Transcriptomes of germinal zones of human and mouse fetal neocortex suggest a role of extracellular matrix in progenitor self-renewal

Simone A. Fietz^a, Robert Lachmann^{b,1}, Holger Brandl^{a,1}, Martin Kircher^c, Nikolay Samusik^a, Roland Schröder^c, Naharajan Lakshmanaperumal^a, Ian Henry^a, Johannes Vogt^{d,2}, Axel Riehn^b, Wolfgang Distler^b, Robert Nitsch^{d,2}, Wolfgang Enard^c, Svante Pääbo^{c,3}, and Wieland B. Huttner^{a,3}

^aMax Planck Institute of Molecular Cell Biology and Genetics, 01307 Dresden, Germany; ^bKlinik und Poliklinik für Frauenheilkunde und Geburtshilfe, Universitätsklinikum Carl Gustav Carus, Technische Universität Dresden, 01307 Dresden, Germany; ^cDepartment of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, 04103 Leipzig, Germany; and ^dCenter for Anatomy, Institute of Cell Biology and Neurobiology, Charité, 10117 Berlin, Germany

Contributed by Svante Pääbo, June 8, 2012 (sent for review May 7, 2012)

The expansion of the neocortex during mammalian brain evolution results primarily from an increase in neural progenitor cell divisions in its two principal germinal zones during development, the ventricular zone (VZ) and the subventricular zone (SVZ). Using mRNA sequencing, we analyzed the transcriptomes of fetal human and embryonic mouse VZ, SVZ, and cortical plate. In mouse, the transcriptome of the SVZ was more similar to that of the cortical plate than that of the VZ, whereas in human the opposite was the case, with the inner and outer SVZ being highly related to each other despite their cytoarchitectonic differences. We describe sets of genes that are up- or down-regulated in each germinal zone. These data suggest that cell adhesion and cell-extracellular matrix interactions promote the proliferation and self-renewal of neural progenitors in the developing human neocortex. Notably, relevant extracellular matrix-associated genes include distinct sets of collagens, laminins, proteoglycans, and integrins, along with specific sets of growth factors and morphogens. Our data establish a basis for identifying novel cell-type markers and open up avenues to unravel the molecular basis of neocortex expansion during evolution.

cerebral cortex | neural stem cells | neurogenesis

N eccortex expansion is a hallmark of mammalian brain evolution. With regard to neuron number, a major cause of this expansion is the increase in the population size of neural stem and progenitor cells (NSPCs) and the number of divisions that each of the various NSPC types undergoes during cortical development (1–4). Two principal classes of these cells can be distinguished based on the location of their mitosis: (*i*) apical progenitors (APs), which undergo mitosis at the luminal surface of the ventricular zone (VZ); and (*ii*) basal progenitors (BPs), which undergo mitosis at an abventricular location, typically in the subventricular zone (SVZ) (2, 5, 6). Neurons born from AP and BP cell divisions migrate radially and settle at the basal (pial) side of the developing cortical wall to form the cortical plate (CP).

Both APs and BPs comprise several types of NSPCs that differ in key cell biological features (e.g., cell polarity, cell processes, cell-to-cell junctions, nuclear migration) and in the principal modes of cell division (symmetric proliferative vs. asymmetric self-renewing vs. symmetric or asymmetric consumptive) (2, 5– 10). APs comprise neuroepithelial cells, which transform into apical radial glial cells (aRGCs) at the onset of neurogenesis (11), and short neural precursors (12). BPs include basal (or outer) radial glial cells (bRGCs), transit amplifying progenitors (TAPs), and intermediate progenitor cells (IPCs) (2, 3, 13).

The evolutionary expansion of the neocortex is associated with an increase in the thickness of the SVZ, which develops into two cytoarchitecturally distinct zones, an inner SVZ (ISVZ) and an outer SVZ (OSVZ) (1–4, 14, 15). The evolutionary increase in the SVZ is accompanied by a change in the proportion of BP subtypes. For example, in the mouse SVZ (mSVZ), ~90% of the BPs are IPCs, which undergo one terminal round of cell division, and bRGCs and TAPs, which can undergo multiple rounds of cell division, constitute a very minor fraction (13, 16–21). In striking contrast, in the human SVZ (hSVZ), about half of the BPs are bRGCs, and TAPs appear to outnumber IPCs (22–24). To gain insight into the molecular mechanisms underlying the differences in the germinal zones (GZs) of the developing neocortex, and in the NSPCs contained therein, we analyzed the transcriptomes of the VZ, ISVZ, OSVZ, and CP of the human fetal neocortex, and of the VZ, SVZ, and CP of the mouse embryonic neocortex.

Results and Discussion

RNA-Seq of Fetal Human and Embryonic Mouse Cortical Zones. The GZs of the developing neocortex are heterogeneous in terms of the NSPC subpopulations they contain (2, 3). However, as molecular markers applicable for the isolation of cortical NSPC subpopulations are at best partly known, we decided to obtain RNA separately from the various cortical zones. Specifically, we isolated total RNA from the VZ, ISVZ, OSVZ, and CP of six 13-16 wk postconception (wpc) human fetuses and from the VZ, SVZ, and CP of five embryonic day (E) 14.5 mouse embryos (Fig. 1A and Fig. S1A), using laser-capture microdissection of Nissl-stained cryosections of dorsolateral telencephalon (Fig. S14). Similar experimental approaches have successfully been applied previously in mouse but not in humans (25, 26). Analysis of the RNA quality revealed an RNA integrity number of 8.5-9.5 for all samples (Fig. S1*B*). Poly(A)⁺ RNA was used as template for the preparation of 24 human and 15 mouse cDNA libraries (Fig. S1C), which were then subjected to single-end 76-bp RNA-Seq (Fig. S1D). Gene-expression levels were quantified as fragments per kilobase of exon per million fragments mapped (FPKM). Relative expression levels across GZs for eight selected human genes as revealed by RNA-Seq were generally

Author contributions: S.A.F., W.E., S.P., and W.B.H. designed research; S.A.F. performed research; R.L., J.V., A.R., W.D., and R.N. provided fetal human tissue; R.S. assisted in laser-capture microdissection; S.A.F., H.B., M.K., N.S., N.L., and I.H. analyzed data; and S.A.F., W.E., S.P., and W.B.H. wrote the paper.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

Data deposition: The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, www.ncbi.nlm.nih.gov/geo (accession no. GSE38805).

¹R.L. and H.B. contributed equally to this work.

²Present address: Institute for Microscopic Anatomy and Neurobiology, University Medicine Mainz, Johannes Gutenberg University Mainz, 55131 Mainz, Germany.

 $^{^3{\}rm To}$ whom correspondence may be addressed. E-mail: paabo@eva.mpg.de or huttner@mpi-cbg.de.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10. 1073/pnas.1209647109/-/DCSupplemental.

Fig. 1. Correlation analysis of fetal human and embryonic mouse cortical zones. (A) Cartoon illustrating the major cell types (APs in blue, BPs in orange and red, neurons in green) in the different GZs and the CP of the developing mouse and human neocortex. (B and C) Heatmap showing pair-wise correlations of gene expression levels between the indicated zones of the fetal human neocortex at the indicated gestational weeks (B, 13-16 wpc) and of the E14.5 mouse neocortex (C). Spearman's rank correlation coefficients (r_s) range from 0.7 (yellow) to 1.0 (blue). Numbers in the blank guadrants indicate the mean r. values ± SD of the zone comparison located mirror-symmetrically to the di-



agonal line; asterisks indicate zone comparison mean r_s values that show a statistically significant difference to each other (P < 0.05).

concordant with those obtained by quantitative PCR of total RNA, thus validating the RNA-Seq data (Fig. S2).

For an initial characterization of the gene-expression patterns in the fetal human cortical zones, we used the FPKM values to calculate pair-wise correlations between the 13-16 wpc samples within a given zone and between the four zones (Fig. 1B). The resulting correlation coefficients were used to cluster the various 13–16 wpc zones into a correlation dendrogram (Fig. S3A). This analysis revealed that the transcriptomes of the three GZs were more closely related to each other than to the CP, with the ISVZ and OSVZ being most closely related to each other. Remarkably, the OSVZ showed a slightly closer relationship to the VZ than did the ISVZ, despite its larger spatial distance to the VZ. Only little differences were found with regard to the various fetal stages (13-16 wpc). Importantly, the same analysis for E14.5 mouse showed that, in contrast to human, the transcriptome of the SVZ was more closely related to that of the CP than was the VZ (Fig. 1*C* and Fig. S3*B*). This finding is consistent with the notion that in human, all three GZs characteristically contain self-renewing NSPCs, whereas in mouse this holds true only for the VZ.

Expression Patterns of Marker Genes. The various human and mouse NSPCs have previously been characterized with regard to the expression of molecular markers, in most cases using immunohistochemistry (2, 3, 5, 6, 8, 27, 28). We examined the mRNA expression levels of selected markers of APs, BPs, and neurons in the different zones (Fig. S4). The mRNA of the AP marker prominin-1 was predominantly expressed in the hVZ and mVZ (Fig. S4 B and I). In both species, mRNA levels for Pax6, nestin, and GLAST, known to be expressed by APs and bRGCs (2, 3, 5, 6, 22, 23, 27–29), were highest in the VZ and lowest in the CP. Interestingly, these mRNAs were relatively stronger expressed in hISVZ and hOSVZ than in mSVZ (Fig. S4 B and I). This finding is consistent with the much higher abundance of bRGCs in hISVZ and hOSVZ than mSVZ (13, 20-23) as revealed, for example, by the greater proportion of Pax6⁺ nuclei in hSVZ (Fig. S4E) than mSVZ (28).

Markers of TAP/IPCs and of TAP/IPC-generation (Tbr2, neurogenin-2, Svet1, Insm1) were highly expressed in all GZs but only weakly in the hCP and mCP (Fig. S4 *C* and *J*). Remarkably, however, the relative levels of the Tbr2 and neurogenin-2 mRNAs were higher in mVZ than mSVZ (Fig. S4*J*), whereas the opposite has been observed for the Tbr2 protein by immuno-histochemistry (27, 28). In contrast, the Tbr2 and neurogenin-2 mRNAs were lower in the hVZ than hISVZ and hOSVZ (Fig. S4*C*), in accordance with Tbr2 immunohistochemistry (Fig. S4*F*). These data are consistent with the notion that in mouse the vast majority of BPs are IPCs, which are generated in the VZ (5, 6, 16–19, 30), whereas in human most Tbr2⁺ BPs are thought to be generated in the SVZ (2, 3, 22).

mRNA levels of neuronal markers (β III-tubulin, Tbr1, NeuroD2, NF-M) were highest in the h/mCP, intermediate in the hISVZ/OSVZ and mSVZ, and lowest in the h/mVZ (Fig. S4 *D* and *K*), as observed for Tbr1 protein by immunohistochemistry (Fig. S4G). This finding is consistent with the notion that in mouse (17, 19, 31) and human (2, 3), BPs generate substantially more neurons than APs. The similar proportions of neuronal marker mRNAs in the mSVZ and hSVZ indicate that the higher similarity of expression patterns between the VZ and SVZ in human compared with mouse (Fig. 1 and Fig. S3) is not because of a different contribution of neurons, but might indeed reflect the common potential for self-renewal.

Genes Differentially Expressed Between Cortical Zones. We used DESeq (Fig. 2A, and Fig. S5 A and C) to identify sets of genes significantly up- or down-regulated in the human (Fig. 2B and Dataset S1) and mouse (Fig. S5B and Dataset S1) cortical zones. [Because only the $poly(A)^+$ RNAs had been used as templates for the preparation of the cDNA libraries, we focused our analysis on protein-encoding RNAs.] In line with the high similarity in overall gene expression between hISVZ and hOSVZ (Figs. 1 and 24, and Fig. S3), these two zones had the smallest numbers of differentially expressed genes (Fig. 2B, red and orange). Interestingly, the 55 genes up-regulated in hOSVZ relative to hVZ, hISVZ, and hCP, as well as the 20 genes up-regulated in the hOSVZ relative to all other human and mouse zones (Dataset S1), included Olig1 and Sox10, which have been implicated in oligodendrogenesis (32, 33). Moreover, both these gene groups included two highly related but functionally poorly understood genes, KIAA1324 and KIAA1324-like, which showed barely detectable expression in the two mGZs (Dataset S1). These genes are thus potential novel markers for hOSVZ.

To gain insight into the biological processes associated with the large number of genes up- or down-regulated in human CP compared with the VZ, ISVZ, and OSVZ, OSVZ compared with VZ and CP, ISVZ compared with VZ and CP, and VZ compared with ISVZ, OSVZ, and CP (Fig. 2B), we used DAVID (34) to cluster the functional gene annotation (FGA) terms significantly overrepresented in these sets of genes (Fig. 2C and Dataset S2). The same analysis was performed among the mouse zones (Fig. 3A and Dataset S2), hCP vs. mCP, hOSVZ or hISVZ vs. mSVZ, and hVZ vs. mVZ (Fig. 3B and Dataset S2). From this comparison, four major findings emerged.

First, the majority (53%) of the five FGA terms most enriched among the genes with higher expression in either the hOSVZ, hISVZ, or hVZ were highly related, or identical, to each other, and concerned cell-to-extracellular matrix (ECM) interaction and the secretory processes underlying both the secretion of ECM constituents and the surface delivery of the corresponding receptors (Fig. 2*C*, *Left*, and Dataset S2). Second, FGA terms associated with ECM interaction and cell adhesion were



Fig. 2. Differential gene-expression analysis and functional annotation clustering of fetal human cortical zones. (A) DESeq scatter plot showing pair-wise comparisons of gene expression between the indicated zones of the 14–16 wpc human neocortex. Each dot represents the mean (five fetuses) expression level for a given gene, with differentially expressed genes (false-discovery rate <0.05) shown in green. Numbers above (below) the diagonal lines refer to the differentially expressed genes up-regulated in the zone indicated on top (on the right). These numbers include the genes expressed in only one of the two zones compared (green dots on vertical and horizontal 0.1 lines). For the corresponding analysis of genes differentially expressed in mouse zones and between human and mouse zones see Fig. S5. (*B*) Venn diagrams showing the numbers of differentially expressed genes (see *A*) that are up-regulated (*Left*, >) and down-regulated (*Right*, <) in a given zone compared with the other zones indicated of the 14–16 wpc human neocortex. (C) Differentially expressed genes (see *A*) that are up-regulated (*Right*, <) in a given zone compared with the other zones indicated of the 14–16 wpc human neocortex. Were analyzed for significantly enriched FGA terms which were clustered, using DAVID. The five clusters with the highest enrichment scores are shown. Note that for certain FGA terms enriched in the various GZs, the same designation were chosen (e.g., secretory pathway), although the individual genes in these clusters were different.

enriched among the genes with lower expression in the mSVZ (Fig. 3*A*, orange *Right*, and Dataset S2). Third, when comparing mVZ with mSVZ and mCP, FGA terms related to cell division and DNA replication (rather than ECM interactions) were strikingly overrepresented (Fig. 3*A*, blue *Left*, and Dataset S2). Fourth, FGA terms associated with ECM interaction, the secretory process, and cell adhesion constituted the vast majority of terms overrepresented among genes with higher expression in hGZs compared with mGZs (Fig. 3*B*, *Left*, and Dataset S2).

Taken together, these data suggest that NSPCs in hGZs are more dependent on ECM interactions and associated secretory pathways than is the case for mouse. Because the majority of neural NSPCs in the developing human neocortex (APs, bRGCs, and TAPs), in contrast to the majority of mouse BPs, are capable of undergoing self-renewing divisions, ECM interactions may promote the capacity for sustained self-renewal of NSPCs in the hGZs. This notion is consistent with previous observations and concepts (35–38) and with our previous findings concerning bRGC maintenance in ferrets and its dependence on integrin signaling (23).

Cluster Analysis of Genes Related to Cell-ECM Interaction. To obtain further insights into the putative role of the ECM for proliferation and self-renewal of NSPCs, we performed mean-shift clustering of those genes for which the DAVID FGA terms included ECM, extracellular region, ECM-receptor interaction, proteoglycan, and basement membrane (referred to as ECM-associated genes). (See Fig. S6 and *SI Results and Discussion* for the results of mean-shift clustering of all genes.) This analysis yielded nine clusters that differed in the pattern of gene-expression levels across the mouse and human zones (Fig. 4 *A–J*). In five of these clusters (Fig. 4 *B–F*), the median gene expression in the hVZ and mVZ, hISVZ, and hOSVZ was higher than in the other zones; these clusters are therefore referred to as "GZ clusters." These expression patterns indicated that the genes in these clusters included those of relevance for NSPC self-renewal. Inspection of the gene lists of the five GZ clusters yielded the following insights.

One GZ cluster (198 genes) (Fig. 4 *A* and *B* and Dataset S3) contained genes most highly expressed in h/mVZ and to a lesser extent in hISVZ and hOSVZ. Enriched FGA terms included growth factor binding, heparin binding, cell proliferation, basement membrane, and the Notch signaling pathway (Dataset S4). As the expression pattern of this cluster reflected the relative abundance of aRGCs and bRGCs in the hGZs and mGZs, the genes contained therein include candidates for supporting self-renewal of these NSPCs in an autocrine manner.

Three GZ clusters contained ECM-associated genes more highly expressed in the GZs than CP of human but not mouse (Fig. 4 A, C, D, and F, and Dataset S3). The first cluster showed a median gene-expression pattern hVZ > hISVZ/hOSVZ (126 genes) (Fig. 4 A and C), and hence may contain genes that specifically support the self-renewal of human-but not mouse-NSPCs. The second cluster contained ECM-associated genes most highly expressed in the hVZ but not hISVZ and hOSVZ (68 genes) (Fig. 4 A and D, and Dataset S3). Similarly, a distinct cluster was obtained for the mVZ (79 genes) (Fig. 4A and E, and Dataset S3). This finding implies that the ECM-associated genes that would specifically support autocrine self-renewal of aRGCs would be distinct for the hVZ and mVZ. The third hGZ cluster showed a median gene-expression pattern hISVZ = hOSVZ but not hVZ (78 genes) (Fig. 4A and F, and Dataset S3). This cluster presumably contains ECM-associated genes involved in hSVZspecific aspects of NSPC behavior. Strikingly, enriched FGA terms related to regulation of vasoconstriction, blood vessel remodeling, and hormone activity were found in this, but not in the other GZ clusters (Dataset S4).



Fig. 3. Functional annotation clustering of genes differentially expressed among mouse cortical zones and between human and mouse cortical zones. Differentially expressed genes identified by DESeq (Fig. S5) that are upregulated (*Left*, >) and down-regulated (*Right*, <) between the indicated zones of the E14.5 mouse neocortex (*A*) and between the 14–16 wpc human and E14.5 mouse neocortex (*B*) were analyzed for significantly enriched FGA terms, which were clustered, using DAVID. The five clusters with the highest enrichment scores are shown. See note on FGA terms in Fig. 2C.

In addition to the five GZ clusters, we obtained four clusters, referred to as "CP clusters," that contained ECM-associated genes more highly expressed in the hCP and/or mCP than in the hGZs and mVZ (Fig. 4 *G–J* and Dataset S3). These genes are likely not of relevance for NSPC self-renewal. The CP clusters consisted of two with genes most highly expressed in the hCP and mCP [43 genes (Fig. 4*H*), 97 genes (Fig. 4*I* and Dataset S3)], and two containing genes most highly expressed only in hCP (65 genes) (Fig. 4*J* and Dataset S3) or mCP (53 genes) (Fig. 4*G* and Dataset S3). Remarkably, genes more highly expressed in the mSVZ than other GZs were found in one of the CP clusters (Fig. 4*I*) rather than the GZ clusters. The vast majority of the enriched FGA terms in the cluster shown in Fig. 4*I* were related to

When focusing on the classic ECM constituents and receptors among the ECM-associated genes, collagen genes were exclusively present in the GZ clusters (Fig. 4K and Dataset S3). Moreover, the vast majority of the laminins (Fig. 4L and Dataset S3), proteoglycans (Fig. 4M and Dataset S3), and integrins (Fig. 4N and Dataset S3) were also found in the GZ clusters. Interestingly, when focusing on individual zones, in human the overwhelming majority of the ECM-associated genes were shared between the GZs, with only a few being specific for either hVZ or hISVZ/OSVZ, and with ECM-associated genes in hCP being clearly distinct (Fig. 4P). In contrast, in mouse the ECMassociated genes were shared between the mCP and mSVZ, but were distinct for the mVZ (Fig. 4P). These data have significant implications with regard to the site of production of specific ECM constituents in the developing human versus mouse neocortex and their potential role in the differential self-renewal capacity of human versus mouse NSPCs, as is discussed below.

Finally, it is important to emphasize that the majority of growth factors and morphogens included with the ECM-associated genes were found in the GZ clusters rather than the CP clusters (Fig. 40 and Dataset S3). This finding points to an interplay between these factors and the ECM/ECM receptors in the signaling that ultimately influences the various types of NSPCs in the developing human and mouse neocortex.

Putative Regulation of ECM-Associated Genes by Transcription Factors.

We sought to obtain clues as to the transcription factors (TFs) that may regulate the ECM-associated genes within the nine clusters (Fig. 4 *B–J*). To this end, for each cluster, the promoter regions of the genes contained therein were analyzed for overrepresentation of specific TF binding sites using oPOSSUM (40, 41). The TFs implicated in the various expression patterns of ECM-associated genes across human and mouse cortical zones were then analyzed for their relative abundance in these zones. Only those TFs were considered that showed a significantly (P < 0.05) higher expression level in the cortical zones with a median log₂ fold-change >0 (Fig. 4 *B–J*, red columns) compared with those with a median log₂ foldchange <0 (Fig. 4 *B–J*, green columns). This analysis revealed strikingly different sets of TFs putatively

This analysis revealed strikingly different sets of TFs putatively regulating the ECM-related genes in hGZs and mVZ vs. h/mCP and mSVZ (Dataset S5). Remarkably, there was substantial overlap in the TFs between the five GZ clusters, suggesting that highly related transcriptional networks may drive the expression of ECM-associated genes in those h/mGZs that are characterized by a high abundance of self-renewing NSPCs. Interestingly, these candidate TFs included not only some already known to be involved in NSPC self-renewal, such as Sox2, but also transcriptional regulators not previously implicated in this process, such as ZNF143 (Dataset S5).

Conclusions

We provide genome-wide RNA-Seq-based expression data for the distinct GZs of the fetal human neocortex. Recently, similar data have been reported for the mouse (26), extending previous gene-expression studies on embryonic cortical NSPCs using microarray analyses (see e.g., refs. 28 and 42). In contrast, for humans, only different areas of the fetal neocortex, without distinguishing GZs and the CP, have been studied (43). The implications of our data with regard to NSPCs in the developing human versus mouse neocortex and the potential role of the ECM in their self-renewal capacity include the following.

First, despite their cytoarchitectonic differences (14), gene expression in the hISVZ and hOSVZ was found to be very similar. This finding presumably reflects the fact that the various



Fig. 4. Mean-shift clustering of genes related to ECM interaction in fetal human and embryonic mouse cortical zones. (*A–J*) Log₂-transformed FPKM fold-changes of ECM-related genes in the 14–16 wpc hVZ, hISVZ, hOSVZ, and hCP (five fetuses) and in the E14.5 mVZ, mSVZ, and mCP (five embryos) relative to the mean expression of each of the genes in all human or mouse zones were subjected to nonhierarchical K-nearest neighbor mean-shift (KNN-MS) cluster analysis, which yielded a total of nine clusters. (*A*) Heat

subtypes of BPs—that is, bRGCs, TAPs, and IPCs—are found in both, hISVZ and hOSVZ, albeit at different relative abundance.

Second, the functional annotations of genes differentially expressed in the various GZs revealed that increased ECM interaction and cell adhesion may underlie the greater self-renewal capacity of NSPCs in hGZs and the mVZ compared with the mSVZ. Potentially relevant ECM-associated genes include distinct sets of collagens, laminins, proteoglycans, and integrins, along with specific growth factors and morphogens. Our findings are in line with previous concepts (35–38) and with the recent findings that (i) mouse NSPCs undergoing terminal neurogenic division down-regulate gene products enabling ECM interactions (28), and (ii) interference with integrin signaling reduces the abundance of ferret bRGCs (23).

Third, a role of ECM constituents and receptors in NSPC selfrenewal during cortical development may contribute to explaining the phenotypes observed in mice and humans carrying mutations in ECM genes. For example, loss of the proteoglycan perlecan results in microcephaly in mouse (44, 45) and human (46). Interestingly, we found that perlecan was more highly expressed in those h/mGZs that are known to be rich in self-renewing NSPCs (Fig. 4*B*). Moreover, the perlecan-containing cluster of ECMassociated genes (Fig. 4*B*) also contained fibulin-4 (EFEMP2), mutations of which also cause microcephaly in human (47).

Fourth, our findings that the three hGZs and the mVZ are major sites of production not only of ECM receptors, but also of ECM constituents, have interesting implications as to the cell biological basis underlying the putative ECM-based promotion of NSPC self-renewal. On the one hand, these cells may contribute, via vesicular transport in their basal processes, to the deposition of ECM constituents at the basal lamina, contact with which likely supports NSPC self-renewal (23). In this context, it is interesting to note that genes, mutations in which result in major basal lamina defects and cortical development disorders, such as fukutin (48), were more highly expressed in those h/mGZs that contain aRGCs and bRGCs (Fig. 4B). On the other hand, ECM constituents may also be deposited locally in the GZs themselves, thereby creating a microenvironment conducive for cell proliferation and self-renewal.

In other words, NSPCs in the three hGZs and mVZ, but not the mSVZ, may generate their own niche. In the first of the above scenarios, this niche is linked to the basal side of the developing cortical wall (i.e., the basal lamina at the pial surface). In the second scenario, this niche is generated locally in those

map showing the nine clusters sorted hierarchically using an average-linkage algorithm. Each gene is represented by a single row and each developmental stage by a single column. Fold-changes range from 5.2 (red) to -6.6 (green). (B-J) Median log₂ fold-changes in expression of the genes (numbers in parentheses) contained in the nine clusters shown in A. (K-O) Occurrence of collagens (K), laminins (L), proteoglycans (M), integrins (N), and growth factors/morphogens (O) in the clusters indicated in B-J of the 14-16 wpc human (five fetuses) or E14.5 mouse (five embryos) neocortex. Numbers in parentheses refer to the total number of collagen, laminin, proteoglycan, integrin, or growth factor/morphogen genes expressed in the fetal human and embryonic mouse neocortex: numbers in the pie chart sectors refer to the number of the respective genes expressed in the indicated cluster. Black sectors, GZ clusters; white sectors, CP clusters. (P) Summary of the distribution of collagens (COL), laminins (LAM), proteoglycans (PG), integrins (INT), and growth factors/morphogens (GFM) between the various zones of fetal human and embryonic mouse neocortex (APs in blue, BPs in orange and red, neurons in green). Numbers in parentheses, total number of genes in respective class, as indicated in K-O; numbers in green/red/blue boxes, number of genes in respective class that are specifically overexpressed (red bars in B-J) only in the hCP or mCP (green), hISVZ, and hOSVZ or mSVZ (red), or hVZ or mVZ (blue); numbers in yellow boxes, number of genes in respective class that are overexpressed in more than one human or mouse zone; numbers in white boxes, number of genes in respective class that are overexpressed in both h/mVZ, hISVZ/ hOSVZ/mSVZ, or h/mCP.

GZs in which the majority of NSPCs can undergo self-renewal. These two scenarios are not mutually exclusive. Moreover, the basal lamina-containing blood vessels that occur throughout the GZs (49–51) may constitute an additional embryonic neural stem cell niche.

Finally, the present catalogs of genes differentially expressed among the various GZs of fetal human and embryonic mouse neocortex will help identify molecular markers for NSPC subtypes, notably BP subtypes, including cell-surface molecules that could be used to isolate NSPC subpopulations. This approach would extend previous microarray studies on NSPC subpopulations isolated from embryonic mouse neocortex (28, 42). Moreover, comparison of gene-expression profiles between the various human and mouse NSPC subtypes may provide clues as to which genes control the number of proliferative and self-renewing divisions, which in turn should lead to insight into the interspecies differences in NSPC behavior during neocortex development that underlie the differences in neocortex expansion among species.

- Rakic P (2009) Evolution of the neocortex: A perspective from developmental biology. Nat Rev Neurosci 10:724–735.
- Fietz SA, Huttner WB (2011) Cortical progenitor expansion, self-renewal and neurogenesis-a polarized perspective. Curr Opin Neurobiol 21:23–35.
- 3. Lui JH, Hansen DV, Kriegstein AR (2011) Development and evolution of the human neocortex. Cell 146:18–36.
- 4. Molnár Z (2011) Evolution of cerebral cortical development. *Brain Behav Evol* 78:94–107.
- Götz M, Huttner WB (2005) The cell biology of neurogenesis. Nat Rev Mol Cell Biol 6: 777–788.
- Kriegstein A, Alvarez-Buylla A (2009) The glial nature of embryonic and adult neural stem cells. Annu Rev Neurosci 32:149–184.
- Huttner WB, Kosodo Y (2005) Symmetric versus asymmetric cell division during neurogenesis in the developing vertebrate central nervous system. *Curr Opin Cell Biol* 17:648–657.
- Farkas LM, Huttner WB (2008) The cell biology of neural stem and progenitor cells and its significance for their proliferation versus differentiation during mammalian brain development. *Curr Opin Cell Biol* 20:707–715.
- 9. Pontious A, Kowalczyk T, Englund C, Hevner RF (2008) Role of intermediate progenitor cells in cerebral cortex development. *Dev Neurosci* 30:24–32.
- Miyata T, Kawaguchi D, Kawaguchi A, Gotoh Y (2010) Mechanisms that regulate the number of neurons during mouse neocortical development. *Curr Opin Neurobiol* 20:22–28.
- 11. Kriegstein AR, Götz M (2003) Radial glia diversity: A matter of cell fate. Glia 43:37-43.
- 12. Corbin JG, et al. (2008) Regulation of neural progenitor cell development in the nervous system. J Neurochem 106:2272-2287.
- 13. Kelava I, et al. (2012) Abundant occurrence of basal radial glia in the subventricular zone of embryonic neocortex of a lissencephalic primate, the common marmoset *Callithrix jacchus*. *Cereb Cortex* 22:469–481.
- Smart IH, Dehay C, Giroud P, Berland M, Kennedy H (2002) Unique morphological features of the proliferative zones and postmitotic compartments of the neural epithelium giving rise to striate and extrastriate cortex in the monkey. *Cereb Cortex* 12:37–53.
- Fish JL, Dehay C, Kennedy H, Huttner WB (2008) Making bigger brains-the evolution of neural-progenitor-cell division. J Cell Sci 121:2783–2793.
- 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.
- Haubensak W, Attardo A, Denk W, Huttner WB (2004) Neurons arise in the basal neuroepithelium of the early mammalian telencephalon: A major site of neurogenesis. Proc Natl Acad Sci USA 101:3196–3201.
- Miyata T, et al. (2004) Asymmetric production of surface-dividing and non-surfacedividing cortical progenitor cells. *Development* 131:3133–3145.
- Attardo A, Calegari F, Haubensak W, Wilsch-Bräuninger 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.
- Shitamukai A, Konno D, Matsuzaki F (2011) Oblique radial glial divisions in the developing mouse neocortex induce self-renewing progenitors outside the germinal zone that resemble primate outer subventricular zone progenitors. J Neurosci 31: 3683–3695.
- Wang X, Tsai JW, LaMonica B, Kriegstein AR (2011) A new subtype of progenitor cell in the mouse embryonic neocortex. Nat Neurosci 14:555–561.
- Hansen DV, Lui JH, Parker PR, Kriegstein AR (2010) Neurogenic radial glia in the outer subventricular zone of human neocortex. Nature 464:554–561.
- 23. Fietz SA, et al. (2010) OSVZ progenitors of human and ferret neocortex are epitheliallike and expand by integrin signaling. *Nat Neurosci* 13:690–699.
- Reillo I, de Juan Romero C, García-Cabezas MA, Borrell V (2011) A role for intermediate radial glia in the tangential expansion of the mammalian cerebral cortex. *Cereb Cortex* 21:1674–1694.
- Wang WZ, Oeschger FM, Lee S, Molnár Z (2009) High quality RNA from multiple brain regions simultaneously acquired by laser capture microdissection. BMC Mol Biol 10:69.

Methods

Fetal human and embryonic mouse cortical zones were obtained by lasercapture microdissection, subjected to RNA-Seq, and analyzed using state-ofthe-art methods. Details can be found in *SI Methods*. See Datasets S6 and S7 for further data regarding functional annotations of genes of each cluster using mean-shift clustering.

ACKNOWLEDGMENTS. We thank the sequencing group of the Max Planck Institute for Evolutionary Anthropology, and Jussi Helppi and other members of the animal facility of the Max Planck Institute of Molecular Cell Biology and Genetics for excellent support; Henrike Heyne, Birgit Nickel, and other members of the S.P. group for experimental advice; Dr. Denise Stenzel of the W.B.H. group for discussion; and Vineeth Surendranath of the Max Planck Institute of Molecular Cell Biology and Genetics for helpful comments regarding data analysis. S.A.F. and N.S. were members of the International Max Planck Research School for Molecular Cell Biology and Bioengineering. W.B.H. was supported by grants from the Deutsche Forschungsgemeinschaft (SFB 655, A2; TRR 83, Tp6) and the European Research Council (250197), by the Deutsche Forschungsgemeinschaft-funded Center for Regenerative Therapies Dresden, and by the Fonds der Chemischen Industrie.

- Ayoub AE, et al. (2011) Transcriptional programs in transient embryonic zones of the cerebral cortex defined by high-resolution mRNA sequencing. *Proc Natl Acad Sci USA* 108:14950–14955.
- Englund C, et al. (2005) Pax6, Tbr2, and Tbr1 are expressed sequentially by radial glia, intermediate progenitor cells, and postmitotic neurons in developing neocortex. J Neurosci 25:247–251.
- 28. Arai Y, et al. (2011) Neural stem and progenitor cells shorten S-phase on commitment to neuron production. *Nat Commun* 2:154.
- Osumi N, Shinohara H, Numayama-Tsuruta K, Maekawa M (2008) Concise review: Pax6 transcription factor contributes to both embryonic and adult neurogenesis as a multifunctional regulator. *Stem Cells* 26:1663–1672.
- Ochiai W, et al. (2009) Periventricular notch activation and asymmetric Ngn2 and Tbr2 expression in pair-generated neocortical daughter cells. *Mol Cell Neurosci* 40: 225–233.
- 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.
- Woodruff RH, Tekki-Kessaris N, Stiles CD, Rowitch DH, Richardson WD (2001) Oligodendrocyte development in the spinal cord and telencephalon: Common themes and new perspectives. Int J Dev Neurosci 19:379–385.
- Li H, Richardson WD (2008) The evolution of Olig genes and their roles in myelination. Neuron Glia Biol 4:129–135.
- Huang da W, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4:44–57.
- Lathia JD, et al. (2007) Patterns of laminins and integrins in the embryonic ventricular zone of the CNS. J Comp Neurol 505:630–643.
- Lathia JD, Rao MS, Mattson MP, Ffrench-Constant C (2007) The microenvironment of the embryonic neural stem cell: lessons from adult niches? *Dev Dyn* 236:3267–3282.
- Loulier K, et al. (2009) Beta1 integrin maintains integrity of the embryonic neocortical stem cell niche. *PLoS Biol* 7:e1000176.
- Marthiens V, Kazanis I, Moss L, Long K, Ffrench-Constant C (2010) Adhesion molecules in the stem cell niche—More than just staying in shape? J Cell Sci 123:1613–1622.
- Bradke F, Dotti CG (1998) Membrane traffic in polarized neurons. *Biochim Biophys* Acta 1404:245–258.
- 40. Ho Sui SJ, et al. (2005) oPOSSUM: Identification of over-represented transcription factor binding sites in co-expressed genes. *Nucleic Acids Res* 33:3154–3164.
- Ho Sui SJ, Fulton DL, Arenillas DJ, Kwon AT, Wasserman WW (2007) oPOSSUM: Integrated tools for analysis of regulatory motif over-representation. *Nucleic Acids Res* 35(Web Server issue):W245–W252.
- Kawaguchi A, et al. (2008) Single-cell gene profiling defines differential progenitor subclasses in mammalian neurogenesis. *Development* 135:3113–3124.
- Johnson MB, et al. (2009) Functional and evolutionary insights into human brain development through global transcriptome analysis. *Neuron* 62:494–509.
- Girós A, Morante J, Gil-Sanz C, Fairén A, Costell M (2007) Perlecan controls neurogenesis in the developing telencephalon. BMC Dev Biol 7:29.
- Barros CS, Franco SJ, Müller U (2011) Extracellular matrix: Functions in the nervous system. Cold Spring Harb Perspect Biol 3:a005108.
- Arikawa-Hirasawa E, et al. (2001) Dyssegmental dysplasia, Silverman-Handmaker type, is caused by functional null mutations of the perlecan gene. Nat Genet 27:431–434.
- Hoyer J, Kraus C, Hammersen G, Geppert JP, Rauch A (2009) Lethal cutis laxa with contractural arachnodactyly, overgrowth and soft tissue bleeding due to a novel homozygous fibulin-4 gene mutation. *Clin Genet* 76:276–281.
- Francis F, et al. (2006) Human disorders of cortical development: From past to present. Eur J Neurosci 23:877–893.
- Javaherian A, Kriegstein A (2009) A stem cell niche for intermediate progenitor cells of the embryonic cortex. Cereb Cortex 19(Suppl 1):i70–i77.
- Saito K, et al. (2009) Ablation of cholesterol biosynthesis in neural stem cells increases their VEGF expression and angiogenesis but causes neuron apoptosis. Proc Natl Acad Sci USA 106:8350–8355.
- Stubbs D, et al. (2009) Neurovascular congruence during cerebral cortical development. Cereb Cortex 19(Suppl 1):i32–i41.

Supporting Information

Fietz et al. 10.1073/pnas.1209647109

SI Results and Discussion

Principal Component Analysis. The results of the correlation analyses of the gene-expression patterns between the cortical zones of fetal human or embryonic mouse neocortex (Fig. 1 B and C, and Fig. S3 A and B) were corroborated when the fragments per kilobase of exon per million fragments mapped (FPKM) values were used to calculate the variance between the various 13–16 wk postconception (wpc) human and embryonic day (E) 14.5 mouse samples and depict the results in 2D plots [principal component analysis (PCA)] (Fig. S3 C-E). The highest variance [first principal component (PC)] between all human and mouse samples was accounted for by the difference between the two species, and the second highest (second PC) by the differences between the zones within a given species (Fig. S3C). When the PCA was restricted to either human or mouse, the distribution of the human (Fig. S3D) and mouse (Fig. S3E) samples revealed three clusters for each species, which corresponded to the cortical plate (CP), subventricular zone (SVZ), and ventricular zone (VZ). Separation of the human SVZ (hSVZ) cluster into inner SVZ (ISVZ) and outer SVZ (OSVZ) subclusters was observed along the second PC axis (Fig. S3D), further supporting the notion that ISVZ and OSVZ are most closely related to each other. Taken together, these data show that differences between zones rather than between individuals and, in the case of human, developmental stages accounted for most of the data variance within a given species.

Putative Regulation of Extracellular Matrix-Associated Genes by MicroRNAs. With regard to regulation of the extracellular matrix (ECM)-associated genes (Fig. 4 *A–J* and Dataset S3) at the translational level, we analyzed the 3' UTRs of the ECM-associated genes contained in each cluster for overrepresentation of specific microRNA binding sites (Microcosm Targets database) (1). In contrast to the results of the transcription factor binding site (TFBS) analysis (Dataset S5), very few microRNAs were shared between any of the nine clusters (Dataset S5). Together with the results of the TFBS analysis, these data suggest that the regulation of expression of ECM-associated genes in the developing neocortex that is common to distinct germinal zones (GZs) and conserved between human and mouse occurs at the transcriptional rather than translational level.

Gene Clusters with Similar Expression Patterns in the Fetal Human and Embryonic Mouse GZs. To obtain functional clues from the differential gene expression in the GZs during cortical development, we performed mean-shift clustering of all genes using the fold-changes in their expression. Two different datasets were used as input: (i) E14.5 mouse SVZ (mSVZ) over mVZ, 13-16 wpc hISVZ over hVZ and hOSVZ over hVZ (Fig. S6 A and B); and (ii) 13-16 wpc hOSVZ over hVZ and hOSVZ over hISVZ (Fig. S6 C and D). Clusters with at least 50 genes were then sorted according to similar temporal and zone-ratio patterns of gene expression, using an average-linkage hierarchical algorithm (Fig. S6 A and C). Gene clusters of interest were selected based on whether they (i) contained a gene implicated in neurogenesis, or (ii) showed a concurrent developmental expression time course (Fig. S6 B and D). These clusters were then subjected to functional gene annotation (FGA) analysis (Datasets S6 and S7).

The cluster analysis encompassing mGZs and hGZs yielded five major findings. First, in the majority of the selected clusters (Fig. S6*B*), the greatest difference in the gene-expression foldchanges (SVZs/VZ) was between mouse and human. At this level of analysis, only small differences were found between the hGZs and developmental stages. Second, FGA terms related to neurogenesis, cell differentiation, and synaptic transmission were enriched in the Svet1, ion channels, and intracellular signaling clusters, which contained genes showing a greater extent of upregulation in the mSVZ over mVZ than in the hSVZs over hVZ (Fig. S6B and Dataset S6). Third, FGA terms related to cell division, cell proliferation, and DNA replication were enriched in the Pax6 and DNA replication clusters, which contained genes showing a greater extent of down-regulation in the mSVZ over mVZ than in the hSVZs over hVZ (Fig. S6B and Dataset S6). Taken together, these data further support the notion that the relative abundance of neural stem and progenitor cells (NSPCs) that lack the ability of self-renewal is higher in the mSVZ than in the hSVZs. Fourth, transport vesicle, plasma membrane and endosome were among the FGA terms enriched in the KIAA1324 cluster, which contained genes showing a greater extent of up-regulation in the hSVZs over hVZ than in the mSVZ over mVZ (Fig. S6B and Dataset S6). As the majority of basal progenitors (BPs) in the hSVZs [basal radial glial cells (bRGCs) and transit amplifying progenitors (TAPs)], in contrast to those in the mSVZ [intermediate progenitor cells (IPCs)], undergo self-renewing divisions, the KIAA1324 cluster might contain genes characteristic of self-renewing BPs. Fifth, the Tbr2 and Caveolae clusters, which contained genes down-regulated in the mSVZ over mVZ but up-regulated in the hSVZs over hVZ (Fig. S6B), showed enrichment of FGA terms related to oxidative phosphorylation, translation, and cell adhesion (Tbr2), and vesicle organization and biogenesis (Caveolae) (Dataset S6). These observations provide general clues about the differences in occurrence of specific biological processes in the various GZs of the developing human and mouse neocortex.

The more refined cluster analysis confined to the hGZs (Fig. S6 C and D) yielded the following main findings. The Svet1-containing cluster (565 genes; OSVZ > VZ, OSVZ = ISVZ) showed enrichment of FGA terms related to cell differentiation, neurogenesis, and synaptic transmission (Dataset S7), thus providing further support for the notion that neuron generation is likely to be more prominent in the ISVZ and OSVZ than the VZ. The KIAA1324 (213 genes) and Olig2 (198 genes) clusters (OSVZ > VZ, OSVZ > ISVZ) provided two clues as to OSVZcharacteristic genes. First, the Olig2 cluster also contained genes such as Olig1, Sox10, and myelin basic protein, which have been implicated in oligodendrogenesis and myelination (2-4). Second, genes of the KIAA1324 cluster showed a greater fold-change in expression in the OSVZ vs. VZ than OSVZ vs. ISVZ. Because bRGCs have been detected at higher relative abundance in the hOSVZ than hISVZ (5, 6), the KIAA1324 cluster may include genes potentially characteristic of bRGCs (Dataset S7).

Two clusters with OSVZ > VZ and OSVZ < ISVZ, translation (508 genes) and Tbr2 (64 genes), showed enrichment of FGA terms related to the regulation of translation and to oxygen supply and utilization (Dataset S7). FGA terms enriched in the Pax6 cluster (1,002 genes; OSVZ < VZ, OSVZ = ISVZ) were related to cell division and cell proliferation, ECM-receptor interaction, growth factor receptor activity (e.g., TGF- β , EGF), and Wnt signaling (Dataset S7). The Nestin (660 genes) and integrin α v (128 genes) clusters (OSVZ < VZ, OSVZ > ISVZ) showed enrichment of FGA terms related to FGF and Notch signaling, as well as cell adhesion (Nestin), and to Wnt signaling as well as cell-matrix adhesion and integrin signaling (Integrin α v) (Dataset S7). These data are consistent with, and extend,

previous studies on fetal human and ferret neocortex showing that self-renewing divisions of basal-process bearing NSPCs, in particular apical progenitors (APs) and bRGCs, are linked to Notch and integrin signaling, respectively (5, 6), as well as previous work on embryonic mouse neocortex showing the role of Notch signaling in AP self-renewal (7). FGA terms associated with Wnt signaling were also enriched, along with those related to microtubule-based movement, in the inversin cluster (54 genes; OSVZ = VZ, OSVZ > ISVZ) (Dataset S7). Taken together, these data provide insight into the spatial and temporal occurrence of specific biological processes in the various GZs of the developing human neocortex.

SI Methods

Tissue. Six human fetuses were obtained from the same institutions and with the same ethical policies as described previously (6) [Technische Universität Dresden: 13 wpc (one fetus), 14 wpc (two fetuses), 15 wpc (one fetus), 16 wpc (one fetus); Charité: 16 wpc (one fetus)]. Tissue collection protocols were approved by the Ethics Committees of the Medical Faculty of the Technische Universität Dresden and the Charité. Human fetuses were placed at 4 °C immediately after abortion, and brains were dissected in ice-cold PBS, placed on dry ice, and stored at -80 °C.

Mice were maintained in the animal facility of the Max Planck Institute of Molecular Cell Biology and Genetics. All animal experiments were performed in accordance with German animal welfare legislation. Five C57BL/6 E14.5 mouse embryos from two litters were used. Brains were dissected in ice-cold PBS, placed on dry ice and stored at -80 °C.

Laser-Capture Microdissection. Telencephalon was cut from frozen human brain at a medium position with regard to the rostrocaudal axis, using a sterile scalpel. Frozen human telencephali and mouse brains were mounted in an RNase-decontaminated cryostat (MICROM International) using Tissue-Tek (Sakura Finetek), and 30-µm coronal cryosections were cut and mounted on nuclease-free MembraneSlide NF 1.0 PEN (Zeiss). For human telencephalon, 15-20 consecutive cryosections starting at the scalpel cut were processed further. For mouse telencephalon, 90 cryosections starting rostrally were cut, and cryosection nos. 30-60 were processed further. Slides were rinsed for 30 s in 70%(vol/vol) ice-cold ethanol, incubated for 30 s in 1% cresyl violet in ethanol, and washed twice for 30 s each in 100% (vol/vol) ethanol. Sections were overlayed with Liquid Cover Glass (Zeiss) and immediately subjected to laser-capture microdissection (LCM) using an Axiovert 200M microscope equipped with a PALM LCM MicroBeam system (Zeiss). GZs (mouse: VZ, SVZ; human: VZ, ISVZ, OSVZ) and CP were microdissected under a 20× objective, proceeding across the entire width of the field (~405 μ m) if possible. Along the apical-basal axis of the cortical wall, microdissected areas were as follows: h/mVZ, apical half only (to prevent a possible contribution of adjacent hISVZ/mSVZ); hISVZ and mSVZ, near-complete, using the near-spherical shape and nonradial orientation of the Nissl-stained nuclei as criteria and excluding the inner fiber layer and intermediate zone, respectively; hOSVZ, near-complete, using the elongated shape and radial orientation of the nuclei as criteria and excluding inner and outer fiber layers; CP, nearcomplete, pooling deep and upper layers when present. For each cryosection, multiple $20 \times$ fields of the dorsolateral telencephalon were subjected to microdissection to harvest the zone of interest almost completely. The 20x fields from multiple cryosections were collected using AdhesiveCap (Zeiss) until $\sim 2 \times 10^6 \ \mu m^2$ microdissected area were contained per single tube. Four tubes per zone were obtained from human telencephalon, two tubes per zone from mouse. Each tube received 350 µL of RLT buffer of RNeasy Micro Kit (Qiagen), was vortexed for 30 s, placed upside-down followed by incubation for 30 min at room temperature, centrifuged for 5 min at 14,800 \times g in a Biofuge pico (Heraeus), and stored at -80 °C.

RNA isolation and cDNA Library Preparation. The laser-microdissected tissue samples were thawed, briefly vortexed, and total RNA was isolated using the Qiagen RNeasy Micro Kit according to the manufacturer's instructions for microdissected cryosections with the following modifications. In step 5, 1.5 volumes of ethanol were added to the tissue lysate; in step 6, for a given zone, the contents of four (human) or two (mouse) tubes was applied to one RNeasy MinElute spin column; in steps 7 and 10, RWT buffer of the Qiagen RNeasy Mini Kit was used instead of RW1 buffer; in step 12, the final washing step using 80% (vol/ vol) ethanol was omitted; in step 14, the column was eluted with two times 20 μ L of 37 °C-warm water. The integrity and concentration of the total RNA thus obtained (range 27.2–82.2 ng) were determined using the Agilent RNA 6000 Pico Kit and a Bioanalyzer (Agilent 2100).

For each zone and human fetus/mouse embryo, a cDNA library was generated from 25 ng of the total RNA using the Illumina TruSeq RNA sample preparation kits according to the manufacturer's instructions. The size distribution and concentration of the cDNAs obtained were determined using the Agilent DNA 1000 Kit and a Bioanalyzer (Agilent 2100).

RNA Sequencing and Quantification. cDNA libraries (10 nM each, starting from the Bioanalyzer data, which takes the DNA size at the fluorescence peak as molecular mass value) were pooled and sequenced on an Illumina Genome Analyzer IIx according to the manufacturer's instructions for Multiplex Single Read sequencing and using 76 + 7 cycles, except that an indexed φ X174 control library was added, yielding about 1% of sequencing reads in each lane. The $\varphi X174$ control reads were aligned to the corresponding reference sequence to obtain a training dataset for the base caller IBIS (8), which was then used to recall bases and quality scores. Raw reads were separated by index, allowing up to one substitution in the index sequence (9) but requiring that all quality scores in the index read were ≥ 10 (PHRED-scale) (10). Reads of the same sample on different sequencing lanes were combined, subjected to adapter trimming, and reads shorter than 60 nucleotides or with more than five bases below a quality score of 15 (PHRED-scale) (10) were removed. The processed reads were aligned to the human reference genome (hg19/GRCh37, release date Feb. 2009, excluding additional haplotypes) (11) and mouse reference genome (mm9/NCBI37, release date Jul. 2007) (12) using TopHat v1.2.0 (13). For a given zone and (if applicable) developmental stage, genes were considered not expressed if no reads were detectable. Genes of the Ensembl release 61 (14) were quantified using Cufflinks v0.9.3 (15), with the resulting RNAseq data being expressed as FPKM values. Orthologous genes were assigned based on annotated Ensembl one-to-one orthologs. Correlation analyses, principal component analyses, and correlation dendrograms were calculated using R (http://www.r-project.org/). Differences between the Spearman's rank correlation coefficients of the various zones were tested for their significance using one-way ANOVA followed by the Tukey-Kramer test for multiple comparisons.

Differential Gene Expression Analysis and Functional Annotation Clustering. Differentially expressed genes were determined using DESeq. (16), with FPKM values as data input to allow an accurate comparison between orthologous genes of different length. We performed this analysis for the sum of all transcripts present for any given gene, and did not distinguish between different transcripts (splice variants). As FPKM calculation may produce noninteger values (which cannot be used as input to DESeq), FPKM values were multiplied by 10 and then rounded to integer values, which were used as input to DESeq. Given that the 13 wpc hOSVZ was not distinct from the 13–16 wpc hISVZ in PCA analysis and the correlation dendrogram (Fig. S3), DESeq analysis was performed only for the 14–16 wpc human samples. Raw *P* values were Benjamini Hochberg-corrected, and a gene was defined as differentially expressed if the resulting adjusted *P* value was <0.05 and the residuals were <15 (16). Numbers of differentially expressed genes in the various sectors of Venn diagrams were calculated using R.

Functional annotation clustering of the differentially expressed genes was performed using DAVID v6.7 (17). When comparing hGZs vs. mGZs, genes not differentially expressed between any of the zones of the human neocortex, or of the mouse neocortex, were removed from the input dataset. Only those functional annotation terms associated with the various sets of differentially expressed genes were clustered that were significantly enriched (*P* value of <0.05) compared with the functional annotation terms associated with the total population of genes expressed in at least one of the human or mouse samples. DAVID default settings were used except that stringency was set to low.

Gene-Expression Profile Clustering and Gene Annotation Enrichment Analysis. For Fig. 4, gene-expression profile clustering was performed for those genes for which the DAVID FGA terms included ECM, extracellular region, ECM-receptor interaction, proteoglycan, or basement membrane (ECM-associated genes). Log₂-transformed FPKM fold-changes in mean expression in the 14-16 wpc hVZ, hISVZ, hOSVZ, and hCP (five fetuses; i.e., 20 samples) and the E14.5 mVZ, mSVZ, and mCP (five embryos; i.e., 15 samples) relative to the mean expression of each of the genes in all human or mouse zones were used as input. An FPKM value of ≤ 1 is typically considered as a marginal geneexpression level (18-20). Hence, only genes with an FPKM value ≥ 1 in at least 5 of the 35 samples were included. Moreover, 0 FPKM values, and FPKM values <1 were set to 1, to be able to calculate a meaningful fold-change, and to avoid erratic deviations in fold-change values, respectively, when low FPKM values were in the denominator. The resulting dataset was clustered using a K-nearest neighbor mean-shift (KNN-MS) algorithm (21, 22) with uncentered Pearson correlation, implemented in the VT-shift software package (23), to determine the similarity between expression profiles. The free parameter K (number of neighbors) was arbitrarily set to 45. Datapoints that over the course of mean-shift iterations converged closer than a correlation of 0.9 were assigned to one cluster. Clusters were grouped according to expression profiles using an average-linkage hierarchical algorithm. Genes of each of the clusters were subjected to gene annotation enrichment analysis using the VT-shift software package. This revealed those Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) annotation terms associated with the genes in the selected clusters that were significantly enriched (P value of <0.05) compared with the annotation terms associated with the total population of genes expressed in the dataset subjected to KNN-MS clustering.

For Fig. S6, showing the clustering of genes according to expression profiles, two different datasets were used as input. The first dataset contained the log₂-transformed FPKM fold-changes in gene expression in the E14.5 mSVZ relative to mVZ, and in the hISVZ relative to hVZ and hOSVZ relative to hVZ from 13 to 16 wpc. The second dataset contained the log₂-transformed FPKM fold-changes in gene expression in the hOSVZ relative to hVZ and hOSVZ relative to hISVZ from 13 to 16 wpc. In the case of 14 and 16 wpc, when two fetuses each had been analyzed, the mean of the two FPKM values and FPKM values <1 were set to 1, to be able to calculate a meaningful fold-change and to avoid erratic deviations in fold-change values, respectively, when low FPKM values were in the denominator.

In the first dataset, genes not differentially expressed between any of the zones of the human neocortex or of the mouse neocortex were removed from the input dataset. In addition, to give equal weight to the mouse and human data, given that the mouse data were one developmental stage and one GZ comparison only, whereas the human data were four developmental stages and two GZ comparisons, the data were normalized as follows. First, a PCA was performed using the log₂-transformed FPKM foldchanges in gene expression of the human samples, and the SD was determined along the first PC. Next, the log₂-transformed FPKM fold-changes in gene expression of the mouse sample were rescaled so that their SD equaled the average SD for the hGZs.

In the second dataset, genes significantly up-regulated in the hCP compared with the hOSVZ, as determined by DESeq analysis, were removed from the input dataset to perform the clustering without the contribution of genes predominantly expressed in neurons.

Each of the resulting two datasets was clustered using a KNN-MS algorithm (21, 22) with uncentered Pearson correlation, implemented in the VT-shift software package (23), to determine the similarity between expression profiles. If the clustering encompassed mGZs and hGZs, the free parameter K (number of neighbors) was arbitrarily set to 5. If clustering encompassed hGZs only, K was arbitrarily set to 34. Datapoints that over the course of mean-shift iterations converged closer than a correlation of 0.9 were assigned to one cluster. Clusters containing at least 50 genes were grouped according to expression profiles using an average-linkage hierarchical algorithm.

Selected clusters were subjected to gene annotation enrichment analysis using the VT-shift software package. This revealed those GO and KEGG annotation terms associated with the genes in the selected clusters that were significantly enriched (P value of <0.05) compared with the annotation terms associated with the total population of genes expressed in the dataset subjected to KNN-MS clustering.

TFBS Analysis. Overrepresented TFBSs were identified using oPOSSUM v3.0 (24, 25) in the 5 kb up- and downstream of the transcription start site of the ECM-associated genes in each of the clusters obtained by mean-shift clustering. Matrix-based TF binding profiles were derived from JASPAR. A conservation cutoff of 40% and a matrix score threshold of 75% were used. A Z-score of >5 and a Fisher score of >5 were used to identify TFBS overrepresented in the input gene set (i.e., the ECM-associated genes) compared with the background gene set (i.e., all genes in the oPOSSUM database). For a given cluster, the overrepresented TFBS obtained were only considered further if the expression of the respective TF matched the median gene expression pattern of that cluster: that is, if the TF showed a significantly (P < 0.05) higher expression level as identified by DESeq in the cortical zones with a median \log_2 fold-change >0 compared with those with a median \log_2 fold-change <0.

MicroRNA Binding Site Analysis. To test for overrepresentation of specific microRNA (miRNA) binding sites (miRBS) in gene clusters, miRNA gene target predictions were taken from the Microcosm Targets database, v5 (http://www.ebi.ac.uk/enrightsrv/microcosm/) (1). All genes in the human and mouse genome with at least one predicted miRBS in their 3' UTR as determined at the transcript level were taken as human and mouse background set, respectively, setting the frequency of occurrence for each predicted miRBS in any given gene to 1, irrespective of whether that miRBS was predicted to occur once or more in the transcripts of that gene. In the case of the human background set, only the predicted human gene targets of human miRNAs were included. Similarly, for the mouse background set, only the predicted mouse gene targets of mouse miRNAs were included. For each individual gene cluster, the miRBS predicted for the human and mouse genes in that cluster, with the frequency of occurrence being calculated as for the background sets, were tested for overrepresentation compared with the human and mouse background set, respectively, using Fisher exact probabilities (26) calculated using R.

Quantitative PCR. Remaining total RNA isolated from hVZ, hISVZ, hOSVZ, and hCP was used as starting material, pooling the 13-16 wpc samples. Single-strand cDNA was prepared from 15 ng each of the four total RNA pools using the SuperScript III First Strand Synthesis System (Invitrogen) according to the manufacturer's instructions. Quantitative PCR for selected genes and, as an internal standard, human β -actin was performed using ABsolute QPCR SYBR Green Mix (Thermo Scientific) on an Mx3000 Real-Time PCR system (Stratagene). Primers were designed using Primer-BLAST software (http://www.ncbi.nlm.nih.gov/tools/primerblast/). Primer sequences were as follows: KIAA1324-fwd, 5'-CT-GAGCCTGGGCACAGCCAC-3', KIAA1324-rev, 5'-TCCCG-GTGGACTCAGCAGCA-3', Integrin av -fwd, 5'-AATGACA-CGGTTGCCGGGCA-3', Integrin av -rev, 5'-TCACTGAGG-GCAAGATCCCGCT-3', FGFR1-fwd, 5'-AACCAACGGCCG-ACTGCCTG-3', FGFR1-rev, 5'-ACACCGGGGTATGGGG-AGCC-3', EGFR-fwd, 5'-ACCAGTGTGCTGCAGGCTGC-3', EGFR-rev, 5'-GGGGGGCAGGTGTCCTTGCAC-3', Talin1-fwd,

- Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ (2008) miRBase: Tools for microRNA genomics. Nucleic Acids Res 36(Database issue):D154–D158.
- Woodruff RH, Tekki-Kessaris N, Stiles CD, Rowitch DH, Richardson WD (2001) Oligodendrocyte development in the spinal cord and telencephalon: Common themes and new perspectives. *Int J Dev Neurosci* 19:379–385.
- Simons M, Trotter J (2007) Wrapping it up: The cell biology of myelination. Curr Opin Neurobiol 17:533–540.
- Li H, Richardson WD (2008) The evolution of Olig genes and their roles in myelination. Neuron Glia Biol 4:129–135.
- Hansen DV, Lui JH, Parker PR, Kriegstein AR (2010) Neurogenic radial glia in the outer subventricular zone of human neocortex. Nature 464:554–561.
- Fietz SA, et al. (2010) OSVZ progenitors of human and ferret neocortex are epitheliallike and expand by integrin signaling. Nat Neurosci 13:690–699.
- Kawaguchi A, et al. (2008) Single-cell gene profiling defines differential progenitor subclasses in mammalian neurogenesis. *Development* 135:3113–3124.
- Kircher M, Stenzel U, Kelso J (2009) Improved base calling for the Illumina Genome Analyzer using machine learning strategies. *Genome Biol* 10:R83.
- Meyer M, Kircher M (2010) Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb Protoc* 2010:pdb prot5448.
- Ewing B, Green P (1998) Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res* 8:186–194.
- 11. Lander ES, et al.; International Human Genome Sequencing Consortium (2001) Initial sequencing and analysis of the human genome. *Nature* 409:860–921.
- Waterston RH, et al.; Mouse Genome Sequencing Consortium (2002) Initial sequencing and comparative analysis of the mouse genome. Nature 420:520–562.
- Trapnell C, Pachter L, Salzberg SL (2009) TopHat: Discovering splice junctions with RNA-Seg. *Bioinformatics* 25:1105–1111.
- Hubbard TJ, et al. (2009) Ensembl 2009. Nucleic Acids Res 37(Database issue): D690–D697.
- 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.

5'-AAACGCTGGGCTGCGTCTCC-3', Talin1-rev, 5'-AGCC-GGCAATGAGCTGTGCA-3', Pleiotrophin-fwd, 5'-GACTGT-GGGCTGGGCACACG-3', Pleiotrophin-rev, 5'- TGCACTCC-GCGCCAAATTGC-3', Tenascin C-fwd, 5'-ACGGCCAGTG-CATCTGCCAC-3', Tenascin C-rev, 5'-GGCAGGAGTGCTG-GCCACAG-3', Ptprz1-fwd, 5'-TCGTTTTGGGAAGCAGGC-TGC-3', Ptprz1-rev, 5'-ACTGTGTCTGTGCAGGGAGGA-3', β-actin -fwd, 5'-CCACGGCTGCTTCCAGCTCC-3', β-actin -rev, 5'-GGGCAGCGGAACCGCTCATT-3'. PCR conditions were as described previously (27) (95 °C for 15 min, followed by 40 cycles of 95 °C for 30 s, 60 °C for 30 s, and 72 °C for 30 s). Foldchanges in the expression of a selected gene relative to β-actin, and between two zones, were determined using the $\Delta\Delta$ Ct method (28).

Immunocytochemistry. The proportion of Pax6⁺, Tbr2⁺, and Tbr1⁺ nuclei in the hVZ, hISVZ, hOSVZ, and hCP was determined after immunocytochemistry on 12- to 14-µm paraffin sections of formalin-fixed 16 wpc human neocortex, using procedures and tissue as described previously (6); the antibodies used were against Pax6 (Covance, PRB-278P; 1:200), Tbr2 (Abcam, ab23345; 1:200), and Tbr1 (Abcam, ab31940; 1:100).

- Anders S, Huber W (2010) Differential expression analysis for sequence count data. Genome Biol 11:R106.
- Huang da W, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4:44–57.
- Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods 5:621–628.
- Brooks MJ, Rajasimha HK, Roger JE, Swaroop A (2011) Next-generation sequencing facilitates quantitative analysis of wild-type and Nrl(-/-) retinal transcriptomes. *Mol Vis* 17:3034–3054.
- Raskatov JA, et al. (2012) Modulation of NF-xB-dependent gene transcription using programmable DNA minor groove binders. Proc Natl Acad Sci USA 109:1023–1028.
- Cheng Y (1995) Mean shift, mode seeking, and clustering. IEEE Trans Pattern Anal Mach Intell 17:790–799.
- Georgescu B, Shimshoni I, Meer P (2003) Mean shift based clustering in high dimensions: A texture classification example. Proceedings of the Ninth IEEE International Conference on Computer Vision 1, Nice, France, 13-16 October 2003:456–463.
- Collinet C, et al. (2010) Systems survey of endocytosis by multiparametric image analysis. Nature 464:243–249.
- Ho Sui SJ, et al. (2005) oPOSSUM: Identification of over-represented transcription factor binding sites in co-expressed genes. Nucleic Acids Res 33:3154–3164.
- Ho Sui SJ, Fulton DL, Arenillas DJ, Kwon AT, Wasserman WW (2007) oPOSSUM: Integrated tools for analysis of regulatory motif over-representation. *Nucleic Acids Res* 35(Web Server issue):W245-52.
- Fleiss JL (1981) Statistical Methods for Rates and Proportions (John Wiley, New York), pp 38–46.
- Ettinger AW, et al. (2011) Proliferating versus differentiating stem and cancer cells exhibit distinct midbody-release behaviour. Nat Commun 2:503.
- Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25: 402–408.



15 w.p.c. E14.5

Fig. S1. Sample preparation for RNA-Seq. (*A*) Nissl staining on 30-µm cryosections of 15 wpc human and E14.5 mouse neocortex. Dashed rectangles, representative areas isolated by LCM; IFL, inner fiber layer; OFL, outer fiber layer; IZ, intermediate zone. (Scale bars, 20 µm.) (*B*) Agilent RNA 6000 electrophoresis (picochip) of total RNA extracted from fetal human and embryonic mouse neocortex after LCM. (*Upper*) Representative electropherogram of human neocortex total RNA. (*Lower*) RNA integrity numbers for the cortical zones of 13–16 wpc human (six fetuses) and E14.5 mouse (five embryos). (*C*) Representative electropherogram obtained by Agilent DNA 1000 electrophoresis of a cDNA library generated from 25 ng of total RNA from the fetal human neocortex. (*D*) Comparison of the total number of sequencing reads (blue) with the total number of reads aligned (green) and of reads uniquely aligned (red) to the human (*Left*) and mouse (*Right*) genome.



Fig. S2. Validation of RNA-Seq data by quantitative PCR. (*A*) Cartoon illustrating the major cell types (APs in blue, BPs in orange and red, neurons in green) in the different GZs and the CP of the developing human neocortex. (*B*) Mean expression levels (FPKM) in the 13–16 wpc hVZ, hISVZ, hOSVZ, or hCP of selected genes as determined by RNA-Seq. (six fetuses). (C) Quantitative PCR (qPCR) in triplicate for the same selected genes as in *B* using pools of total RNA extracted from 13–16 wpc hVZ, hISVZ, hOSVZ, or hCP. FPKM-values (*B*) and normalized C_T-values (C) obtained for the OSVZ were divided by those obtained for the VZ (blue), ISVZ (orange), and CP (green). Data are log-transformed (base 2) means of the resulting fold-changes; bars indicate SEM (*B*, *n* = 6) or SD (*C*, *n* = 3).



Fig. S3. Correlation dendrogram and principal component analysis of fetal human and embryonic mouse cortical zones. (*A* and *B*) Correlation dendrograms showing the relationship of gene-expression levels between the indicated zones of the fetal human neocortex at the indicated gestational weeks (*A*, six fetuses) and of the E14.5 mouse neocortex (*B*, five embryos). Correlation analysis was performed using an average-linkage algorithm; *r_{sr}* Spearman's rank correlation coefficient. (*C*–*E*) PCA of gene-expression levels of the indicated zones (see key) of the fetal human neocortex at the indicated gestational weeks (*C* and *D*: 13–16 wpc, circles) and of the E14.5 mouse neocortex (*C* and *E*: triangles).



Fig. S4. Comparison of the expression of selected genes in the fetal human and embryonic mouse cortical zones. (*A*) Cartoon illustrating the major cell types (APs in blue, BPs in orange and red, neurons in green) in the different GZs and the CP of the developing human neocortex. (*B*–*D*) Comparison of the expression level of selected genes characteristic of APs or bRGCs (*B*), TAPs/IPCs and TAP/IPC-generation (*C*), or neurons (*D*) in the VZ (blue), ISVZ (orange), OSVZ (red), and CP (green) of 13–16 wpc human neocortex (six fetuses). For each gene, the FPKM value in a given zone is expressed as percentage of the sum of values in all four zones; bars indicate SEM. NF-M, neurofilament medium polypeptide. (*E*–*G*) Quantification of Pax6⁺, Tbr2⁺, and Tbr1⁺ nuclei in the 16 wpc hVZ, hISVZ, hOSVZ, and hCP, expressed as percentage of total nuclei in the respective zone as revealed by DAPI staining. Data are the mean of five (*E*), nine (*F*), and two (*G*) 150-µm-wide images counted (from the same fetus); bars indicate SD (*E* and *F*) or the variation of the individual values from the mean (*G*). (*H*) Cartoon illustrating the major cell types (APs in blue, BPs in orange and red, neurons in green) in the different GZs and the CP of the developing mouse neocortex. (*I*–*K*) Comparison of the expression level of selected genes, characteristic of APs (*I*), TAPs/IPC-generation (*J*), and neurons (*K*) in the VZ (blue), SVZ (orange), and CP (green) of E14.5 mouse neocortex (five embryos). For each gene, the FPKM value in a given zone is expressed as percentage of the sum of values in given join the different GZ sone the percentation (*J*), and neurons (*K*) in the VZ (blue), SVZ (orange), and CP (green) of E14.5 mouse neocortex (five embryos). For each gene, the FPKM value in a given zone is expressed as percentage of the sum of values in all three zones; bars indicate SEM.



Fig. 55. Differential gene-expression analysis of embryonic mouse cortical zones and comparison with fetal human cortical zones. (*A*) DESeq scatter plot showing pair-wise comparisons of gene expression between the indicated zones of E14.5 mouse neocortex. Each dot represents the mean (five embryos) expression level for a given gene, with differentially expressed genes (false-discovery rate <0.05) shown in green. Numbers above (below) the diagonal lines refer to the differentially expressed genes up-regulated in the zone indicated on top (*Right*). These numbers include the genes expressed in only one of the two zones compared (green dots on vertical and horizontal 0.1 lines). (*B*) Venn diagrams showing the numbers of differentially expressed genes (see *A*) that are up-regulated (*Left*, >) and down-regulated (*Right*, <) in a given zone compared with the other zones of the E14.5 mouse neocortex. (*C*) DESeq scatter plot showing pair-wise comparisons of gene expression between the indicated zones of the 14–16 wpc human (five fetuses) and E14.5 mouse (five embryos) neocortex. Further specifications are as in *A*.



Fig. 56. Mean-shift gene-expression clustering of fetal human and embryonic mouse GZs. (*A* and *B*) Log_2 -transformed FPKM fold-changes in gene expression in E14.5 mSVZ/VZ (five embryos) and 13–16 wpc hISVZ/VZ and hOSVZ/VZ (six fetuses) were subjected to nonhierarchical KNN-MS cluster analysis, which yielded a total of 359 clusters. (*A*) Heat map showing the 16 clusters containing at least 50 genes, sorted hierarchically using an average-linkage algorithm. Each gene is represented by a single row and each developmental stage by a single column; only genes with a phenoscore of at least 0.95 are shown. Fold-changes range from 9.9 (red) to -10.7 (green). (*B*) Plots showing the median fold-changes in expression of all genes (numbers in parentheses) contained in 10 selected clusters, as indicated in *A*. Designations were chosen based on selected single genes or biological features, as deduced from enriched functional annotation terms present in a given cluster. (*C* and *D*) Log_2 -transformed FPKM fold-changes in gene expression in 13–16 wpc hOSVZ/VZ and hOSVZ/VZ (six fetuses) were Legend continued on following page

DNAS

subjected to KNN-MS cluster analysis, which yielded a total of 735 clusters. (C) Heat map showing the 62 clusters containing at least 50 genes, sorted hierarchically using an average-linkage algorithm. Fold-changes range from 11.5 (red) to -9.5 (green). Further specifications are as in A. (D) Plots showing the median fold-changes in expression of all genes (numbers in parentheses) contained in 10 selected clusters as indicated in C. Further specifications are as in B.

Dataset S1. Genes differentially expressed between cortical zones

Dataset S1 (xlsx)

Differentially expressed genes that are up-regulated (>) or down-regulated (<) as identified by DESeq (P < 0.05) in a given zone compared with the other zones of either the 14–16 wpc human neocortex or the E14.5 mouse neocortex, or in the 14–16 wpc hOSVZ compared with all other zones of the 14–16 wpc human and E14.5 mouse neocortex.

Dataset S2. Functional annotations of differentially expressed genes between cortical zones

Dataset S2 (xlsx)

Differentially expressed genes that are up-regulated (>) or down-regulated (<) in a given zone compared with the other zones of either the 14–16 wpc human neocortex (Fig. 2B) or the E14.5 mouse neocortex (Fig. S5B), or between the indicated zones of the 14–16 wpc human and E14.5 mouse neocortex (Fig. S5C), were analyzed for significantly enriched FGA terms, which were clustered using DAVID. For each of the distinct sets of differentially expressed genes, the resulting clusters of the significantly enriched FGA terms are listed.

Dataset S3. ECM-associated genes contained in each cluster obtained by mean-shift clustering

Dataset S3 (xlsx)

Log₂-transformed FPKM fold-changes of ECM-associated genes in the 14–16 wpc human and E14.5 mouse cortical zones relative to the mean expression of these genes in all human or mouse zones were subjected to mean-shift clustering, which yielded a total of nine clusters (Fig. 4 A–J). The genes contained in each of the nine clusters are listed. Violet, collagen; green, laminin; blue, proteoglycan; red, integrin; gray, growth factor/morphogen.

Dataset S4. Functional annotations of ECM-associated genes of each cluster obtained by mean-shift clustering

Dataset S4 (xlsx)

ECM-associated genes from each of the nine clusters obtained by mean-shift clustering (Fig. 4 A–J and Dataset S3) were subjected to gene annotation enrichment analysis. For each cluster, those GO and KEGG annotation terms are listed that were significantly enriched (P < 0.05) compared with the annotation terms associated with the total population of genes subjected to mean-shift clustering.

Dataset S5. Overrepresented transcription factors and microRNA binding sites in sets of ECM-related genes

Dataset S5 (xlsx)

Promoter regions or 3' UTRs of the human and/or mouse ECM-associated genes contained in each of the nine clusters obtained by mean-shift clustering (Fig. 4 A–J and Dataset S3) were analyzed for overrepresentation of binding sites for either TFs showing higher expression levels in the relevant cortical zones, or microRNAs, respectively. Boxed fields, TFBSs or miRNA binding sites found in more than one GZ cluster; yellow fields, TFBSs or miRNA binding sites found in both human and mouse cortical zones of the same cluster.

Dataset S6. Functional annotations of genes of each cluster obtained by mean-shift clustering encompassing all mouse and human orthologous genes

Dataset S6 (xlsx)

 Log_2 -transformed FPKM fold-changes in gene expression in E14.5 mSVZ/VZ and 13–16 wpc hISVZ/VZ and hOSVZ/VZ were subjected to mean-shift clustering, which yielded a total of 359 clusters. Genes of each of the 10 selected clusters (Fig. S6 *A* and *B*) were then subjected to gene annotation enrichment analysis. For each of the clusters, those GO and KEGG annotation terms are listed that were significantly enriched (P < 0.05) compared with the annotation terms associated with the total population of genes subjected to mean-shift clustering.

Dataset S7. Functional annotations of genes of each cluster obtained by mean-shift clustering of human genes

Dataset S7 (xlsx)

PNAS PNAS

Log₂-transformed FPKM fold-changes in gene expression in 13–16 wpc hOSVZ/VZ and hOSVZ/ISVZ were subjected to mean-shift clustering, yielding a total of 735 clusters. Genes of each of the 10 selected clusters (Fig. S6 C and D) were subjected to gene annotation enrichment analysis. For each of the clusters, those GO and KEGG annotation terms are listed that were significantly enriched (P < 0.05) compared with the annotation terms associated with the total population of genes subjected to mean-shift clustering.