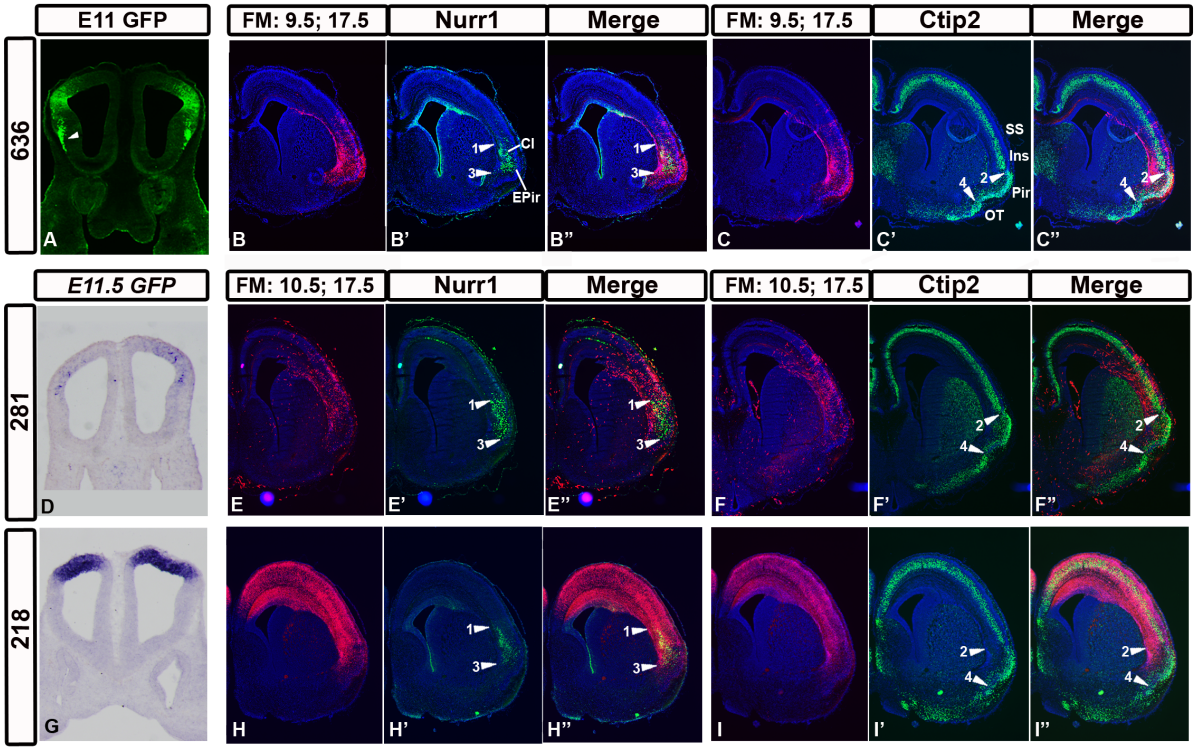

E11.5 GFP
Fatemapping: Tx E10.5; H E17.5
Medial Prefrontal Cortex
192
1056
348
Dorsal Prefrontal Cortex
1050
840
Lateral Prefrontal Cortex
636
281
1172
CR
TT
TT/IG
IG
SP
DLGE
MOrb
IL/PrL
MPF
PSe
MPFC
DPFC
M
FP
LO
Pir
VP
LOrb
Ins
Enhancer
Medial
Dorsal
Lateral
Orbital
Pole
TT/IG
Ventral
Medial
Lateral
High
Low
J
M/F
SS
LO/I
Pir
MOrb
TT/IG
I
C
IL/PrL
TT/IG
PS
Er
Fr
Gr
Hr
Ir
Ch/F

Figure 4



J

Enhancer	Rostral			Medial				Caudal	
	Neo	Insula	Piriform	Neo	Insula	Piriform	P Amgy.	Neo	ERh
1318									
840									
218									
1050									
1006									
281									
1172									
636									

K

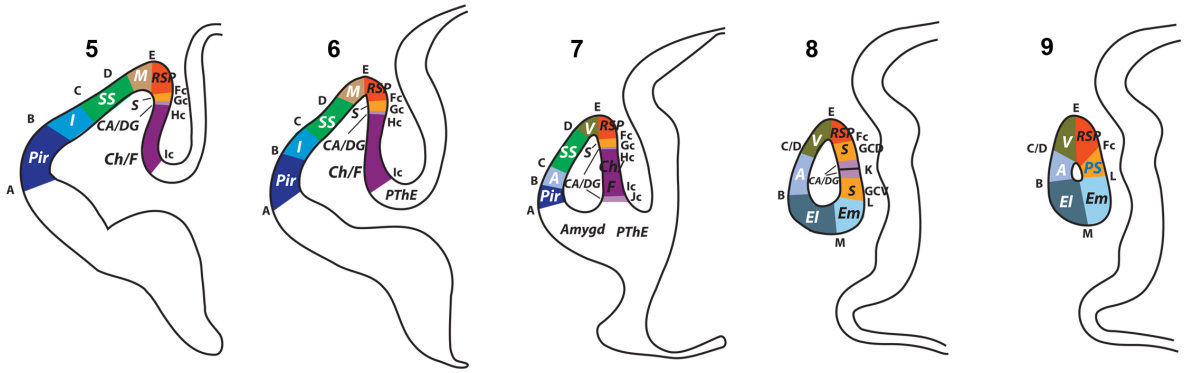


Figure 5

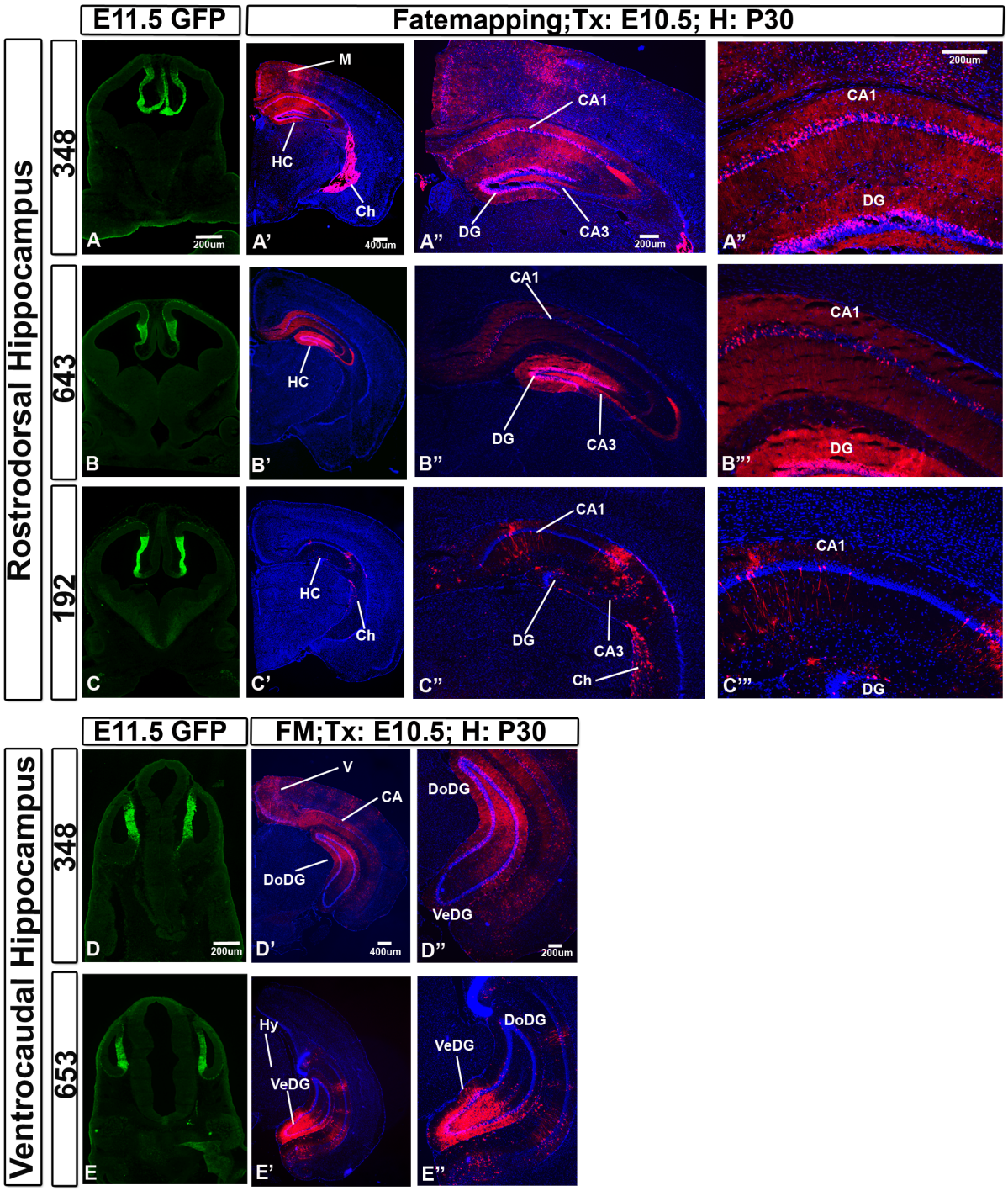

E11.5 GFP
Fatemapping;Tx: E10.5; H: P30
Rostrodorsal Hippocampus
348
643
192
M
HC
Ch
CA1
DG
CA3
200um
400um
A
A'
A''
A'''
B
B'
B''
B'''
C
C'
C''
C'''
E11.5 GFP
FM;Tx: E10.5; H: P30
Ventrocaudal Hippocampus
348
653
V
CA
DoDG
VeDG
Hy
D
D'
D''
E
E'
E''

Figure 6

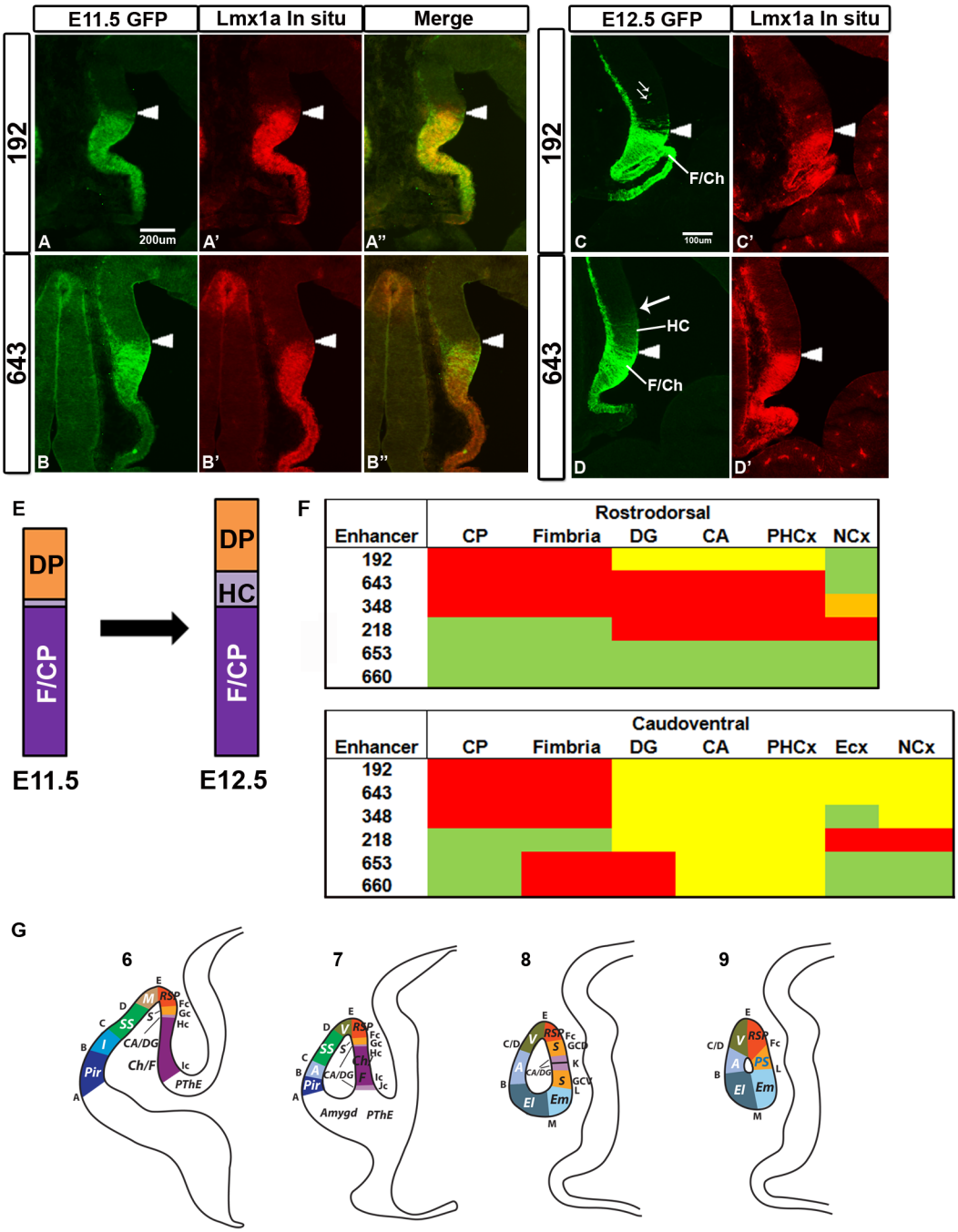
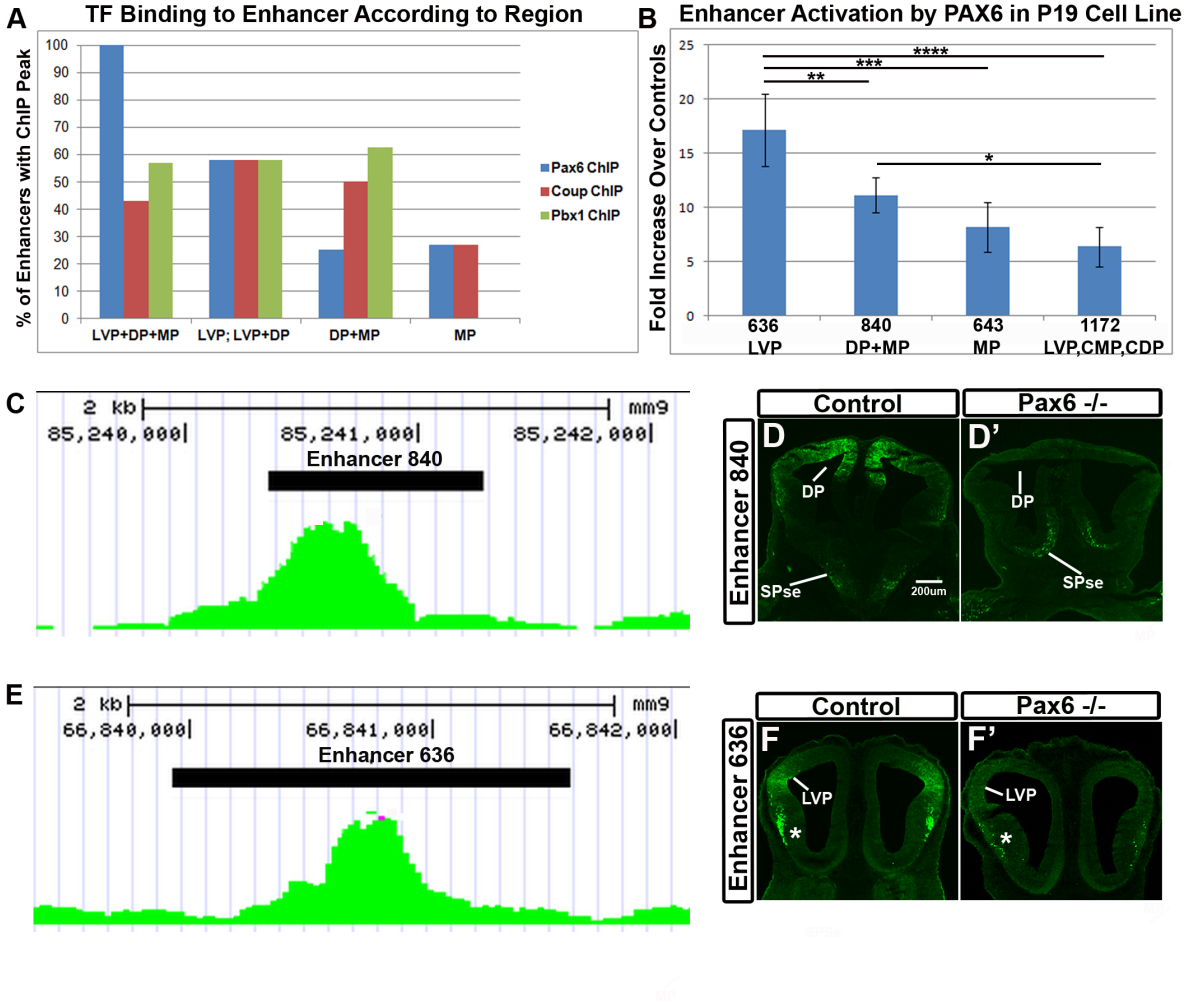

E11.5 GFP
Lmx1a In situ
Merge
E12.5 GFP
Lmx1a In situ
192
643
A
200um
A'
A"
C
100um
C'
F/Ch
B
B'
B"
HC
F/Ch
D
D'
E
DP
F/CP
DP
HC
F/CP
E11.5
E12.5
F
Rostrodorsal
Enhancer
CP
Fimbria
DG
CA
PHCx
NCx
192
643
348
218
653
660
Caudoventral
Enhancer
CP
Fimbria
DG
CA
PHCx
Ecx
NCx
192
643
348
218
653
660
G
6
7
8
9
Pir
CA/DG
Ch/F
PThE
Amygd
PThE
El
Em
PS

Figure 7


A
TF Binding to Enhancer According to Region
% of Enhancers with ChIP Peak
100
90
80
70
60
50
40
30
20
10
0
LVP+DP+MP
LVP; LVP+DP
DP+MP
MP
Pax6 ChIP
Coup ChIP
Pbx1 ChIP
B
Enhancer Activation by PAX6 in P19 Cell Line
Fold Increase Over Controls
25
20
15
10
5
0
****
***
**
*
636
LVP
840
DP+MP
643
MP
1172
LVP,CMP,CDP
C
2 kb
mm9
85,240,000
85,241,000
85,242,000
Enhancer 840
Control
Pax6 -/-
Enhancer 840
D
D'
DP
SPse
200um
E
2 kb
mm9
66,840,000
66,841,000
66,842,000
Enhancer 636
Control
Pax6 -/-
Enhancer 636
F
F'
LVP
*

Supplemental Information

Figure S1: Enhancer activity assays at E11.5 of transient transgenics expressing β-galactosidase from the LacZ gene (enhancer numbers and nearby genes are indicated). *: Stable transgenic lines were made using these enhancers. Coronal sections across the rostrocaudal telencephalon were studied for 36 different enhancers. A) 10 enhancers with a nested pattern of LacZ expression in the latero-ventral pallium. B) 15 enhancers with a nested pattern of LacZ expression in the dorsal pallium. C) 11 enhancers with a nested pattern of LacZ expression in the medial pallium.

Figure S2 (14 parts; A-N): Comprehensive compilation of data is assembled for each of the 14 enhancer stable transgenic lines, including: 1) schema of flattened E11.5 pallial progenitor domains indicating in grey the approximate E11.5 expression of the enhancer in the stable transgenic lines; 2) fate map of E11.5 enhancer activity using schema of flattened E17.5/P30 pallial regional domains; a red-green (high-low) scale indicates the density of $tdTomato^{+}$ cells; 3) coronal sections (rostrocaudal series) of E11.5 forebrain showing activity of enhancer (x-gal reaction product in blue) in transient transgenic; 4) for some enhancers, expression of nearby gene using *in situ* RNA hybridization coronal sections (rostrocaudal series) of E11.5 forebrain; 5) expression of GFP from stable enhancer transgene at E11.5, E12.5, E14.5 and E17.5 shown only for enhancers with GFP expression (E10.5 for some enhancers); 6) tamoxifen-independent Cre activation of tdTomato expression assayed at E17.5 shown only for enhancers with tdTomato expression; 7) tdTomato expression for $CreER^{T2}$ fate map analysis with tamoxifen given at E10.5 and analysis at E17.5 (and P30 for some enhancers). Abbreviations: see legend to Figure 2.

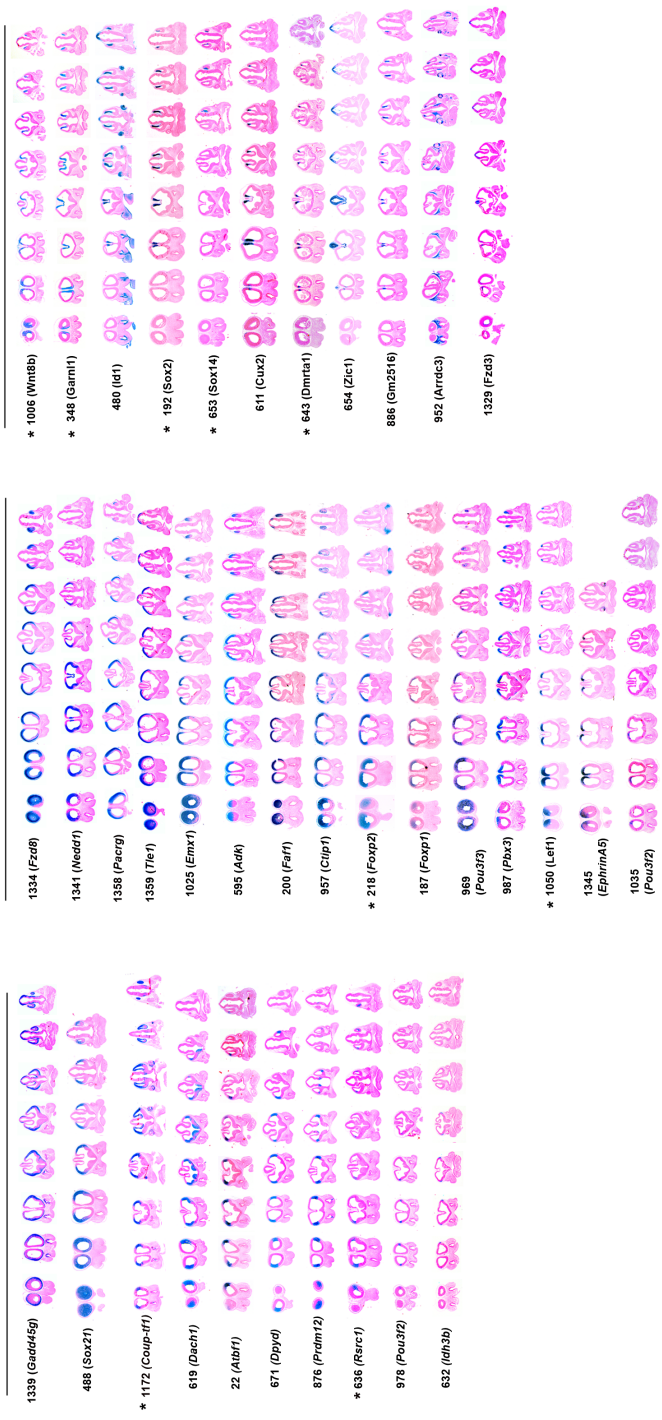
Figure S3: Informatic method to identify TF binding sites enriched in enhancers with different regional activities. A. Area under the Receiver Operator Characteristic (ROC) curve for the two classifiers employed in the analysis, describing their performance at discriminating MP enhancers from background sequences. The curve is generated by plotting the true positive rate (TPR) against the false positive rate (FPR). Model 1 was trained to distinguish between 9 MP enhancers, 15 non-MP enhancers, and 480 random genomic sequences. Model 2 was trained to distinguish between 9 MP enhancers, 15 non-MP enhancers, and a set of background sequences consisting of 480 random genomic sequences and 765 sequences from the VISTA browser that were negative for enhancer activity (see Methods). B. Motifs ranked according to their importance (highest and lowest), according to each of the two classifiers (see Methods). Only 15 most important motifs were selected from the motif rankings generated by each model. High positive values indicate that the corresponding motif is exclusively enriched in MP enhancers, and thus, relevant for correctly discriminating MP enhancers. In contrast, negative importance values suggest that the motif is most likely depleted among MP enhancers, and therefore, can have a detrimental impact on the classification.

Figure S4: Transcription factors identified using the informatic approach described in Results and Figure S3, that show expression: A) limited to the MP; B) in the DP and low/absent from the MP (Ch/F). Data is E11.5 in situ hybridization from the Allen Brain Institute Atlas of the Developing Mouse Brain; http://developingmouse.brain-map.org/.

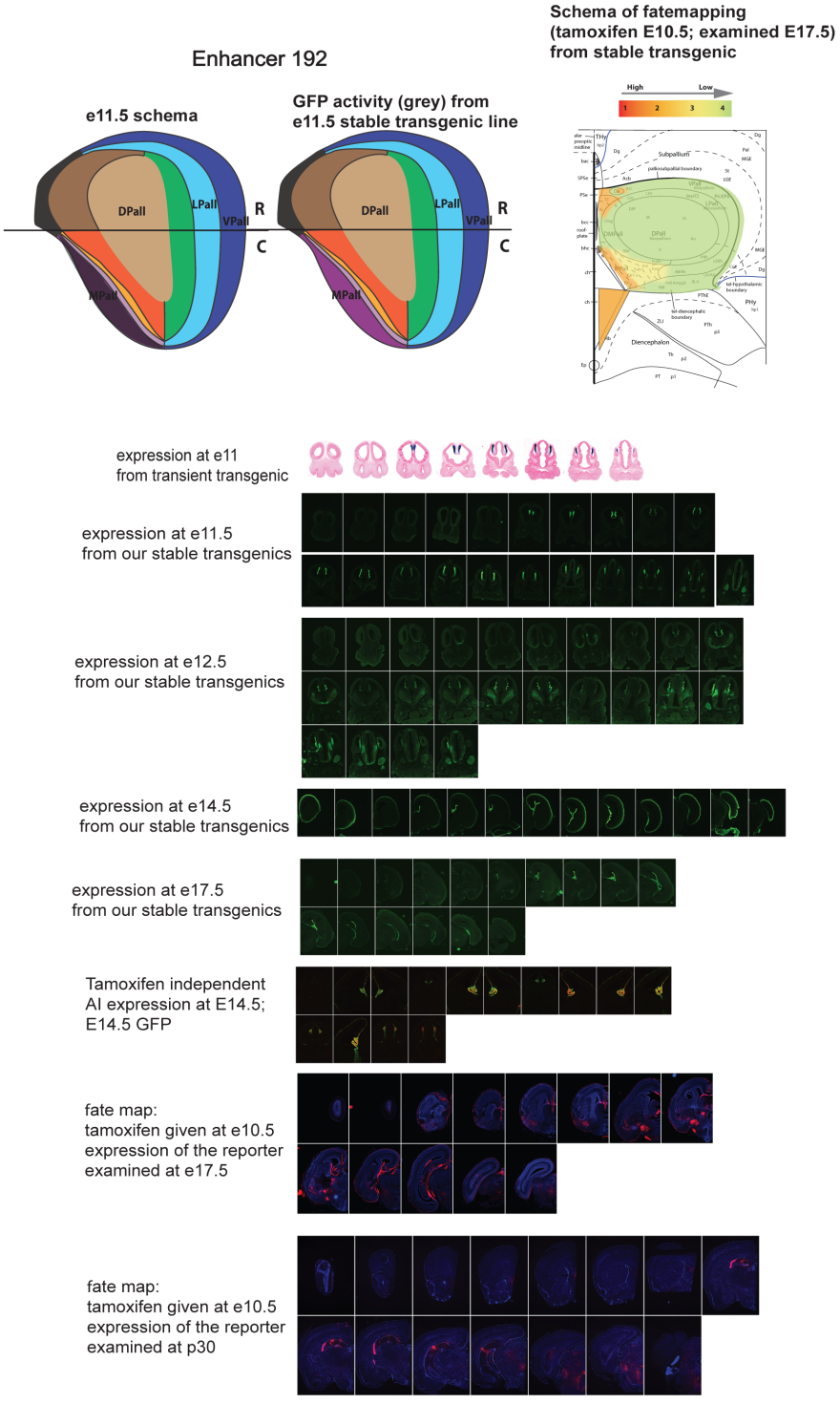
Figure S5: TF ChIP-Seq analysis. A: PAX6; B: COUPTFI; C. PBX1. First tier: mouse genome (mm9) coordinates and scale. Second tier: black bar indicating the length of the enhancer under investigation (PAX6: 281, 636, 643, 840, 1172; COUPTFI: 281, 840, 1172; PBX1: 1050, 1318).

Third tier: TF ChIP-Seq peaks that were in the same position as the enhancers. Controls: Fourth tier: ChIP that included a PAX6 or PBX1 peptide that binds to the anti-PAX6 antibody blocks the peaks (for COUPTFI IgG control); Fifth tier input control for PAX6 and PBX1. Sixth tier (fifth tier for COUPTF1): Evolutionary conservation of the mouse genome region shown using the UCSC genome browser. A'-A''': GFP IHC expression at E11.5 for enhancers 636(A'), 643 (A''), and 840 (A'') in controls and *Pax6* mutants (*Sey/Sey*), placed next to the picture of the enhancer's ChIP-Seq peaks.

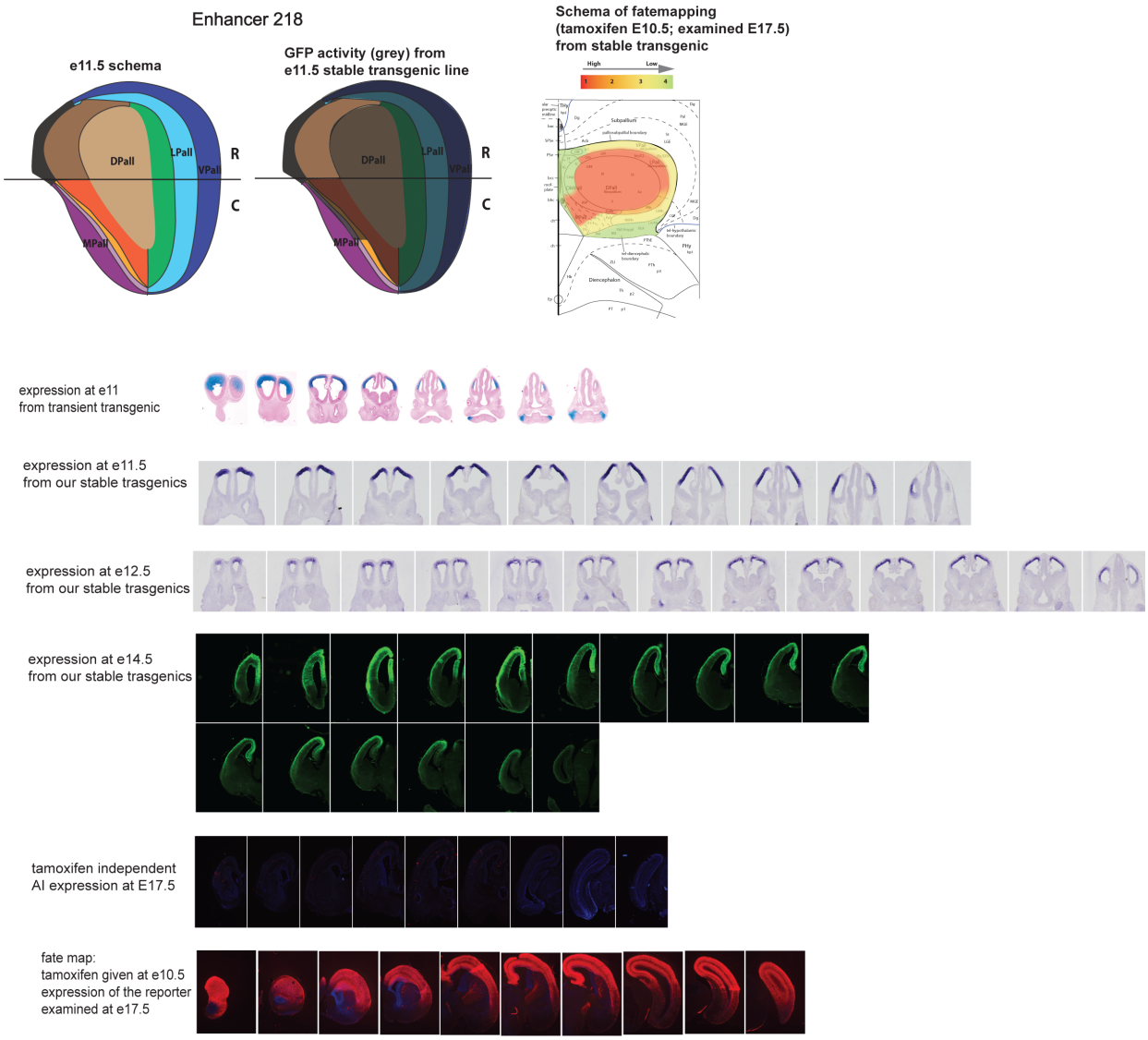
Supplemental Figure 1



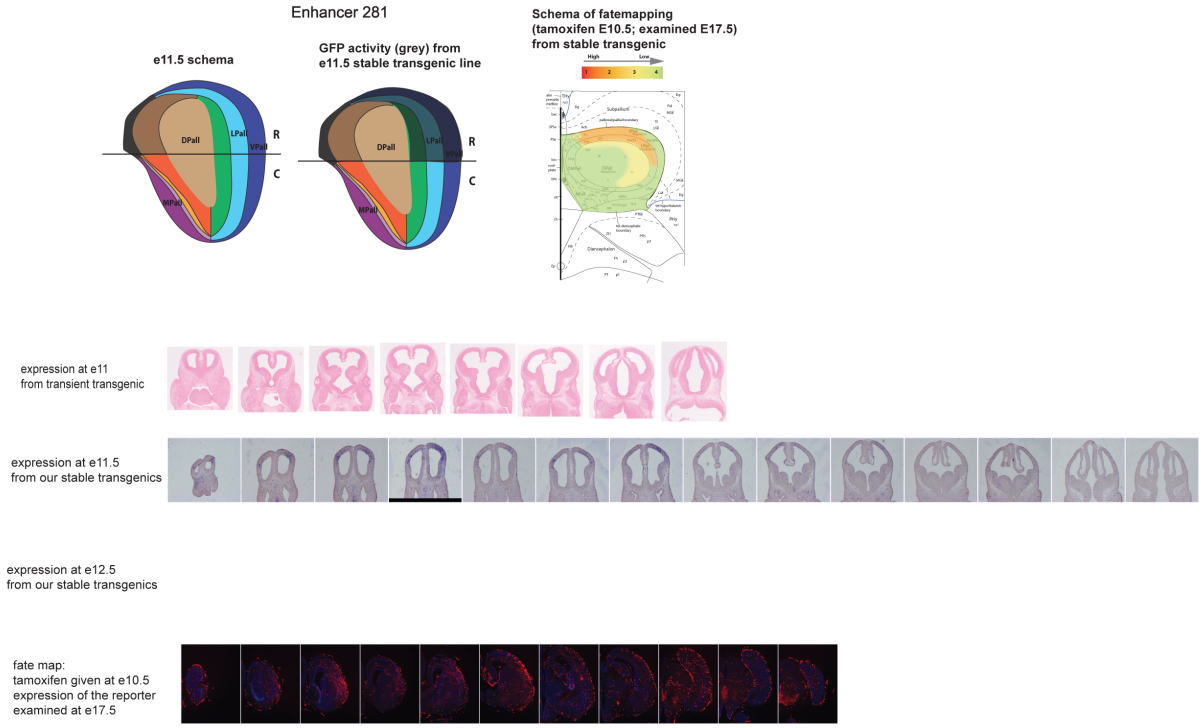
Supplemental Figure 2A



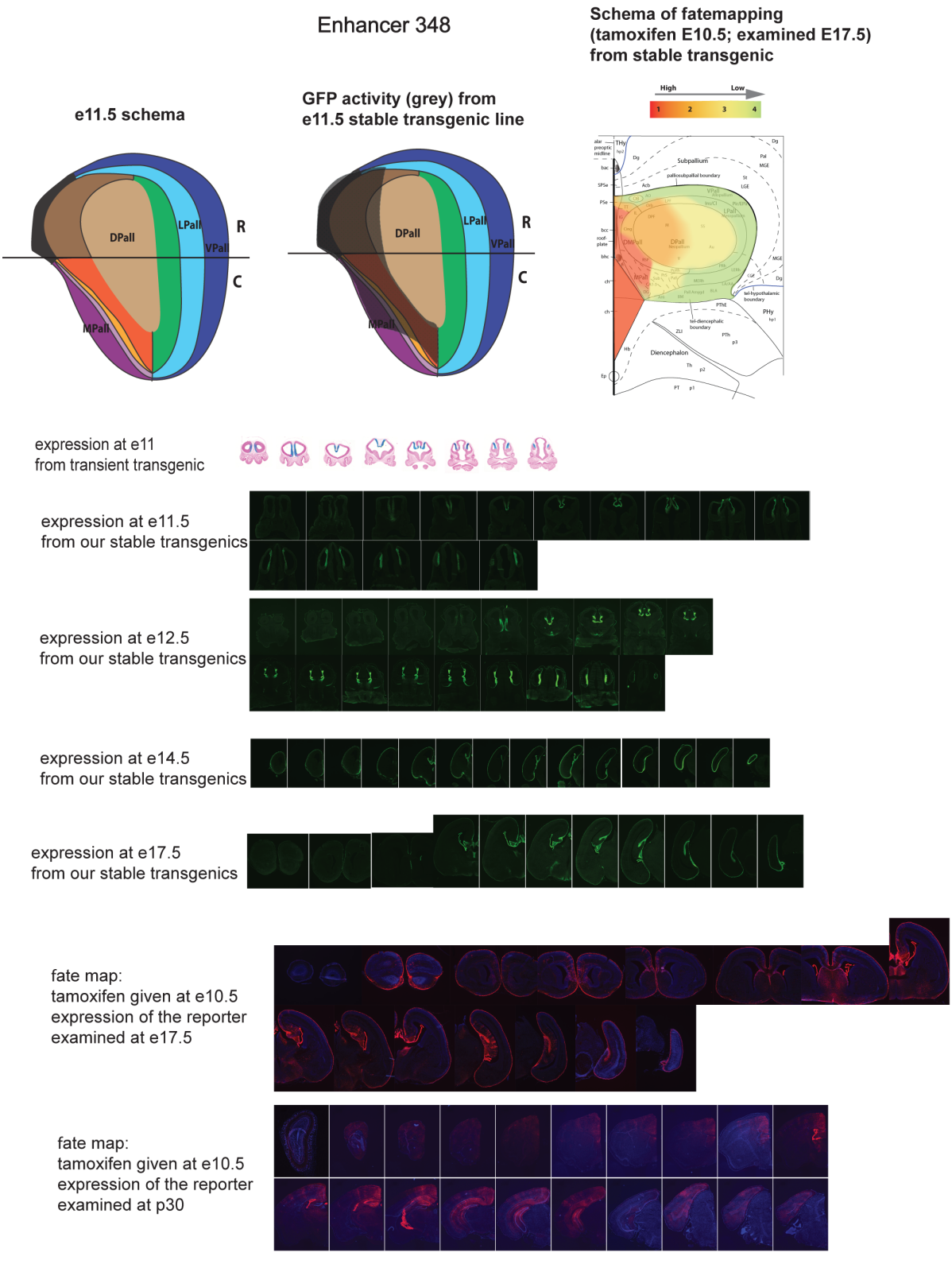
Supplemental Figure 2B



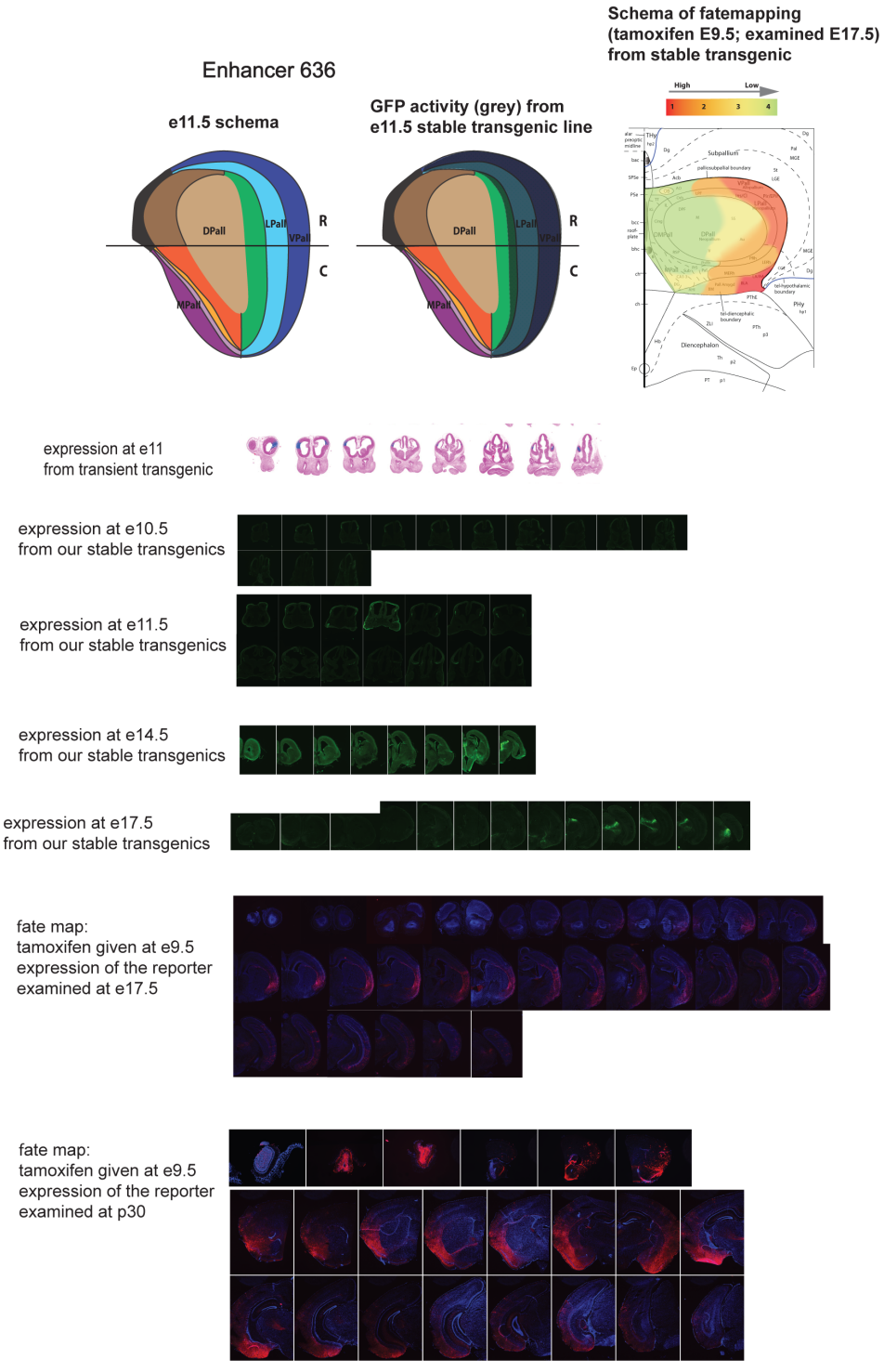
Supplemental Figure 2C



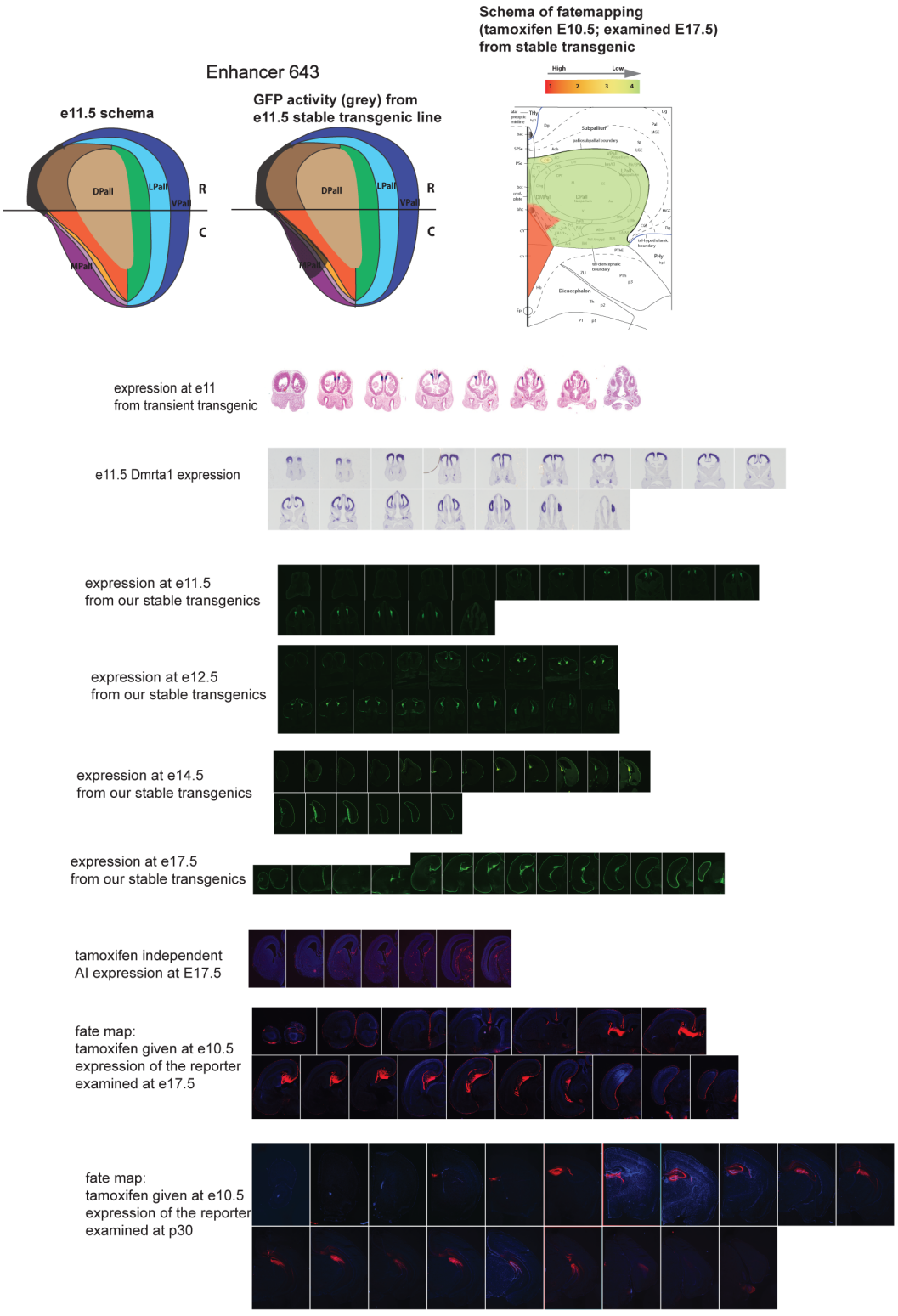
Supplemental Figure 2D



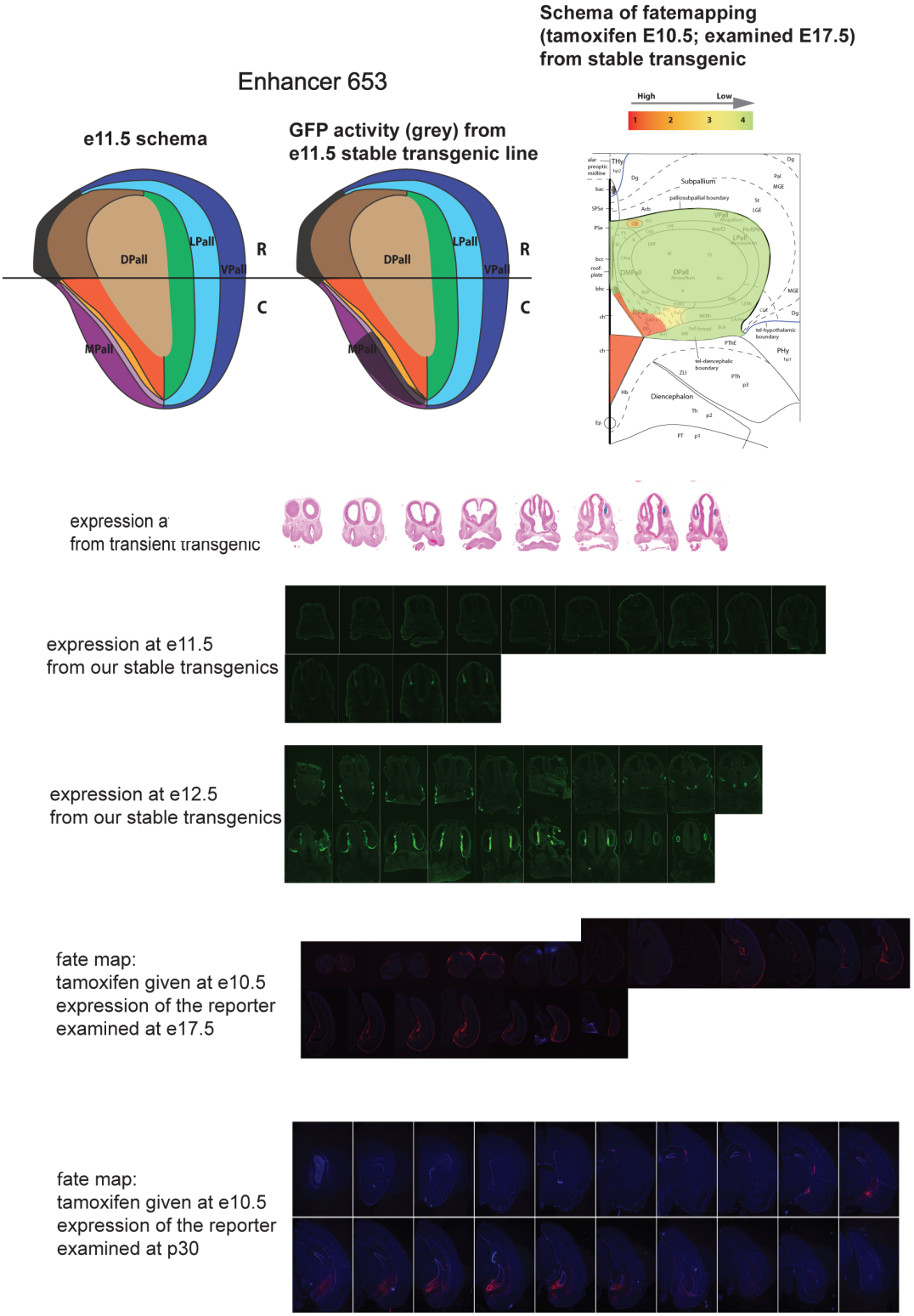
Supplemental Figure 2E



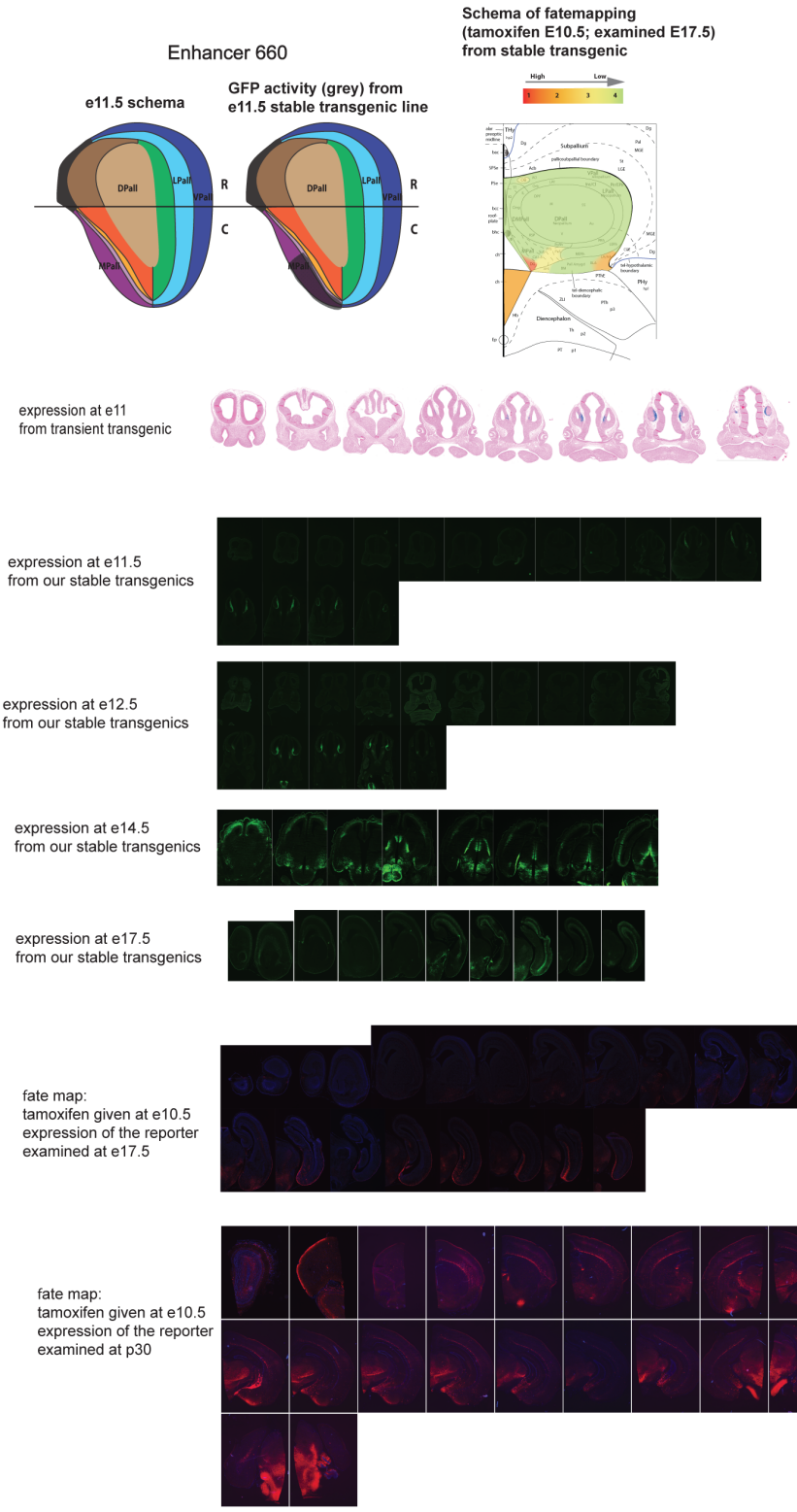
Supplemental Figure 2F



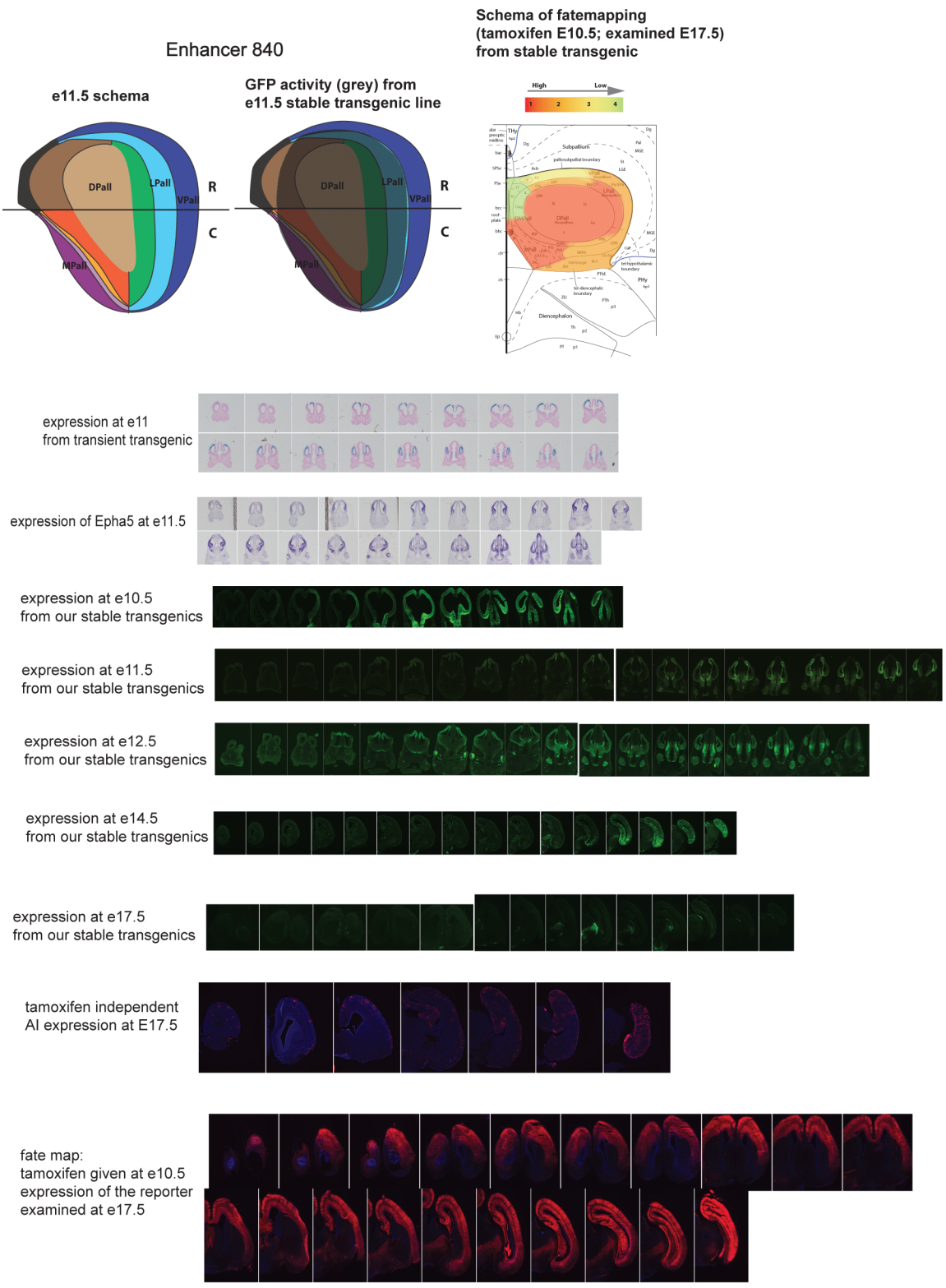
Supplemental Figure 2G



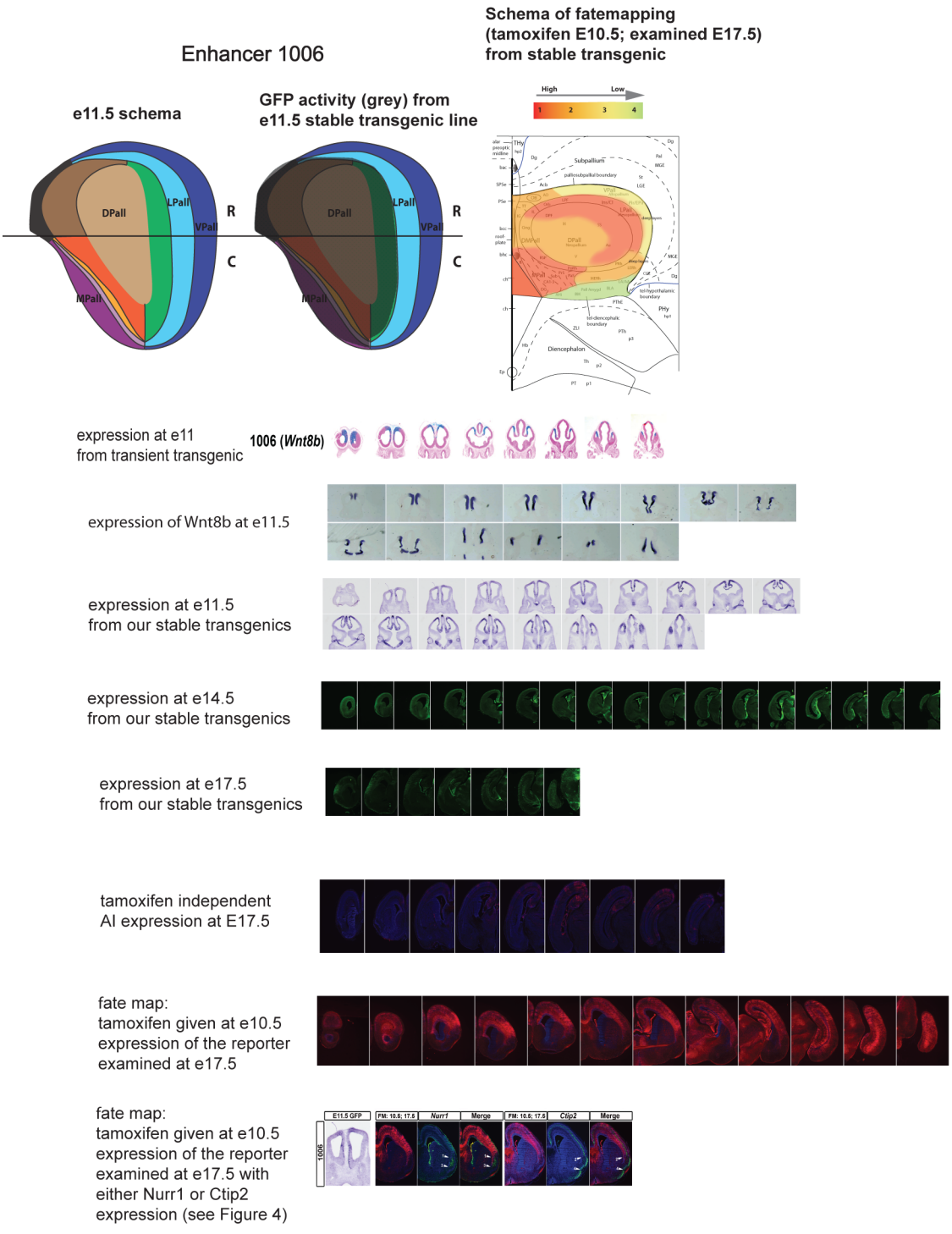
Supplemental Figure 2H



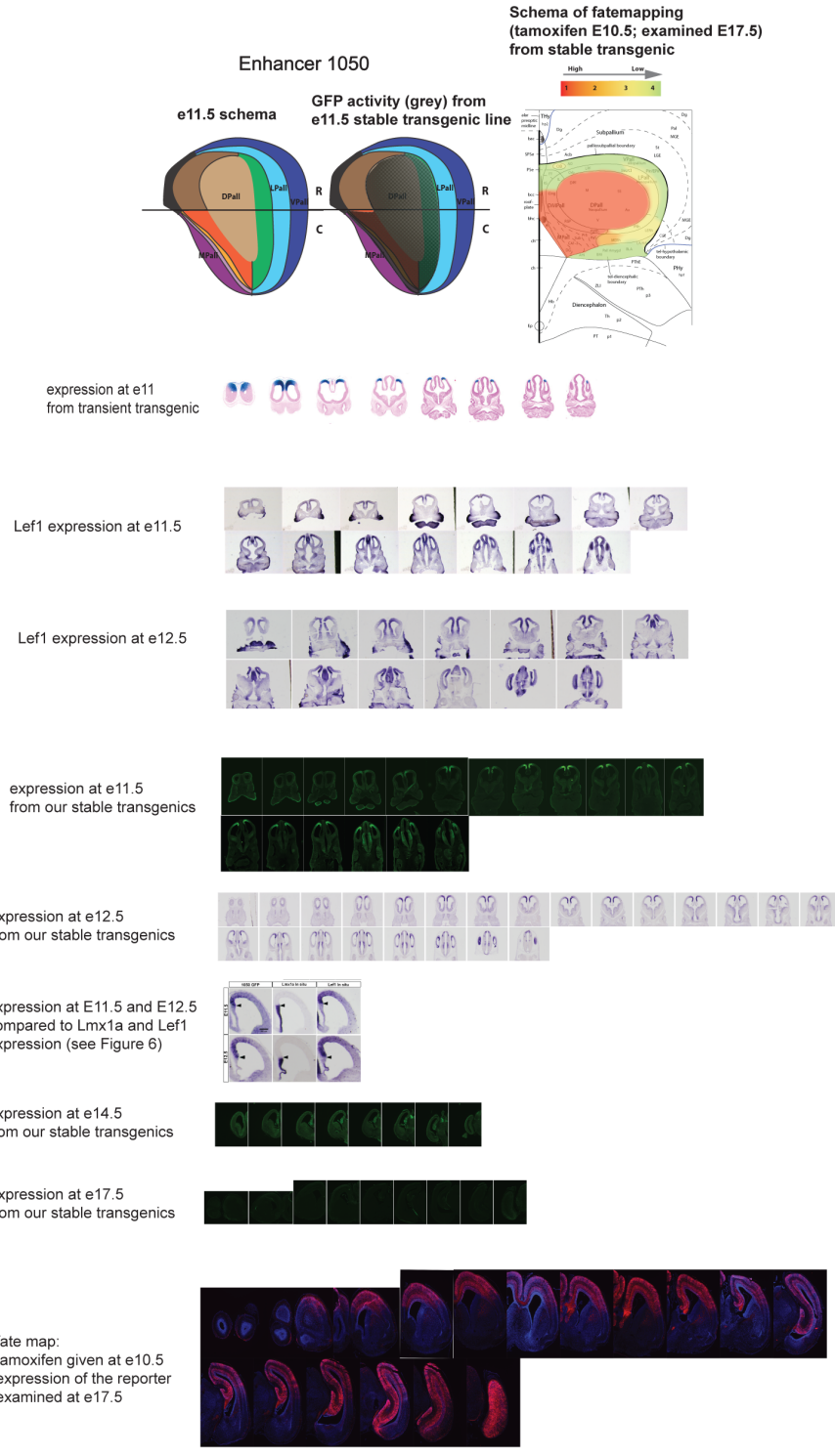
Supplemental Figure 2I



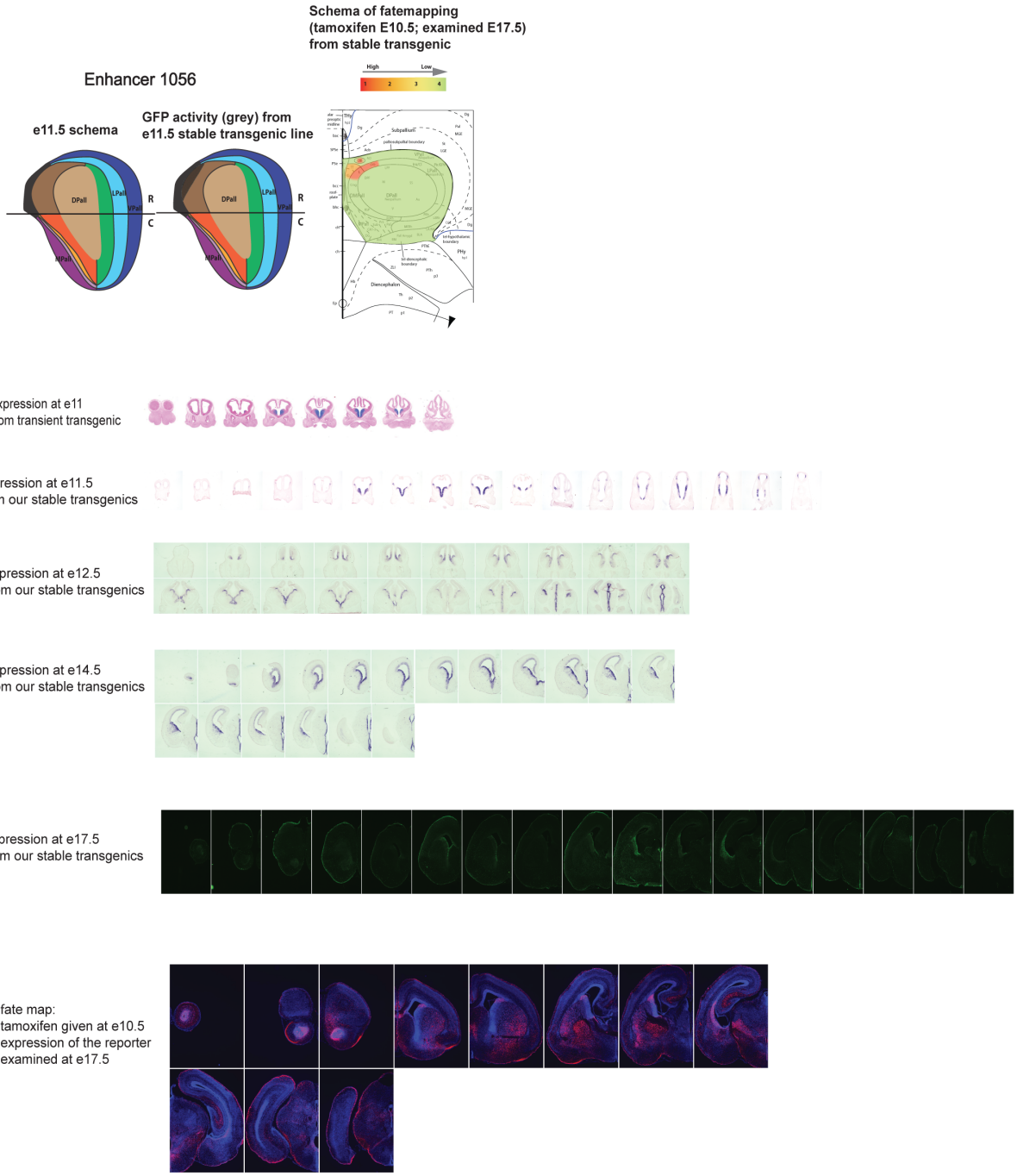
Supplemental Figure 2J



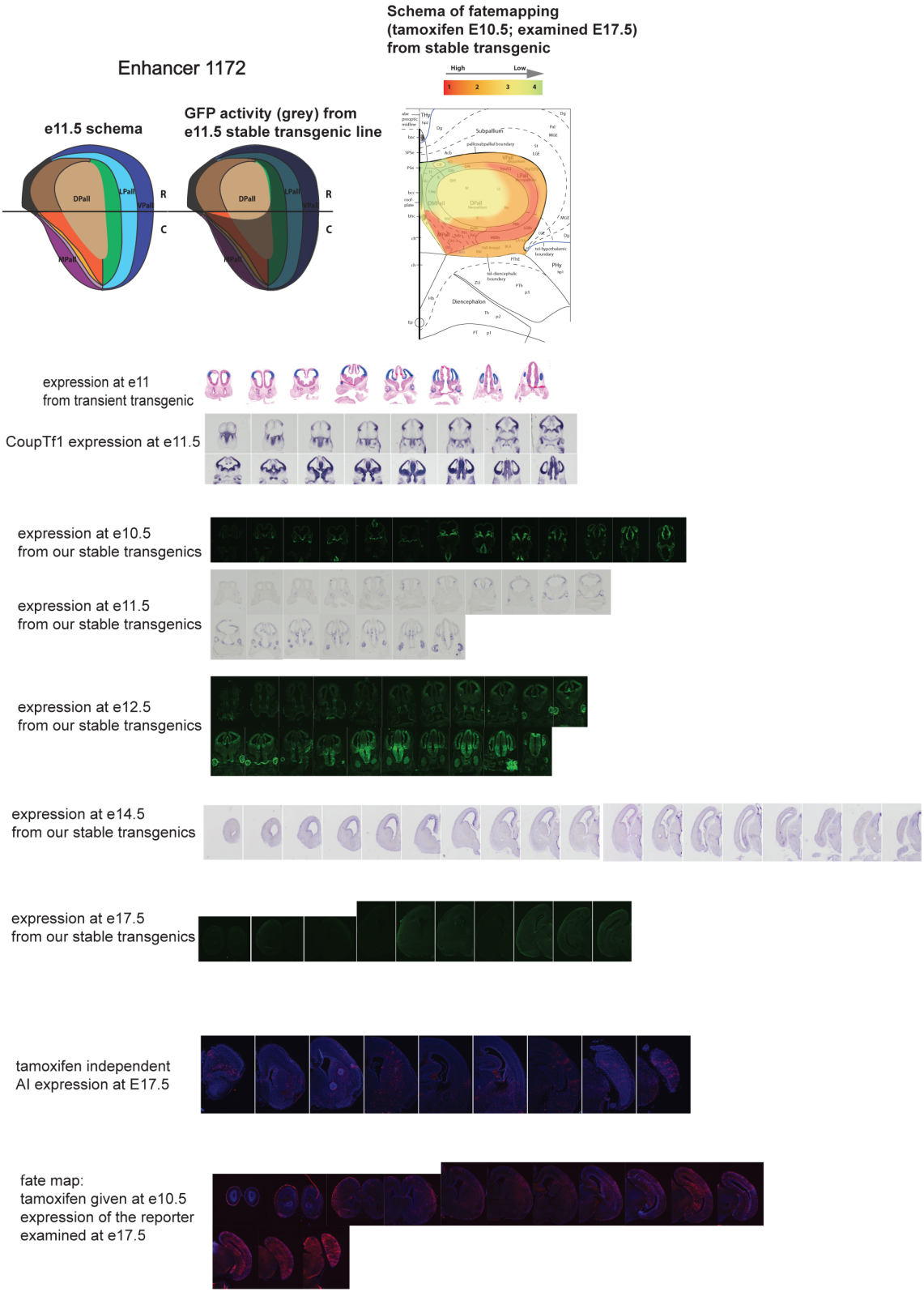
Supplemental Figure 2K



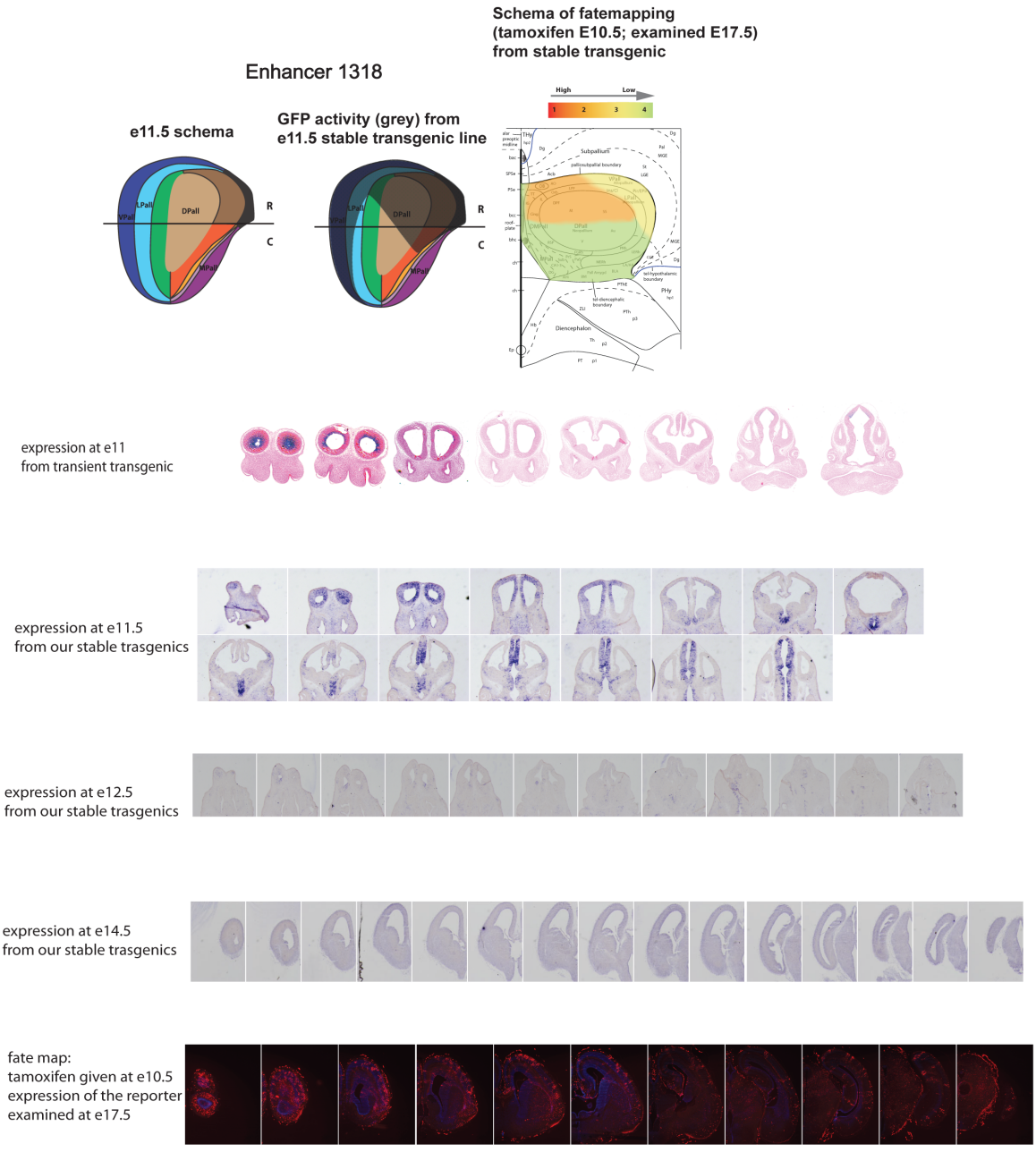
Supplemental Figure 2L



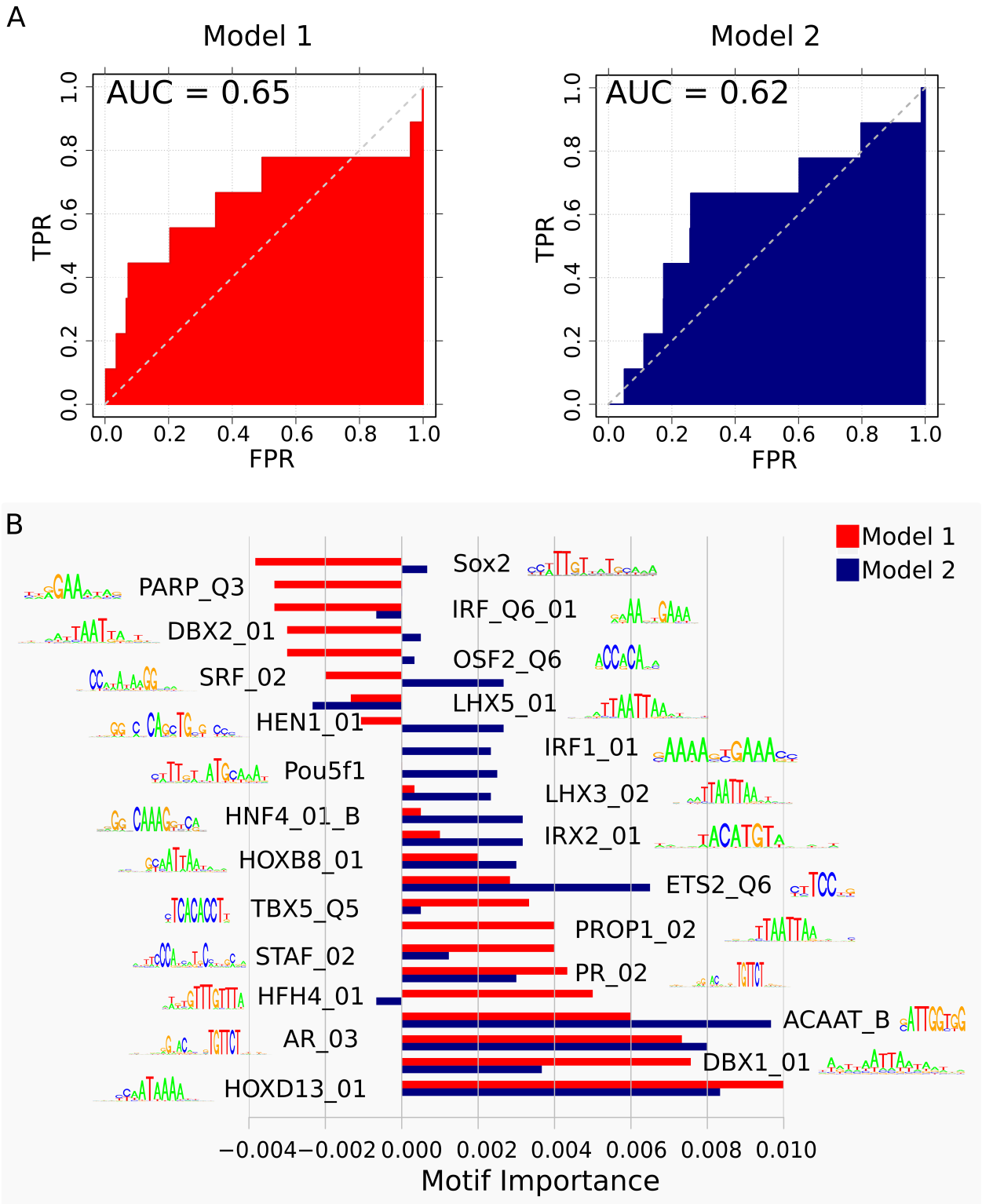
Supplemental Figure 2M



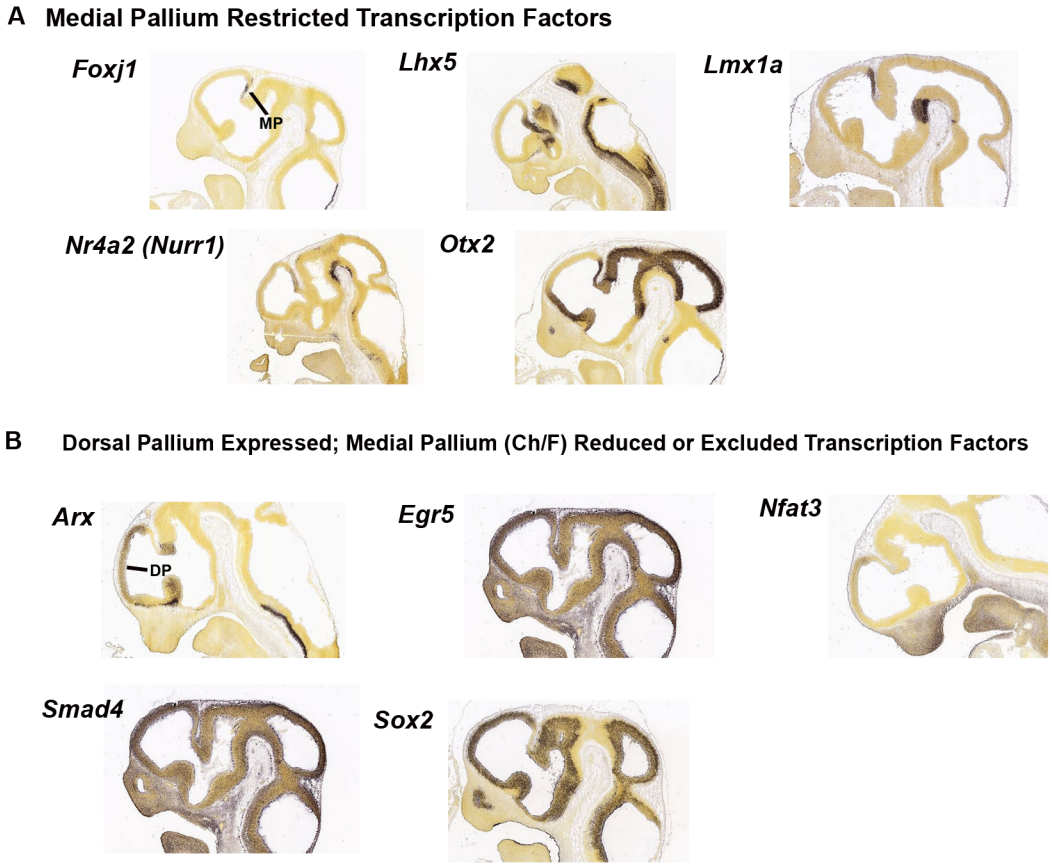
Supplemental Figure 2N



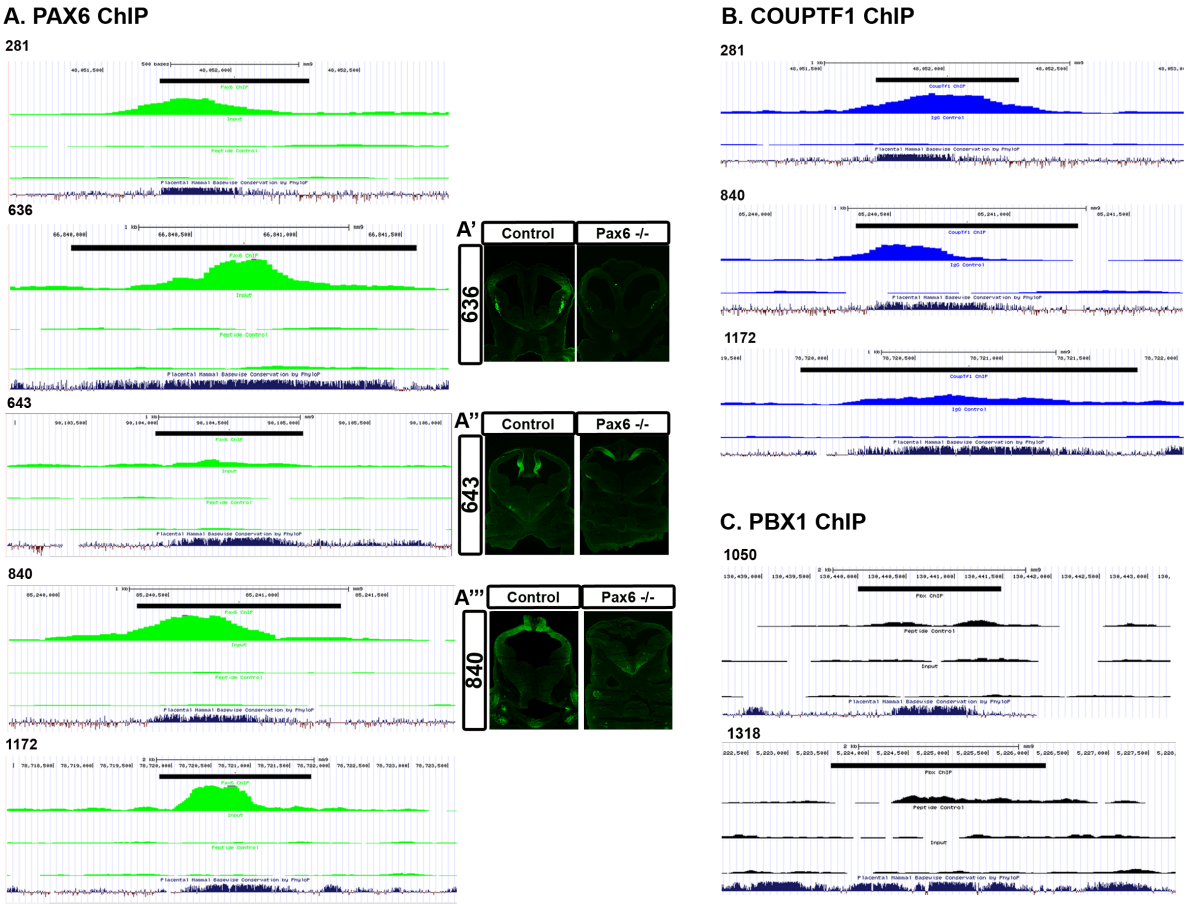
Supplemental Figure 3



Supplemental Figure 4



Supplemental Figure 5



Supplemental Table Legends

Table S1. Boundaries of enhancer activity observed in E11.5 transient transgenic assays (LacZ expression) measured in eight planes of section. See Figure 1D for the boundaries and Legend to Figure 1 for definitions of abbreviations.

Table S2. Information about quality of the sequence mapping and peak calling for each of the experiments: ChIP, input and negative control. Column 1: Description of library. Column 2: Total number of reads. Column 3: Number of reads after filtering. Column 3: Number of unique regions read in mm9 genome. Column 4: Total number of called peaks. Column 5: Number of peaks after filtering to remove peaks where a significant proportion of the reads originated in repetitive regions of the genome.

Table S3. Enhancer binding by PAX6, COUPTF1 and PBX1 (ChIP-Seq peaks). Column 1: enhancer number. Column 2: gene closest to specified enhancer. Columns 3, 4, and 5: presence of a called peak in the PAX6, COUPTF1, and PBX1 ChIP-Seq. The enhancers are organized by E11.5 activity pattern: LVP; LVP+DP, DP+MP, MP, and rostral pallium. Total number of enhancer peaks and occupancy % per region is calculated for each group. Column 6: Additional notes about the expression pattern.

Table S1

Transient transgenic activity protodomains based on LacZ expression at E11.5								
Enhancer	Section 1	Section2	Section 3	Section 4	Section 5	Section 6	Section 7	Section 8
192	None	None	F-H	F-H	F-H	F-H	F-H	F-H
218	B-F	B-F	B-F	B-F	B-E	B-D	B-D	B-D
281	None	B-C	B-C	B-C	None	None	None	None
348	F-H	F-SP septum	F-SP septum	E-I	E-DETh	E-I	E-I	E-I
636	A-C	A-C	A-C	B-C	B-C	B-C	B-C	B-C
643	None	None	F-G	F-G	F-G	F-G	F-G	F-G
653	None	None	None	None	None	None	None	G-I
660	None	None	None	None	None	None	G-I	G-I
840	C-F	B-F	B-G	B-G	B-G	A-I	A-I	A-I
1006	D-H	D-H	D-H	D-H	D-H	D-H	D-H	D-H
1050	D-G	D-G	D-F	D-F	D-F	D-F	D-F	D-F
1056	None	None	None	None	None	None	None	None
1172	None	LGE-D	LGE-D	LGE-D	LGE -D	LGE-I	I	A-I
1318	A-I (VZ)	A-C	None	None	None	None	None	None

Table S2

Library	Total Reads	Total Filtered Reads	Unique mm9 Mapped Reads	Peaks	Filtered Peaks
ChIP					
E13.5 Pax6	23784414	20746883	8119123	15352	12176
E12.5 Pax6 replicate 1	29723465	25840233	8788046	10656	6717
E12.5 Pax6 replicate 2	32625840	28396912	9221025	6889	3564
E12.5 CoupTf1	23711807	20179454	7777784	16023	11691
E12.5 Pbx	32623391	24484910	7165828	5977	4739
Input					
Input (E13.5 Pax6)	24599395	21443956	9321761	1731	260
Input (E12.5 Pax6 rep 1)	40175400	34880798	12232292	1761	176
Input (E12.5 Pax6 rep2)	31766143	27504004	9654826	1714	178
Input (E12.5 CoupTf1)	29199234	25338108	10973032	1476	152
Input (E12.5 Pbx)	42380742	31731256	9917026	1767	389
Negative Control					
Peptide (E12.5 Pax6 rep1)	29598294	25744786	8680860	1545	158
Peptide (E12.5 Pax6 rep2)	31028763	26927775	9120999	1567	154
IgG (E12.5 CoupTf1)	26980148	22760024	9143483	1503	204
Peptide (E12.5 Pbx)	30122258	22085488	5861940	1325	189

Table S3

Summary of ChIP-Seq Data					
		Pax6 ChIP	Coup ChIP	Pbx1 ChIP	Notes
LVP+DP+MP Enhancers					
840	Epha5	Y	Y	N	
1172	CoupTf1	Y	Y	N	Caudal
200	Faf1	Y	Y	Y	
595	Adk	Y	N	Y	Not Ch/F
957	Ctip1	Y	N	Y	Not Ch/F
122	Arx	Y	N	N	Not Ch/F
853	Cyclin h	Y	N	Y	
Total Peaks		7	3	4	
Occupancy %		100	43	57	
LVP; LVP+DP Enhancers					
636	Rsrc1	Y	N	N	
281	Foxp4	Y	Y	N	
619	Dach1	N	Y	Y	
22	Atbf1	N	Y	Y	
671	Dpyd	Y	Y	Y	
876	Prdm12	N	N	N	

978	Brn2	Y	N	Y	
1339	Gadd45g	N	Y	N	
488	Sox21	Y	N	Y	
1334	Fzd8	N	Y	Y	
1341	Nedd1	Y	Y	Y	
187	Foxp1	Y	N	N	
Total Peaks		7	7	7	
Occupancy %		58	58	58	
DP+MP Enhancers					
1006	Wnt8b	N	N	N	
1050	Lef1	N	N	Y	
218	Foxp2	N	N	N	
1358	Pacrg	Y	Y	Y	
1359	Tie1	Y	Y	Y	
1025	Emx1	N	Y	Y	
987	Pbx3	N	N	N	
1345	EphrinA5	N	Y	Y	
Total Peaks		2	4	5	
Occupancy %		25	50	62.5	
MP Enhancers					

348	Garnl1	N	N	N	
643	Dmrta1	Y	N	N	
653	Sox14	N	N	N	
660	Smad3/6	N	N	N	
611	Cux2	Y	Y	N	
886	Gm2516	N	N	N	
1329	Fzd3	N	Y	N	
192	Sox2	N	N	N	
480	Id1	N	N	N	
654	Zic1	Y	Y	N	
952	Arrdc3	N	N	N	
Total Peaks		3	3	0	
Occupancy %		27	27	0	
Rostral Enhancers					
1318	Zfhx4	N	N	Y	
1035	Pou3f2	N	Y	N	Rostrodorsal
1056	Sall3	N	N	N	Rostrodorsal
Total Peaks		0	1	1	
Occupancy %		0	33.3	33.3	

Extended Experimental Methods

Generation of Stable Enhancer Transgenic Mice

Enhancer regions were amplified by PCR from human genomic DNA (Visel et al., 2013). Enhancers were subcloned into *Hsp68-CreERT2-IRES-GFP* (Visel et al., 2013). CreERT2 was obtained from Pierre Chambon (IGBMC/GIE-CERBM) (Feil et al., 1997). Stable transgenic mice were generated by pronuclear injection at the UCSF Transgenic Core. Founders were screened using CreERT2 specific primers:

(F: 5'- TGGGCCAGCTAAACATGCTTC-3'; R: 5'-GCAGCTCTCATGTCTCCAGCAG-3'). CreERT2 activity of positive founders was screened by crossing to *Rosa26-LacZ* reporter mice and assaying β-galactosidase activity using X-gal (Soriano, 1999). To induce nuclear Cre activity, tamoxifen (in corn oil, 5mg/40 g body weight) was administered orally to pregnant dams at E9.5 or E10.5. Embryo dissection, sectioning and *LacZ* expression staining at E12.5 were performed according to standard protocols.

Mouse Husbandry and Lines

Mouse colonies were maintained at the University of California, San Francisco, in accordance with National Institutes of Health and UCSF guidelines. For staging of embryos, midday of the vaginal plug was calculated as embryonic day 0.5 (E0.5). Stable transgenic mice were maintained on *CD1* background (Jackson Laboratories). *Ai14* reporter mice (Madisen 2010) and *Pax6* mutant line (Sey; Hill, 1991) was maintained on *C57BL/6J*. Stable transgenic mice were crossed on to *Ai14* reporter and *Pax6* mutant background for further analysis.

Tamoxifen Independent Cre Activation

To determine the fidelity of the tamoxifen control of CreER^{T2} activity, we assayed *Ai14 (tdTomato)* activation at E17.5 independent of tamoxifen (Figure S2). For 9/14 lines we found very little or no tdTomato expression (Table S2), demonstrating tight temporal control of CreER^{T2} activity mediated by tamoxifen. For 5 lines, we detected some recombination independent of tamoxifen (192, 643, 840, 1006, 1172), although the recombination pattern did not extend beyond the boundaries defined by the E10.5-E17.5 fate map analysis. 4/5 of the lines that had tamoxifen-independent Cre activity had the strongest E11.5 GFP expression, suggesting that high levels of CreER^{T2} lead to its entry into the nucleus, independently of estrogen/tamoxifen treatment (Liu et al., 2010).

Thus, because the GFP expression domains at E12.5, E14.5 and E17.5 showed largely the same regional pattern (with respect to the regions defined by E11.5 GFP expression), and because the pattern of CreER^{T2} activation (tdTomato expression) independent of tamoxifen largely did not extend beyond the boundaries defined by the E10.5-E17.5 fate mapping, we have confidence that the fate maps reflect enhancer activity at E11.5.

Analysis of Enhancer GFP Activity and Fate Mapping

Enhancer transgenic lines were crossed with CD1 strain mice; embryos were collected at a variety of ages for examination of GFP expression by three methods: 1) immunohistochemistry (chicken anti-GFP, Aves); 2) fluorescent GFP imaging; or 3) in situ hybridization targeted to GFP. For fate mapping, enhancer transgenic lines were crossed to *Ai14* Cre-reporter mice that express tdTomato. To promote increased survival a reduced dose of tamoxifen (3mg/40 g body

weight) was administered at E9.5 or E10.5. Embryos or adults were collected at a variety of ages; sections were assayed by immunohistochemistry (rabbit DsRED or rat DsRED, rabbit Ctip2, and goat Nurr1). For the adult fate mapping, enhancer transgenic mouse on the Ai14 background was crossed to CD1 females (this promoted increased survival). Anatomical assignments of regional GFP expression and fate maps (tdTomato expression) were made independently by at least three, of four, investigators (KP, OG, JR and LP).

Immunohistochemistry

Immunohistochemistry was performed as previously described (Flandin, 2010) using rabbit (Clontech) or rat (Chromotek) antibodies for DsRED, rabbit anti-Ctip2 (Abcam), goat anti-Nurr1 (R&D Systems), and goat anti-Prox1 (Gift from Sam Pleasure) and rabbit anti-PH3 (Millipore).

In Situ Hybridization

RNA *in situ* hybridization was performed as previously described (Jeong et al., 2008). Riboprobes were provided by M.-J. Tsai (Nr2f1), R. Grosschedl (Lef1), S. Retaux (Lmx1a), and S. Pleasure (Wnt8b). Fluorescent In situ/immunohistochemistry was performed as described by Jeong et al., 2008, with the following modifications. Triton X was used instead of Proteinase K. GFP antibody (chicken anti-GFP, Aves) was added overnight along with Anti-Digoxigenin Antibody. GFP signal was amplified using a biotinylated goat anti-Chicken (Vector Labs) and Streptavidin Alexa 488 (Invitrogen). The in situ was developed using the HNPP Fluorescence Detection set (Roche).

Image acquisition and analysis

Fluorescent images were taken using a Coolsnap camera (Photometrics) mounted on a Nikon Eclipse 80i microscope using NIS Elements acquisition software (Nikon). Brightness and contrast were adjusted and images merged using Adobe Photoshop software.

Identification of Region Specific Motifs

Enhancer datasets

Experimentally assayed MP and non-MP enhancers (n=9 and 15 respectively) were separated according to the reporter gene expression patterns driven. In addition, for each of the 24 enhancers we sampled 20 random sequences from the human genome, with matching (within a 5% error) length, GC- and repeat content. Finally, we included in the analysis a set of 765 human sequences from the VISTA browser that were negative for enhancer activity (Visel et al., 2007). On average, MP enhancers are 1155 bp-long, and non-MP enhancers are 1389 bp-long. In total, control sequences randomly sampled from the genome have an average length of 1301 bp. Finally, the sequences extracted from the VISTA browser that were negative for enhancer activity had an average length of 1495 bp. The sequences were not examined for minimal enhancer activity, and presumably include flanking sequences in addition to the core enhancer.

Enhancer representation

Enhancers were transformed into 1064-dimensional feature vectors, where each feature corresponds to a binding site in the TRANSFAC (Matys et al., 2006) database. Significant occurrences of binding sites in the sequences were determined with MAST (Bailey and Gribskov, 1998). Each feature represents the number of occurrences of a given binding site per base pair of sequence.

Random forest classifier

A random forest (RF) trains a set of decision trees on subsets of features. Each tree in the forest assigns a class to each of the enhancers. The final classification of a given enhancer is decided by a simple majority vote. In the construction of the decision tree, a subset of n out of the total N features are randomly selected at each split, and the feature with maximum information gain out of the n is used to split the node. We constructed a RF with 500 decision trees, and randomly selected 10 out of the total 1064 features to split the nodes. We used the RF implementation from the 'randomForest' R package (Liaw and Wiener, 2002). During the construction of a RF, the out-of-bag (OOB) data, approximately one-third of the enhancers, are then used to estimate the prediction accuracy. Small classification errors would indicate classes of enhancers with strong tissue-specific signatures (Narlikar et al., 2010). We trained two different models. Both models were trained to distinguish between three classes of sequences. In the case of the first model, the classes consisted of 9 MP enhancers, 15 non-MP enhancers, and 480 random genomic sequences with matching length, GC- and repeat content. For this classifier we obtained an OOB estimate of the error rate of 9.92%. The second model was trained to distinguish between 9 MP enhancers, 15 non-MP enhancers, and a set of background sequences consisting of 480 random genomic sequences with matching length, GC- and repeat content and 765 sequences from the VISTA browser that were negative for enhancer activity. For this classifier we obtained an OOB estimate of the error rate of 11.72%.

Motif rankings

For each of the models, we used the RF to rank the importance of the binding sites for the classification of the different groups of sequences. RF classifiers can be used to rank the importance of variables in a natural way, as described by [Breiman, 2001]. Basically, to measure the importance of the *j*th variable, the values of the *j*th variable are permuted among the training data, and the OOB error is computed. This OOB error is then compared with the OOB error obtained for the original data. This is repeated several times. The importance of the

variable is defined as the average of the differences between the OOB error on the permuted and original data, divided by the standard deviation of these differences.

E11.5 expression analysis

Expression patterns of the top TF motifs were analyzed using the E11.5 developing mouse ISH data in the Allen Brain Atlas (http://developingmouse.brain-map.org). The subdomains of the E11.5 pallium were divided according to E11.5 sagital reference atlas and previously described gene expression patterns.

Chromatin immunoprecipitation (ChIP)

ChIP was performed using Pax6 (Millipore), CoupTf1 (R&D systems), and Pbx1/2/3 (C-20, Santa Cruz Biotechnology) similarly as in a published method (McKenna et al., 2011) with few modifications. Briefly, E12.5 or E13.5 CD1 cortices were dissected, fixed in either 1% (Pbx1/2/3) or 1.5% (Pax6 and CoupTf1) formaldehyde for 10 min. (Pbx1/2/3) or 20 min (Pax6 and CoupTf1) and neutralized with glycine. Fixed chromatin was lysed and sheared into 200-1000 bp fragments using a bioruptor (Diagenode). Immunoprecipitation (IP) reactions were performed in duplicates using either Mouse IgG (CoupTf1; Millipore 12-371) or 200x molar fold blocking peptide (Pbx1/2/3, sc-888, Santa Cruz Biotechnology; Pax6, AG983, Millipore) as negative controls. Precipitated fractions were purified using Dynabeads (Invitrogen). Libraries were prepared using an Ovation Ultralow DR Multiplex System (Nugen), size selected in the range of

200-300 bp on a LabChip (Lifesciences) and lastly quality control tested on a Bioanalyzer (Agilent). Information about quality of the sequence mapping and peak calling is provided in Table S7.

Readouts were aligned to the mouse genome (mm9) using the BWA aligner (PMID: 19451168).

Call: bwa aln -t 6 -l 25 mm9 sample.fastq.gz

Peaks were called using MACS (PMCID: PMC2592715): call: macs14 -t chip.bam --control=input.bam –name=chip_output --format=BAM --gsize=mm --tsize=50 --bw=300 --mfold=10,30 --llocal=20000 --shiftsize=100 --pvalue=1e-5.

Peak calling was performed using both matched input and IgG as control samples. After peak calling, enriched regions were filtered to remove peaks where a significant proportion of the reads originated in repetitive regions of the genome. Peaks were mapped to UCSC genome browser (mm9), and pallial enhancers were screened for called peaks. The pallial enhancers were organized by regional activity (Table S3). The peak occupancy percentage was compared between regions using a non-parametric binomial test.

Luciferase Assay

P19 cells (Farah et al., 2000) in 24-well plates were transfected with 400ng of DNA when there were $5*10^5$ cells per well. Of the total 400ng of DNA, we used 200ng *pCAGGs* (empty) or *pCAGGs-Pax6*, and 200ng Promega *pGL4.23* luciferase reporter (empty) or containing an enhancer element upstream of the luciferase gene (pGL4.23-enhancer). In addition, 0.5ng of Promega *pGL4.73*-Renilla luciferase was added to measure transfection efficiency, to enable the normalization calculation. Each condition was assayed in triplicates. Relative luciferase

expression was normalized to measurements from *pGL4.23-enhancer+pCAGGS-empty* co-transfections, and is reported as fold-change relative to transfections of *pGL4.23-empty+pCAGGs-TF*. One-way ANOVA with a tukey post-hoc test was used to compare activation of the different enhancers.

References

Bailey, T.L., and Gribskov, M. (1998). Combining evidence using p-values: application to sequence homology searches. Bioinformatics. 14(1), 48-54

Breiman, L. (2001). Random Forests. Machine Learning. 45, 5-32.

Hill, R.E., Favor, J., Hogan, B.L., Ton, C.C., Saunders, G.F., Hanson, I.M., Prosser, J., Jordan, T., Hastie, N.D., and van Heyningen, V. (1991). Mouse small eye results from mutations in a paired-like homeobox-containing gene. Nature. 354, 522–525.

Feil, R., Wagner, J., Metzger, D., and Chambon, P. (1997). Regulation of Cre recombinase activity by mutated estrogen receptor ligand-binding domains. Biochem. Biophys. Res. Commun. 237, 752–757.

Liaw, A., and Wiener, M. (2002). Classification and regression by random forest. R News. 2,18-22.

Liu, Y., Suckale, J., Masjkur, J., Magro, M.G., Steffen, A., Anastassiadis, K., and Solimena, M. (2010). Tamoxifen-independent recombination in the RIP-CreER mouse. PLoS One. 5, e13533.

Matys, V., Kel-Margoulis O.V., Fricke, E., Liebich, I., Land S, Barre-Dirrie, A., Reuter, I., Chekmenev, D., Krull, M., Hornischer, K., et al. (2006). TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. Nucleic Acids Res. 34, D108-110.

Narlikar, L., Sakabe, N.J., Blanski, A.A., Arimura, F.E., Westlund, J.M., Nobrega, M.A., Ovcharenko, I. (2010). Genome-wide discovery of human heart enhancers. Genome Res. 3, 381-392

Soriano, P. (1999). Generalized lacZ expression with the ROSA26 Cre reporter strain. Nat. Genet. 21, 70-71.

Visel, A., Minovitsky, S., Dubchak, I., Pennacchio L.A. (2007). VISTA Enhancer Browser—a database of tissue-specific human enhancers. Nucleic Acids Res.35, D88-92

Chapter 4: Conclusions and Future Directions

The focus on understanding cortical patterning has been on the role of individual transcription factors. We aimed to further our understanding of cortical patterning by identifying TF networks involved in patterning. We approached this by trying to identify other TFs, cis-regulatory elements, and TF interactions involved in patterning. We have made progress in each of these areas, as we have identified at least 40 novel TFs, several enhancers with cortical activity, and interactions between several TFs specifically Pax6 and CoupTf1.

We propose that sharp boundaries in the adult cortex are set up in the ventricular zone by cortical enhancers. These cortical enhancers integrate gradients of transcription factors, which are set up by patterning centers in the telencephalon. Our work provides support for this concept. However, more work required to further understand the mechanism of cortical patterning. Most importantly, further studying the role of these enhancers in development is important. This can be achieved by studying the affect of removing these enhancers on patterning. In addition, identifying the gene or genes regulated by these cortical enhancers will further elucidate their role in development and possibly another gene important to cortical patterning. This can be accomplished by using chromosome conformation capture (3-C). Understanding how these enhancers form sharp boundaries in the VZ would also be extremely interesting. This requires identifying the TFs that bind each enhancer. We unsuccessfully attempted to answer this question by using an *in vitro* assay and mass spectrometry to identify the proteins that bind. In this thesis, we used TF ChIP-seq to study which enhancers are bound by Pax6 and CoupTf1. Theoretically, we could use ChIP-seq to analyze where every cortically TF binds. However, this is extremely expensive and time consuming, and there are not antibodies for most TFs. Alternatively, we could mutate the sequence of the enhancer and identify domains important for boundary formation. This is still expensive and time-consuming, but technically feasible (De Val et al. 2008). Through this iterative process, we can

narrow down these larger enhancer sequences into the functionally important core domain. These proposed experiments are large-scale projects.

This work also highlighted that genetic interaction of Pax6 and CoupTf1. In the ventral cortex, Pax6 is important for maintaining expression of CoupTf1/2, while CoupTf1/2 repress Pax6. This network fits the previously described findings that manipulate levels of Pax6 (citations), as well as work on the CoupTf1 (Faedo et al. 2008). Generally in the cortex, Pax6 and CoupTf1 binding sites are enriched for regions near genes important for nervous system and cortical development. In addition, there is overlap between Pax6 and CoupTf1, and these genomic regions are further enriched for TFs and nervous system development. Analysis of TF binding sites in these regions showed a lack of Pax6 paired domain and CoupTf1 nuclear receptor domain, but the presence of homeobox binding sites. Further analysis of this complex would be interesting in understanding how patterning TFs work together in the brain.

Understanding cortical patterning has relevance beyond development. Both autism and schizophrenia are diseases of brain development. Postmortem RNA expression profiling of 17 cortices from autistic patients and age-matched controls showed reduced differences between the frontal and temporal cortices (Voineagu et al. 2011). This change could be attributed to changes in patterning. In addition, with the advance in stem cells and IPSCs, there is interest in therapies involving cell transplantation. In the brain, cell transplantation could be used to repair damaged cortices after traumatic brain injury or stroke. However, the milieu of the adult brain is very different than the developing brain. Thus, transplantation of undifferentiated cells will most likely be able to correctly integrate into the damaged cortex. Understanding how the brain develops and patterns can be harnessed to improve integration and bias for the correct identity.

Bibliography

De Val S, Chi NC, Meadows SM, Minovitsky S, Anderson JP, Harris IS, Ehlers ML, Agarwal P, Visel A, Xu SM, Pennacchio LA, Dubchak I, Krieg PA, Stainier DY, Black BL. (2008) Combinatorial regulation of endothelial gene expression by ets and forkhead transcription factors. Cell. 135(6):1053-1064

Faedo, A., Tomassy, G.S., Ruan, Y., Teichmann, H., Krauss, S., Pleasure, S.J., Tsai, S.Y., Tsai, M.J., Studer, M., and Rubenstein, J.L. (2008). COUP-TF1 coordinates cortical patterning, neurogenesis, and laminar fate and modulates MAPK/ERK, AKT, and beta-catenin signaling. Cereb. Cortex. 18, 2117-2131.

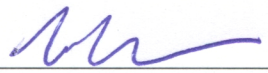
Voineagu I, Wang X, Johnston P, Lowe JK, Tian Y, Horvath S, Mill J, Cantor RM, Blencowe BJ, Geschwind DH. (2011) Transcriptomic analysis of autistic brain reveals convergent molecular pathology. Nature. 474(7351), 380-384

Publishing Agreement

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

Please sign the following statement:

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.



Author Signature

3/28/14.

Date