

# Transcriptional programs in transient embryonic zones of the cerebral cortex defined by high-resolution mRNA sequencing

Albert E. Ayoub<sup>a</sup>, Sunghee Oh<sup>b</sup>, Yanhua Xie<sup>a</sup>, Jing Leng<sup>b,c</sup>, Justin Cotney<sup>b</sup>, Martin H. Dominguez<sup>a</sup>, James P. Noonan<sup>b,c,d</sup>, and Pasko Rakic<sup>a,d,1</sup>

Departments of <sup>a</sup>Neurobiology and <sup>b</sup>Genetics, Yale University School of Medicine, New Haven, CT 06520; and <sup>c</sup>Program in Computational Biology and Bioinformatics and <sup>d</sup>Kavli Institute for Neuroscience, Yale University, New Haven, CT 06520

Contributed by Pasko Rakic, August 1, 2011 (sent for review June 20, 2011)

**Characterizing the genetic programs that specify development and evolution of the cerebral cortex is a central challenge in neuroscience.** Stem cells in the transient embryonic ventricular and subventricular zones generate neurons that migrate across the intermediate zone to the overlying cortical plate, where they differentiate and form the neocortex. It is clear that not one but a multitude of molecular pathways are necessary to progress through each cellular milestone, yet the underlying transcriptional programs remain unknown. Here, we apply differential transcriptome analysis on microscopically isolated cell populations, to define five transcriptional programs that represent each transient embryonic zone and the progression between these zones. The five transcriptional programs contain largely uncharacterized genes in addition to transcripts necessary for stem cell maintenance, neurogenesis, migration, and differentiation. Additionally, we found intergenic transcriptionally active regions that possibly encode unique zone-specific transcripts. Finally, we present a high-resolution transcriptome map of transient zones in the embryonic mouse forebrain.

radial glia | pyramidal neurons | cortical development |  
laser microdissection

**O**ne of the main characteristics of the developing neocortex is the generation of various cellular zones, most of which begin to disappear by late gestation (1, 2). During corticogenesis, progenitor cells in the ventricular zone undergo two distinct modes of division (3, 4). Symmetric divisions expand the surface area of the cortex by increasing the number of progenitors, whereas asymmetric divisions produce intermediate progenitors (INPs) or postmitotic neurons (PNs) that pass through the subventricular zone (SVZ) and intermediate zone (IZ), respectively (5, 6). These newly generated daughter cells then produce the outermost zone, the cortical plate (CP), where neurons align in an inside-out pattern to form the six-layered cerebral cortex. Recent studies highlight the dependence of normal corticogenesis on genetic programs involving a multitude of genes and that the disruption of these genetic programs is thought to cause the majority of brain disorders and malformations (7–9).

Significant progress in developmental neuroscience has often been tied to technological breakthroughs and creative use of sophisticated techniques (10). The past decade witnessed significant advances in the developmental neuroscience field as a result of the increasing adoption of genome-wide approaches. For example, global analyses of gene expression with the use of microarrays have identified genes involved in progenitor cell proliferation (11) and specification of cortical layers (12, 13). Although several studies have divided the developing brain into functional areas to study arealization (14, 15) and gene networks (16), to date we are aware of no study that has determined the full repertoire of transcripts necessary for the progression of cells from birth to differentiation.

Whole-transcriptome mRNA sequencing (mRNA-seq) is a breakthrough technology that provides quantitative and sen-

sitive detection with a broad dynamic range (17), which was recently demonstrated by Han and colleagues (18). This study described global gene and alternative isoform expression using mRNA-seq on mouse neocortex at late embryonic (embryonic day [E] 18) and early postnatal (postnatal day [P] 7) ages. It became clear that overcoming limits in spatial resolution is a necessary step to determine transcriptional programs that define the maintenance of neural stem cells in the VZ, the migration of neurons through the SVZ–IZ, and finally the differentiation of neurons in the CP. Therefore, we performed differential expression analysis after mRNA-seq on cells collected by laser microdissection (LMD) from each transient embryonic zone. We characterized the full repertoire of zone-specific transcribed genes, alternative isoforms, and novel transcriptionally active regions (nTAR) at a level of resolution that has not, to the best of our knowledge, been previously achieved. By using this approach, we were able to expose genetic programs that define milestones of corticogenesis.

## Results

**High-Resolution Differential Gene Expression.** Resolving expression differences among small cellular groups and detecting genes expressed at low levels are common limitations of whole-cortex studies. To overcome these obstacles, we performed mRNA-seq on RNA isolated from small cellular groups by LMD (Fig. 1A). We determined the abundance of genes by calculating reads per kilobase of exon model per million mapped reads (RPKM) values for each annotated isoform, and summed all isoform-level RPKMs for each gene (19, 20). RPKM values were highly correlated for all cellular zones, irrespective of sequencing lanes or flow cells (Spearman correlation > 0.92; Fig. 1C). Principal component analysis (PCA) indicates that the primary source of variance in our mRNA-seq data are caused by expression differences among transient embryonic zones, with biological or technical replicates for each zone likely contributing little additional variance (Fig. 1D and Figs. S1–S3; PCA values in Table S1). These exploratory analyses demonstrate that our LMD-mRNA-seq approach preserves the spatial resolution of gene expression in the developing cortex.

We assessed differential expression after statistical testing of pooled data sets by using a log-linear Poisson model, which has been shown to be highly effective at identifying differentially expressed genes (DEGs) in mRNA-seq data (21, 22). We identified a total of 2,677 DEGs upon imposing a conservative

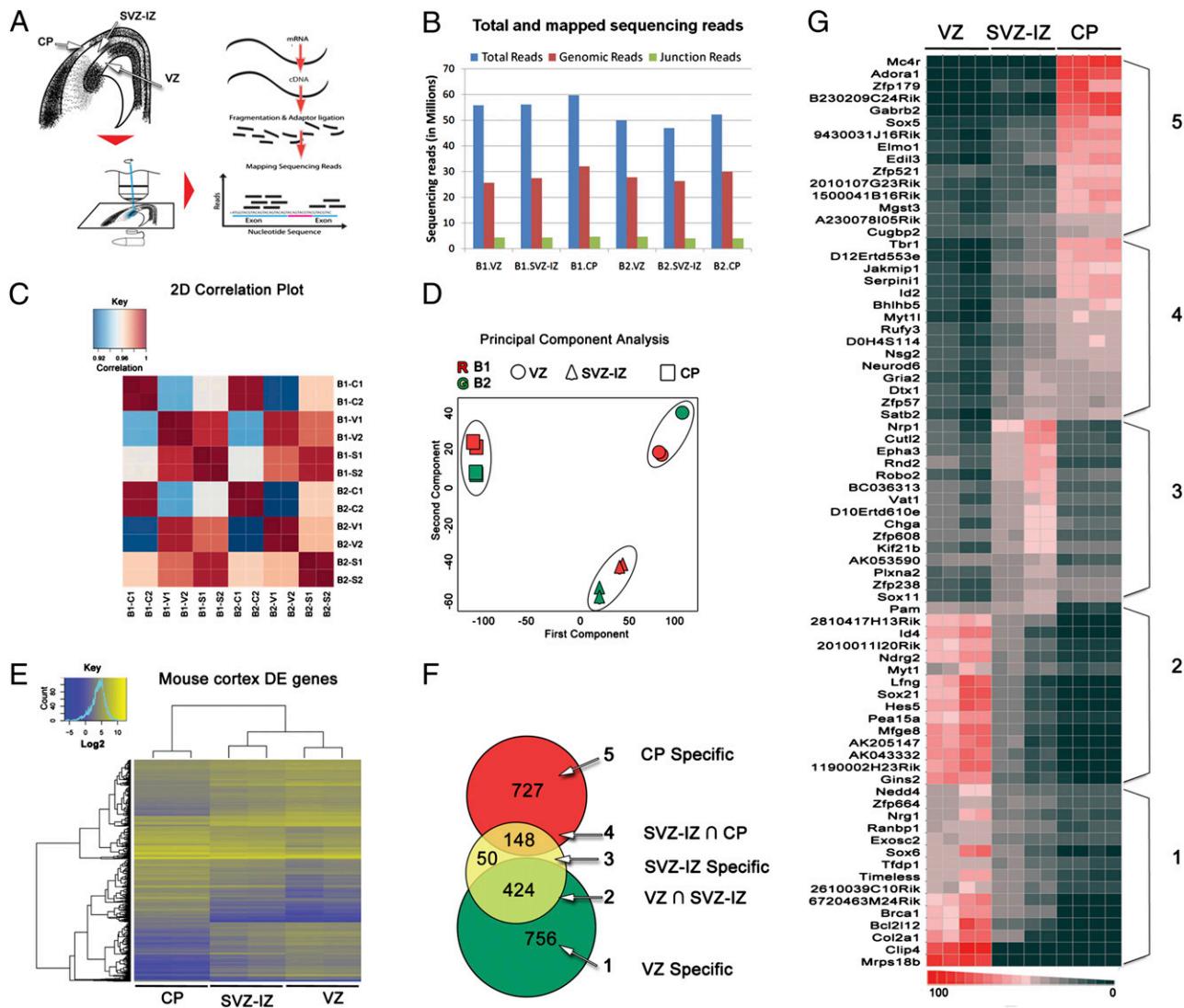
Author contributions: A.E.A. and P.R. designed research; A.E.A., Y.X., J.C., and M.H.D. performed research; J.P.N. contributed new reagents/analytic tools; A.E.A., S.O., J.L., and J.P.N. analyzed data; and A.E.A., J.P.N., and P.R. wrote the paper.

The authors declare no conflict of interest.

Data deposition: The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, [www.ncbi.nlm.nih.gov/geo](http://www.ncbi.nlm.nih.gov/geo) (accession no. [GSE30765](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE30765)).

<sup>1</sup>To whom correspondence should be addressed. E-mail: [pasko.rakic@yale.edu](mailto:pasko.rakic@yale.edu).

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1112213108/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1112213108/-DCSupplemental).



**Fig. 1.** Determination of zone-specific differential expression. (A) Schematic shows experimental steps: we used LMD to isolate each of the cellular zones (VZ, SVZ-IZ, and CP) from E14.5 brains in coronal sections chosen by systematic random sampling. We performed single-end 75-bp RNA sequencing on the Illumina Genome Analyzer IIx. (B) Graph shows reads uniquely mapped to the mouse genome and splice junctions compared with the total number of sequencing reads that passed quality checks. (C) Two-dimensional correlation plot of gene expression levels between sequencing lanes. Spearman  $R^2$  for each technical replicate and zone are color-coded from 0.92 (blue) to 1 (red). Labels starting with "B1" and "B2" indicate biological replicas 1 and 2 respectively; C, CP; S, SVZ-IZ; V, VZ. (D) Principal components analysis plot of  $\log_2$ -RPKM for 12 samples confirms that the correlation of technical replicates is higher than biological replicates, which in turn is much higher than different tissue samples. (E) Unsupervised hierarchical clustering of DEGs. RPKM values were  $\log_2$ -transformed and clustered using the gplots and heatmap2 packages in R. The color scale (Top Left) indicates levels of expression, with yellow corresponding to high expression levels and blue corresponding to low. (F) Venn diagram shows zone-specific gene expression after pairwise comparisons. We applied a Benjamini-Hochberg  $P$  value ( $bhp < 10^{-5}$ ) and a twofold change in expression to determine DEGs. (G) Heat map shows the fractional change in expression for a subset of DEGs in each of the five groups in B. Each row represents the fraction of expression from all flow cells for a particular gene (red is highest; black is lowest).

$P$  value ( $<10^{-5}$ ) and a twofold change in expression ( $\log_2 > 1$ ) between at least two cellular zones. We were able to isolate five groups of DEG, three of which were specific to only one zone (groups 1, 3, and 5; Fig. 1F). Two additional groups (groups 2 and 4) were differentially expressed in two zones versus the third (VZ  $\cap$  SVZ-IZ and SVZ-IZ  $\cap$  CP), whereas the intersection of VZ- and CP-specific genes did not return any common DEG. Furthermore, we performed unsupervised clustering of all DEGs to identify mRNA specific to each neocortical zone, which supported our differential expression analysis (Fig. 1E). The expression patterns across the cortex for a representative set of genes from each group are shown in Fig. 1G.

**nTARs Are Differentially Distributed Among Embryonic Zones.** Transcriptional activity has been reported from regulatory elements

located in intergenic regions in cultured E16.5 neurons (23) but not *in vivo*. Our data contained a significant pool of reads that mapped outside of annotated genes. We defined nTARs based on contiguous coverage and zero overlap with all known genes and 3'UTRs (Materials and Methods). After quantification, we analyzed the differential distribution of nTARs using the same statistical methods described later. We found 1,621 differentially expressed nTARs (Table S2) after pairwise comparisons between neocortical zones. We analyzed nTARs for significant overlap with known genomic elements such as enhancers, promoters, and remaining intergenic regions including gene deserts. Only two nTARs overlapped with suspected forebrain enhancers, near *Ngn2* and *Fzd8*, respectively, and one overlapped with a known limb enhancer (24), suggesting that nTARs in our data are less likely transcribed from enhancers. Although 172 nTARs over-

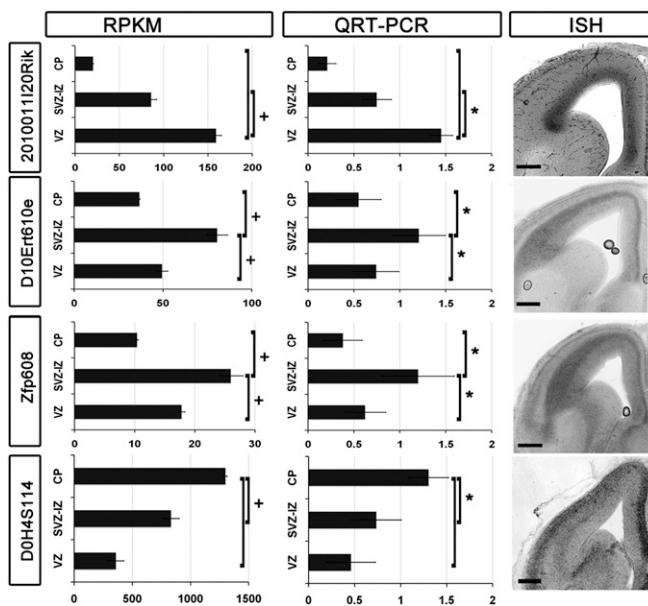
lapped with promoters near zone-specific genes, the majority of nTARs mapped to intergenic areas excluding gene deserts, suggesting the existence of putative unique transcripts showing zone-specific dynamic expression in the mouse neocortex.

**Verification of Zone Specific Expression.** We chose several zone-specific genes to validate expression levels determined by mRNA-seq. First, we carried out quantitative real time RT-PCR (qRT-PCR) on RNA isolated from each cellular zone by LMD following the same procedures used for mRNA-seq. We performed *in situ* hybridization (ISH) to map the histological distribution of these genes. The level of gene expression predicted by our analysis of mRNA-seq data were concordant with levels determined by qRT-PCR and tissue distribution determined by ISH (Fig. 2 and Fig. S4). To extend our validation process, we also turned to publicly available resources—the Allen Brain Atlas (ABA) and GenePaint (GP). With the exception of novel genes and transcripts with multiple splice isoforms, all 60 targets we searched for in public databases showed tissue distribution similar to that predicted by our mRNA-seq expression analysis.

**Functional Analysis of Zone-Specific Genes.** To gain insight into the biological processes and pathways enriched in each transitional zone during corticogenesis, we used each of the five groups of DEGs (Fig. 1E) as input to perform Gene Ontology (GO) analysis (Table S3). The most enriched categories in the VZ were related to cell division, which included genes such as cyclins and cyclin-dependent kinases in addition to genes affecting stem-cell renewal such as *Ctnnb1*, *Notch1*, and *Spread1*. Notably, genes involved in transcriptional regulation, such as *Pax6*, *Sox2*, and *Hes1*, were more prominent in group 2 (VZ  $\cap$  SVZ–IZ). The top categories of transcripts enriched in the SVZ–IZ included axonal guidance and cell migration genes such as *Nrp1*, *Plxna2*, and *Dcx*. Genes contributing to axonogenesis and cell migration represented the largest fraction in group 4 (SVZ–IZ  $\cap$  CP), including *Cdk1*, *Satb2*, and *Cdk5rl1*. Analysis of the CP transcriptome showed overrepresentation of functional categories relating to synaptogenesis, exocytosis, and axonogenesis that

included genes such as *Cpxl1*, *Nav1*, and *Fezf2*. We also investigated overrepresented canonical pathways in addition to GO categories. In the VZ, the most prominent pathways centered on cell cycle regulation, transition through the cell cycle, and termination of the cell cycle. Pathways enriched in the SVZ–IZ included ephrin-dependent axonal guidance, extracellular  $\text{Ca}^{2+}$ - and nicotine-dependent inhibition of cell cycle progression, as well as regulation of Erk signaling. In the CP, the most prominent pathways were GABA-A receptor trafficking, Reelin/CDK5-mediated neuronal growth, and PKA signaling.

**Self-Organizing Maps Reveal Functional Insights Into Novel Zone-Specific Genes.** Our analysis generated a significant pool of novel, zone-specific genes that were excluded from GO and pathway analyses because of missing annotations. To explore potential relationships between known and novel genes, we used AutoSOME to isolate modules of coexpressed genes (25). First, we performed cluster analysis by comparing expression levels by using the five groups of DEGs (Fig. 1E) as input. We then queried the GeneMania (26) database of functional data to identify networks based on self-organizing map (SOM) modules (Table S4). We analyzed stable modules at 45% confidence level and  $P \leq 0.05$  (Monte Carlo  $P$  value). For example, we examined a module in group 2 (VZ  $\cap$  SVZ–IZ) that contained *Sox2* and *Pax6* (Fig. 3A), which have been shown to affect the balance of stem cells and neurogenesis (27). Network analysis shows that *Sox2* and *Pax6* have first-order interactions with *Hes5*, a downstream effector of Notch, and *Sox21*, as well as potential second-order interaction with *Nde1* and *Mfge8* (Fig. 3A). We determined potential interactions between the novel SVZ–IZ gene *D10ERT610E* (Fig. 2), and other module members *Nrp1*, *Sema3c*, and *Plxna2* that have been shown to play a role in axonal guidance (Fig. 3B). Finally, we assembled a SOM-based network including the novel CP-specific gene *B230209C24RIK* (Fig. 3C), which uncovered potential interactions with genes involved in neurotransmitter trafficking and release such as *Cpxl1* and *Dync1i1* (Fig. 3C). However, the full benefit of gene networks based on SOM modules that contain novel genes can only be achieved with functional studies.



**Fig. 2.** Verification of zone-specific genes. A row of three panels is shown for every gene named on the left side. The first panel shows mean RPKM expression levels ( $\pm$ SEM). Significance is based on pairwise comparisons and  $bhp < 10^{-5}$  (+). The second column of graphs shows qRT-PCR on RNA isolated by LMD from E14.5 cryosections. Mean fold change ( $\pm$ SEM) is plotted for every cellular zone (\* $P < 0.001$ ). The third column shows ISH on E14.5 mouse sections.

**Contribution of Alternative Splicing to Corticogenesis.** Although it is presently difficult to gauge the full contribution of alternative splicing to corticogenesis, its importance cannot be dismissed, as studies on *Pax6* and *Tnc* have demonstrated (28, 29). Differential alternative splicing has also been observed in different areas of the normal and diseased human cortex (15, 30). To elucidate how alternative splicing contributes specifically to gene-level expression and globally to corticogenesis, we explored the extent of alternative isoform use by examining isoform-level RPKM values between cortical zones. First, we calculated the relative abundance of transcripts using Cufflinks (20), and then analyzed differentially expressed isoforms (DEIs; Fig. 4A) following the same statistical workflow mentioned earlier. Overall, the proportion of DEGs with two or more isoforms was highest in the CP (15.7%) compared with the VZ (11.8%) and the SVZ–IZ (12.8%; Table S2).

We separated zone-specific splice isoforms into six groups based on their differential expression in at least one zone compared with the others (Fig. 4A), which highlighted differences between the distribution of DEIs and DEGs. Even though the contribution of splice isoforms to gene-level expression is additive, gene-level expression is confounded by switching events in which an isoform is dominant in one zone but negligible in another. For example, individual splice variants of *Wdr61*, which falls in group 6 (Fig. 4A), are differentially expressed (Fig. 4B), but *Wdr61* is not differentially expressed at the gene level given our cutoff criteria ( $P < 10^{-5}$  and  $\log_2 > 1$ ). Switching of dominant isoforms does not always result in subthreshold differential expression. For instance, of three expressed isoforms for *Mfge8*, which falls in group 2 (VZ  $\cap$  SVZ–IZ) according to our analysis, two are differentially expressed but with different zonal distribution. Isoform 1 of *Mfge8* (UC009hyb.1)

**Fig. 3.** Network analysis of zone-specific SOM modules. Networks were generated by using the GeneMania large functional association database based on SOM module membership. Edges connecting different nodes (i.e., genes) represent data derived from other studies. (A) Representative cluster from  $VZ \cap SVZ$ -IZ highlighting all edges emanating from *Sox2* and *Pax6* to other members of the SOM module. (B) Network from SVZ-IZ module highlighting all edges connecting the novel gene *D10Ert610e* to *Nrp1* and *Sema3c*, which are involved in axonogenesis. (C) SOM-based network from CP-specific modules shows extensive connections between the novel gene *B230209C24RIK* and *Cpx1* and *Dync1i1*, which are involved in exocytosis.

is highest in the VZ whereas isoform 2 (UC009hya.1) is highly expressed in the VZ and SVZ-IZ versus the CP (Fig. 4D). Inspection of *Mfge8*'s histological distribution on ABA and GP revealed VZ-only expression. To investigate this discrepancy, we tested the expression of both isoforms *in situ*, which confirmed the distribution predicted by our expression analysis. However, for other DEGs with multiple isoforms, such as of *B230209C24RIK*, *Cugbp2*, and *Hes6* (Fig. 4), the expression of splice isoforms changed concomitantly in every zone.

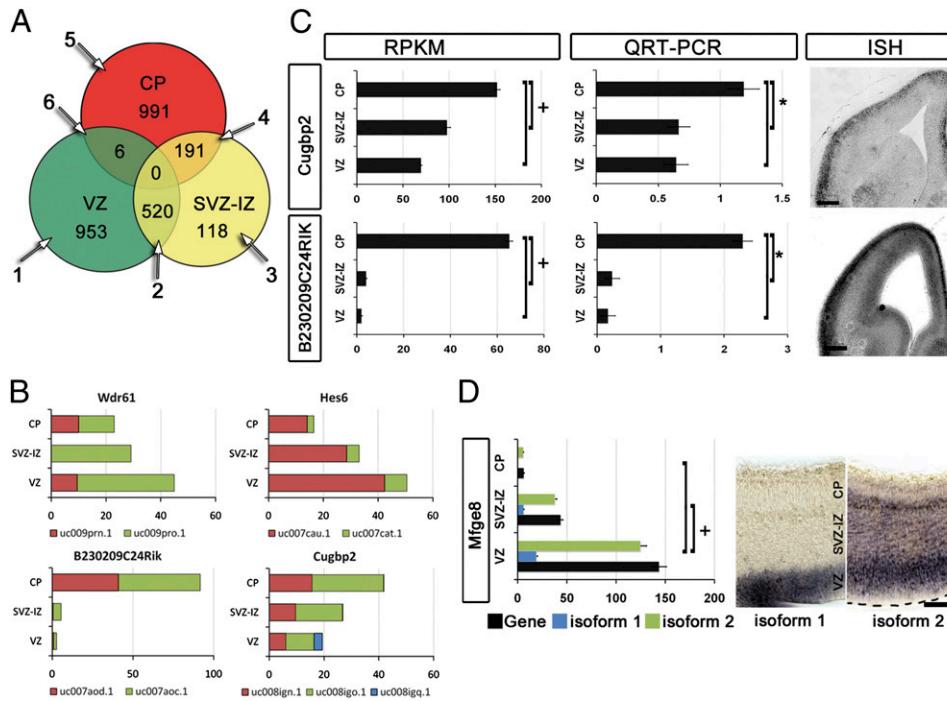
**Zone-Specific Transcriptional Regulation.** To better understand dynamic developmental events, we sought to identify transcriptional enhancers located near cortex DEGs. Both extreme evolutionary conservation and analysis of the genome-wide binding patterns of the enhancer-associated factor p300, followed by mouse transient transgenic assays, have identified enhancers active in the E11.5 mouse forebrain (24). However, there is little information available about zone-specific enhancers in the cortex or whether enhancers active in E11.5 cortex regulate gene expression at E14.5. With these caveats in mind, we computationally identified candidate zone-specific p300-enriched forebrain enhancers (31) nearest to the transcription start site of DEGs in the E14.5 cortex. We found 181 DEGs with nearby candidate enhancers (Table S5). We also found several verified enhancer elements near DEGs (e.g., *Sox21* and *Tle4*) on the Vista Enhancer Browser. In contrast to alternative

promoter use and alternative splicing, enhancer regulation of zone-specific differential expression is a level of genomic regulation that is yet to be empirically explored in the developing neocortex.

## Discussion

A fundamental feature of the developing neocortex is that the fate and layer specification of neurons is defined at the time of their birth from neural stem cells in the proliferative VZ and SVZ (1). The transcriptional content of cells as they progress from neural stem cells to differentiated neurons remains to be determined. Although differential expression contributing to the arealization of the neocortex has been described by several groups (14, 15), to date it has not been performed with next-generation sequencing or at the resolution of cellular zones. Here, we provide a comprehensive transcriptome analysis of the birth-to-differentiation axis in vivo with an unprecedented resolution by combining two technologies, mRNA-seq and LMD. The data go beyond a traditional genome-scale screen of genes and splice variants by providing the distribution of nTARs enriched in each embryonic zone and putative *cis*-regulatory elements near DEGs.

In the present study, we collected cells from the neocortex by LMD to better resolve the genetic basis of neuronal birth, migration, and differentiation. As expected, we were able to resolve the transcriptome for each cellular zone at the transcript



**Fig. 4.** Distribution and verification of splice isoforms. (A) Venn diagram shows the distribution of splice isoforms in six groups after pairwise comparisons,  $\log_2$  fold change, and  $bhp < 10^{-5}$ . (B) The expression levels of different isoforms for the same gene are additive. However, isoforms for *Wdr61* have opposing gradients, whereas *Hes6*, *B230209C24RIK*, and *Cugbp2* isoforms have similar expression gradients across zones. (C) We verified the expression of *B230209C24RIK* and *Cugbp2* by qRT-PCR after LMD. The first column shows mean RPKM expression levels ( $\pm$ SEM; +,  $bhp < 10^{-5}$ ). The second column of graphs shows mean fold change in qRT-PCR ( $\pm$ SEM) plotted for every cellular zone (\* $P < 0.001$ ). ISH shows that tissue distribution is comparable to the expression gradient determined by the differential analysis. (D) *Mfge8* has two DEIs in the neocortex at E14.5 with different expression gradients. The first column shows mean RPKM expression levels ( $\pm$ SEM; +,  $bhp < 10^{-5}$ ). We also confirmed the difference in the distribution of isoforms by ISH in the mouse cortex. ISH was performed in parallel by using probes designed to differentiate between the two isoforms.

level. As we did not separate cytoarchitectonic regions, we expected our results to be blind to rostrocaudal and mediolateral expression gradients; yet, we were able to detect genes expressed in such gradients (e.g., *Nsg2* and *D0H4S114*; Fig. 2 and Fig. S4). To extend the validation process, we examined the additional DEGs in publicly available databases (GP and ABA), which confirmed the expression gradient predicted with mRNA-seq. However, there was little information available in public databases about the histological distribution of novel genes and splice isoforms. Alternative splicing is a major source of protein diversity that affects cellular expression capacity and gene regulation (32). Our data indicate that, in most cases, splice isoforms increase or decrease concomitantly among different embryonic zones (Fig. 4). We also observed switching events whereby the distribution of splice isoforms conflicted with gene level distribution. In these cases, deducing zone-specific distribution of genes is confounded by splice isoform use, alternative transcription-start sites, and complex *cis*-regulation. The data we present here provide a higher level of detail that is necessary, yet currently unavailable, for investigating the combinatorial expression of splice isoforms that may impart slightly different functions during development.

Two major advantages of mRNA-seq versus microarray technology are its ability to provide a digital readout of mRNA levels in cells (17), as well as its high sensitivity in detecting even few copies of mRNA species. The first apparent feature in our data are that more than 65% of detected genes are expressed in all three developmental zones. In our analysis, we found that several mRNAs (e.g., *Sox2* and *Pax6*) expressed by stem cells in the VZ were present in CP neurons days after their birth in the VZ. As the CP is known to be devoid of stem cells at this stage, we speculate that the degradation of mRNA or silencing of gene expression during development proceeds at a much slower pace than the migration of neurons to their destination in the cortex. We also detected the expression of genes involved in neuronal differentiation (33) (e.g., *NeuroD* and *NeuroD2*) at low levels in the VZ, suggesting an anticipatory program that restricts VZ stem cells to a neocortical neuronal fate. These data lends support to the observation that early activation of differentiation genes occurs in VZ progenitors, which was observed in live imaging experiments of the  $\beta$ -III tubulin-driven GFP (34), and that a restrictive program ensues at the onset of neurogenesis in the dorsal telencephalon (35, 36). Alternatively, the detection of differentiation genes may be a result of heterogeneity of cells in the VZ (37, 38).

Indeed, the VZ at midcorticogenesis contains a heterogeneous population of cells, including radial glia, INPs, and short neuronal progenitors (38, 39) that generate upper layer neurons (4, 33, 40). The predominant function deduced from the VZ transcriptome is cell cycle regulation, which reflects the density and principal role of neural stem cells in the VZ (Table S3). Our functional enrichment analysis also supports a great body of work assigning fate determination and diversification of neuronal types to progenitors in the VZ (1, 41). Furthermore, our analysis identified several pathways enriched in the VZ and SVZ–IZ that seem to be disrupted in cancerous cells. Indeed, a big proportion of infantile supratentorial malignancies originate from stem cells that reside in the VZ and SVZ–IZ (42). As the mouse remains the most used animal model for normal and pathological development, data presented here raise the possibility of finding new tumor-promoting regulatory cascades, and therefore novel therapeutic agents.

The SVZ–IZ region is the major throughway in the developing neocortex, as it contains PNs and INPs (1). PNs initiate the

formation of axons soon after entering this zone (5, 43), and engage in complex navigational maneuvering to find their ultimate destination (44). INPs amplify the number of neurons generated in the VZ (6, 45) and are thought to be responsible for the evolutionary expansion in cortical thickness (46, 47), but whether they contribute to increased diversification in cortical layers or solely to increasing neuronal numbers is not known. Our data indicate that genes involved in cell cycle regulation are enriched in both the VZ and SVZ–IZ. This suggests that the total size of the neuronal population depends on the mitotic activity in both zones. Fate-determination genes were enriched in both zones, albeit more so in group 2 (VZ  $\cap$  SVZ–IZ), indicating that INPs in the SVZ–IZ may also contribute to fate specification. Our results support the currently accepted theory that fate specification occurs at the time of birth in the VZ (1) but raises the possibility that further specification ensues in the SVZ–IZ. Determining the function of SVZ–IZ–specific novel transcription factors (e.g., *Zfp608*, *Myt1*, and *Zfp238*) may explain whether they contribute to an independent specification program or an intermediary implementation of the protomap. In the mouse cortex, a small population of *Sox2*<sup>+</sup> cells are scattered in the SVZ–IZ, but only the most lateral region of the SVZ–IZ contains a sizable pool of *Pax6*<sup>+</sup> cells. However, it is not known if these *Sox2*<sup>+</sup>/*Pax6*<sup>+</sup> cells perform the same function as stem cells that reside in the much-expanded SVZ of the human embryonic brain (46, 47).

The CP at this stage of development is composed of pyramidal neurons that project to subcortical areas including the spinal cord, pons, and midbrain. These neurons grow elaborate dendritic trees to receive proper input. Recent evidence suggests that final differentiation mechanisms do not occur until neurons settle in the CP (48, 49). Indeed, our transcriptome analysis exposed CP-enriched genes and splice isoforms involved in the maturation programs, including dendritic growth, synaptogenesis, exocytosis, and axonogenesis. Furthermore, our analysis uncovered genes associated with tumorigenesis (*Sox11*) (50), mental retardation (*Kirrel3*) (51), and schizophrenia (*Gabrb2*) (52), but their contribution to normal differentiation and signaling processes in the cortex is not yet understood.

Through high-resolution transcriptome analysis, we define five transcriptional programs that contain genes necessary for stem cell maintenance, neurogenesis, migration, and differentiation (Fig. 1F). We expect our functional analyses to facilitate *in vivo* testing of gene network relationships and candidate *cis*-regulatory sequences. Finally, our high-resolution transcriptome map of the mouse transient zones opens the door for determining evolutionary conserved processes, as well as species-specific differences, with embryonic human transient cellular zones.

## Materials and Methods

**Hi-Resolution mRNA-seq.** Cells were collected from fresh frozen tissue by LMD. Sequencing libraries were prepared after RNA extraction and then sequenced using the Illumina GA IIx.

Further experimental details can be found in *SI Materials and Methods*.

**ACKNOWLEDGMENTS.** The authors acknowledge the technical assistance of Mariamma Pappy, as well as Paul Zumbo and Shrikant Mane at the Yale Center for Genome Analysis. We also thank members of the Rakic and Noonan laboratories for critical comments on the manuscript. This work was supported by National Institutes of Health (NIH) Grants DA02399 (to P.R.) and GM094780 (to J.P.N.), the Patterson Trust Fellowship in Brain Circuitry (A.E.A.), and the Kavli Institute for Neuroscience at Yale University. The Yale University Biomedical High Performance Computing Center is supported by NIH Grant RR19895.

1. Rakic P, Ayoub AE, Breunig JJ, Dominguez MH (2009) Decision by division: Making cortical maps. *Trends Neurosci* 32:291–301.
2. Bystron I, Blakemore C, Rakic P (2008) Development of the human cerebral cortex: Boulder Committee revisited. *Nat Rev Neurosci* 9:110–122.
3. Rakic P (1988) Specification of cerebral cortical areas. *Science* 241:170–176.
4. Takahashi T, Goto T, Miyama S, Nowakowski RS, Caviness VS, Jr. (1999) Sequence of neuron origin and neocortical laminar fate: relation to cell cycle of origin in the developing murine cerebral wall. *J Neurosci* 19:10357–10371.
5. Noctor SC, Martínez-Cerdeño V, Ivic L, Kriegstein AR (2004) Cortical neurons arise in symmetric and asymmetric division zones and migrate through specific phases. *Nat Neurosci* 7:136–144.
6. Kowalczyk T, et al. (2009) Intermediate neuronal progenitors (basal progenitors) produce pyramidal-projection neurons for all layers of cerebral cortex. *Cereb Cortex* 19:2439–2450.
7. Gaitanis JN, Walsh CA (2004) Genetics of disorders of cortical development. *Neuroimaging Clin N Am* 14:219–229.

8. Gohlke JM, et al. (2008) Characterization of the proneural gene regulatory network during mouse telencephalon development. *BMC Biol* 6:15.

9. Sansom SN, et al. (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.

10. Rakic P (2006) A century of progress in corticogenesis: From silver impregnation to genetic engineering. *Cereb Cortex* 16(suppl 1):i3–i17.

11. Ajioka I, Nakajima K (2005) Birth-date-dependent segregation of the mouse cerebral cortical neurons in reaggregation cultures. *Eur J Neurosci* 22:331–342.

12. Chen JG, Rasin MR, Kwan KY, Sestan N (2005) Zfp312 is required for subcortical axonal projections and dendritic morphology of deep-layer pyramidal neurons of the cerebral cortex. *Proc Natl Acad Sci USA* 102:17792–17797.

13. Tachikawa K, Sasaki S, Maeda T, Nakajima K (2008) Identification of molecules preferentially expressed beneath the marginal zone in the developing cerebral cortex. *Neurosci Res* 60:135–146.

14. Kudo LC, Karsten SL, Chen J, Levitt P, Geschwind DH (2007) Genetic analysis of anterior posterior expression gradients in the developing mammalian forebrain. *Cereb Cortex* 17:2108–2122.

15. Johnson MB, et al. (2009) Functional and evolutionary insights into human brain development through global transcriptome analysis. *Neuron* 62:494–509.

16. Winden KD, et al. (2009) The organization of the transcriptional network in specific neuronal classes. *Mol Syst Biol* 5:291.

17. Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: A revolutionary tool for transcriptomics. *Nat Rev Genet* 10:57–63.

18. Han X, et al. (2009) Transcriptome of embryonic and neonatal mouse cortex by high-throughput RNA sequencing. *Proc Natl Acad Sci USA* 106:12741–12746.

19. Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10:R25.

20. Trapnell C, et al. (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* 28:511–515.

21. Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y (2008) RNA-seq: An assessment of technical reproducibility and comparison with gene expression arrays. *Genome Res* 18:1509–1517.

22. Bullard JH, Purdom E, Hansen KD, Dudoit S (2010) Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments. *BMC Bioinformatics* 11:94.

23. Kim TK, et al. (2010) Widespread transcription at neuronal activity-regulated enhancers. *Nature* 465:182–187.

24. Visel A, et al. (2009) ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* 457:854–858.

25. Newman AM, Cooper JB (2010) AutoSOME: A clustering method for identifying gene expression modules without prior knowledge of cluster number. *BMC Bioinformatics* 11:117.

26. Warde-Farley D, et al. (2010) The GeneMANIA prediction server: Biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res* 38 (web server issue):W214–W220.

27. Guillemot F (2007) Cell fate specification in the mammalian telencephalon. *Prog Neurobiol* 83:37–52.

28. von Holst A, Egbers U, Prochiantz A, Faiszner A (2007) Neural stem/progenitor cells express 20 tenascin C isoforms that are differentially regulated by Pax6. *J Biol Chem* 282:9172–9181.

29. Berger J, et al. (2007) Conditional activation of Pax6 in the developing cortex of transgenic mice causes progenitor apoptosis. *Development* 134:1311–1322.

30. Twine NA, Janitz K, Wilkins MR, Janitz M (2011) Whole transcriptome sequencing reveals gene expression and splicing differences in brain regions affected by Alzheimer's disease. *PLoS ONE* 6:e16266.

31. Blow MJ, et al. (2010) ChIP-Seq identification of weakly conserved heart enhancers. *Nat Genet* 42:806–810.

32. Black DL (2000) Protein diversity from alternative splicing: A challenge for bioinformatics and post-genome biology. *Cell* 103:367–370.

33. Mattar P, et al. (2008) Basic helix-loop-helix transcription factors cooperate to specify a cortical projection neuron identity. *Mol Cell Biol* 28:1456–1469.

34. Attardo A, Calegari F, Haubensak W, Wilsch-Brauninger M, Huttner WB (2008) Live imaging at the onset of cortical neurogenesis reveals differential appearance of the neuronal phenotype in apical versus basal progenitor progeny. *PLoS ONE* 3:e2388.

35. McCarthy M, Turnbull DH, Walsh CA, Fishell G (2001) Telencephalic neural progenitors appear to be restricted to regional and glial fates before the onset of neurogenesis. *J Neurosci* 21:6772–6781.

36. Shen Q, et al. (2006) The timing of cortical neurogenesis is encoded within lineages of individual progenitor cells. *Nat Neurosci* 9:743–751.

37. Gal JS, et al. (2006) Molecular and morphological heterogeneity of neural precursors in the mouse neocortical proliferative zones. *J Neurosci* 26:1045–1056.

38. Stancik EK, Navarro-Quiroga I, Sellke R, Haydar TF (2010) Heterogeneity in ventricular zone neural precursors contributes to neuronal fate diversity in the postnatal neocortex. *J Neurosci* 30:7028–7036.

39. Soriano E, Dumesnil N, Auladell C, Cohen-Tannoudji M, Sotelo C (1995) Molecular heterogeneity of progenitors and radial migration in the developing cerebral cortex revealed by transgene expression. *Proc Natl Acad Sci USA* 92:11676–11680.

40. Nieto M, et al. (2004) Expression of Cux-1 and Cux-2 in the subventricular zone and upper layers II-IV of the cerebral cortex. *J Comp Neurol* 479:168–180.

41. Desai AR, McConnell SK (2000) Progressive restriction in fate potential by neural progenitors during cerebral cortical development. *Development* 127:2863–2872.

42. Grondin RT, Scott RM, Smith ER (2009) Pediatric brain tumors. *Adv Pediatr* 56:249–269.

43. Hatanaka Y, et al. (2009) Distinct roles of neuropilin 1 signaling for radial and tangential extension of callosal axons. *J Comp Neurol* 514:215–225.

44. Torii M, Hashimoto-Torii K, Levitt P, Rakic P (2009) Integration of neuronal clones in the radial cortical columns by EphA and ephrin-A signalling. *Nature* 461:524–528.

45. Sessa A, Mao CA, Hadjantonakis AK, Klein WH, Broccoli V (2008) Tbr2 directs conversion of radial glia into basal precursors and guides neuronal amplification by indirect neurogenesis in the developing neocortex. *Neuron* 60:56–69.

46. Hansen DV, Lui JH, Parker PR, Kriegstein AR (2010) Neurogenic radial glia in the outer subventricular zone of human neocortex. *Nature* 464:554–561.

47. Fietz SA, et al. (2010) OSVZ progenitors of human and ferret neocortex are epithelial-like and expand by integrin signaling. *Nat Neurosci* 13:690–699.

48. Joshi PS, et al. (2008) Bhlhb5 regulates the postmitotic acquisition of area identities in layers II-V of the developing neocortex. *Neuron* 60:258–272.

49. Kwan KY, et al. (2008) SOX5 postmitotically regulates migration, postmigratory differentiation, and projections of subplate and deep-layer neocortical neurons. *Proc Natl Acad Sci USA* 105:16021–16026.

50. Stuart JE, et al. (2010) Identification of gene markers associated with aggressive meningioma by filtering across multiple sets of gene expression arrays. *J Neuropathol Exp Neurol* 70:1–12.

51. Bhalla K, et al. (2008) Alterations in CDH15 and KIRREL3 in patients with mild to severe intellectual disability. *Am J Hum Genet* 83:703–713.

52. Pun FW, et al. (2010) Imprinting in the schizophrenia candidate gene GABRB2 encoding GABA(A) receptor beta(2) subunit. *Mol Psychiatry* 16:557–568.