Distinct Molecular Programs Define Human Radial Glia Subtypes During Human Cortical Development

Tomasz J. Nowakowski, PhD, and Alex A. Pollen, PhD

Department of Neurology
Eli and Edythe Broad Center of Regeneration Medicine and Stem Cell Research
University of California, San Francisco
San Francisco, California
Introduction
The cell and the genome represent two of the fundamental units in biology. The evolution of multicellular organisms relied on the ability of individual cells to adopt distinct morphologies, physiological properties, and functional states by utilizing the same genome in different ways through diverse transcriptional regulation mechanisms. However, until recently, technical limitations prevented large-scale analysis of gene expression signatures in individual cells. Instead, unbiased sampling of the transcriptome required combining thousands or millions of cells. The average landscape of gene expression from a mixture of cells may obscure the distinct molecular features of diverse cell types or states. Recent advances in microfluidics and molecular biology now enable routine transcriptional profiling from single cells using high-throughput sequencing. These advances support powerful strategies for decoding the mutations, genes, and pathways that distinguish diverse cell lineages and cell types. Here we discuss how we have applied single-cell RNA sequencing (scRNA-seq) to studying the development of the human nervous system.

Astonishing Diversity
More than 100 years ago, Ramon y Cajal appreciated the astonishing diversity of cells in the nervous system, using silver stains to visualize single cells (Cajal, 2002). Cajal likened himself to an entomologist searching for colorful butterflies, “whose beating of wings may one day reveal to us the secrets of the mind” (Garcia-Lopez, 2012). We now know that the human cerebral cortex contains more than 16 billion neurons and even more non-neuronal cells (Herculano-Houzel, 2009) distributed across dozens of cortical areas, and that a single region of mouse cortex may contain more than 40 cell types (Petilla Interneuron Nomenclature Group et al., 2008; Zeisel et al., 2015; Tasic et al., 2016) distributed across six layers. These diverse neuronal cell types emerge from an initially homogenous neuroepithelium during embryonic development.

In human development, the primordial forebrain emerges at around gestational week 7 (GW7) after sequential waves of induction and patterning of the neural tube. Neuroepithelial stem cells undergo symmetric expansive divisions as the neural tube grows before giving rise to the founder neural stem cell population of radial glia along the ventricular zone (VZ) at~GW11. These ventricular radial glia (vRG) can divide asymmetrically to generate intermediate progenitor cells (IPCs) that will subsequently give rise to postmitotic neurons (Fig. 1A). In early development, radial glia in the dorsal forebrain generate glutamatergic neurons that migrate to deep cortical layers and generally project to subcortical regions, whereas neurons born at later stages migrate to upper cortical layers and tend to project to neurons within the cortex. In contrast, inhibitory cortical interneurons are generated by radial glia in the ventral forebrain and migrate tangentially toward the cortical plate (Fig. 1B).

The human neocortex also contains a second population of radial glia cells concentrated in the outer subventricular zone (OSVZ) that emerge around GW14, and are rarely observed in mouse. These outer radial glia (oRG) cells differ from vRG cells with respect to position, morphology, and dynamic cell behavior during cell division (Hansen et al., 2010) (Fig. 1C). In the VZ, vRG cells are bipolar and possess apical processes that directly contact the lateral ventricle and transduce signals from the CSF. These signals are critical for their survival, proliferation, and neurogenic capacity (Lehtinen et al., 2011). In synchrony with the cell cycle, cell bodies of vRG cells undergo interkinetic nuclear migration (INM) (Fig. 1C). In contrast, oRG cells have unipolar morphology, lack apical junctions, and undergo a distinct dynamic behavior—mitotic somal translocation (MST)—that directly precedes cytokinesis (Hansen et al., 2010) (Fig. 1C). Thus, vRG and oRG cells reside in distinct niches defined by differences in anatomical location, provision of growth factors, and their ability to be further distinguished based on morphology and cell behavior (Fietz et al., 2010). Although oRG cells are able to generate the majority of cortical neurons (Smart et al., 2002), the molecular features sustaining neural stem cell properties of oRG cells in the OSVZ niche are largely unknown, and the long-term proliferative capacity of these cells has not been examined.

Applying Single-Cell RNA Sequencing to Developmental Heterogeneity
We sought to identify genes and pathways distinguishing vRG and oRG cells during cortical development (Pollen et al., 2015). Previous studies attempted to find markers for oRG cells by comparing gene expression between microdissected samples (Fietz et al., 2012; Miller et al., 2014) or between cell populations expressing particular surface proteins (Florio et al., 2015; Johnson et al., 2015). However, inferring the oRG signature in heterogeneous tissue...
Figure 1. Single-cell mRNA-seq reveals molecular heterogeneity among cells in the developing human brain. A, B, Schematics show major modes of neurogenesis of cortical neurons. C, Comparison of the closely related radial glia subpopulations. D, Strategy for unbiased sampling of single cells from the germinal zone. E, Unbiased clustering of 393 single cells highlights major cell populations distinguished by nonoverlapping patterns of marker gene expression. F, Overlapping patterns of radial glia and neuronal marker gene expression highlight IPCs as the intermediate step during neurogenesis. G, H, Iterative analysis of the molecular variation within classically defined radial glia reveals cell cycle as the dominant source of transcriptional variation. I, CRYAB is a candidate G1 marker in vRG cells expressed in cells that do not express M-phase marker pHH3 or incorporate S-phase-specific BrdU tracer. J, K, Stem cell niche contributed to transcriptional variation within radial glia and highlights candidate vRG and oRG marker genes. TPM, transcripts per million. L–N, Ideal vector correlation analysis identifies radial glia subtype–specific genes. CP, cortical plate; IZ, intermediate zone; MZ, marginal zone; SG, subpial granular layer; SP, subplate. Adapted with permission from Pollen AA et al. (2015) Molecular identity of human outer radial glia during cortical development, Cell 163:55–67, their Figs. 1, 2, and 3. Copyright 2015, Elsevier.
proved challenging. We chose to examine gene expression in single cells captured from the VZ and OSVZ without additional enrichment steps for specific cell types. We then isolated radial glia from other cell types in silico by analyzing thousands of genes that vary across cell types (Zeisel et al., 2015) and examined the major sources of variation among radial glia. In contrast to this approach of capturing diverse cell types without enrichment, another elegant study used flow cytometry to enrich for cells in the G1 phase of the cell cycle that expressed radial glia markers; the study then analyzed single-cell gene expression specifically among radial glia to identify a similar set of oRG marker genes (Thomsen et al., 2016).

To capture single cells, we used the Fluidigm C1 system (Fluidigm, South San Francisco, CA) and analyzed single cells from paired regions of VZ and OSVZ microdissected from three samples (Fig. 1D). After sequencing, we further analyzed 393 nonoutlier cells based on the number of genes detected, a metric that largely overlapped with several outlier removal approaches. To determine cell-type identity, we performed principal component (PC) analysis and used expectation-maximization clustering to group cells based on their position in PC space. Major clusters of cells were visualized using t-distributed stochastic neighbor embedding. Based on the expression of known marker genes, we interpreted groups to represent cells along the cortical excitatory lineage and inhibitory interneurons (Fig. 1E).

Among the excitatory lineage cells (Fig. 1F), we identified groups of cells that robustly expressed markers of human radial glia (SLC1A3, PAX6, SOX2, HES1, and GLI3; yellow bar, Fig. 1F) and groups of cells that expressed markers of postmitotic neurons (RBFOX1, NEUROD2, and STMN2). In addition, we identified groups of cells that retained a reduced level of some radial glia markers but also expressed early neuronal markers such as STMN2 and NEUROD6. These cells were defined by a gene expression module that included known and novel markers for intermediate progenitor cells, including EOMES (TBR2), ELAVL4, NEUROG1, NEUROD1, NEUROD4, PPP1R17, and PENK (magenta bar, Fig. 1F) (Hevner et al., 2006; Kawaguchi et al., 2008). This analysis demonstrated that scRNA-seq successfully recovers the major sources of molecular variation expected to distinguish cells at different stages of projection neuron differentiation.

**Major Sources of Transcriptional Variation Among Radial Glia Relate to Cell Cycle and Stem Cell Niche**

We next analyzed variation in gene expression across 107 cells that robustly expressed canonical markers of radial glia but not markers of other major cell populations (Fig. 1G). Given the high proliferative capacity of neural progenitors, we anticipated that cell cycle would be a major source of transcriptional variation across single cells we profiled. Indeed, genes involved in cell-cycle regulation, mitosis, and DNA replication explained most variation along PC1, PC2, and PC4. Clustering radial glia based on variation along these axes revealed cell groups representing G1, G1/S checkpoint, and G2/M checkpoint (Fig. 1H), and we confirmed that select markers were specific for stages of the radial glia cell cycle (Fig. 1I). Thus, differentiation and cell cycle are major sources of transcriptional heterogeneity among cells in the germinal zone, and single-cell analysis reveals novel molecular features of these states.

Given the distinct morphologies and mitotic behaviors of vRG and oRG cells, we hypothesized that niche occupancy would also contribute to variation among radial glia. Indeed, we found that the spatial source of radial glia was significantly associated with the position of cells along PC3 (Fig. 1J). By clustering radial glia based on the 1% of genes most strongly loading PC3, we identified two transcriptionally distinct groups: one almost purely composed of cells from the VZ, which we interpreted as vRG cells, and another composed of cells from both the VZ and the subventricular zone (SVZ), which we interpreted as oRG cells (Fig. 1K).

**Predicted Markers Relate to Position, Morphology, and Behavior of oRG Cells**

To relate these distinct transcriptional states to the stem cell niches of the developing neocortex, we first searched for genes likely to distinguish predicted radial glia subtypes. We measured the specificity of genes by their correlation with an ideal marker gene uniformly expressed in only one putative radial glia subpopulation across all 393 cells. We identified 67 candidate marker genes strongly correlated with the oRG population, 33 candidate genes strongly correlated with the vRG population, and 31 genes strongly correlated with both radial glia populations (Fig. 1L, orange, green, and yellow boxes, respectively). In support of these
predictions, we observed that candidate vRG markers showed higher expression in the VZ, whereas candidate oRG markers showed higher expression in the SVZ across human cortical tissue samples profiled by the Allen Brain Institute using microarray technology (Fig. 1M) (Miller et al., 2014). In addition, the expression levels of predicted oRG and vRG markers were inversely correlated across radial glia cells (Fig. 1N).

To further investigate candidate marker genes, we performed in situ hybridization in primary tissue samples. We found that expression of vRG candidates CRYAB, PDGFD, TAGLN2,FBXO32, and PALLD was strongest in the VZ, while expression of oRG candidates HOPX, PTPRZ1, TNC, FAM107A, and MOXD1 was strongest in the OSVZ (Fig. 2A). We confirmed specificity by co-immunolabeling stained tissue samples with antibodies against the classical radial glia marker SOX2 (SRY [sex-determining region-Y]-box2) and intermediate progenitor marker EOMES (eomesodermin) (Fig. 2B). In contrast to the radial-glia-specific expression of these transcripts, expression of the predicted novel marker of IPCs, PPP1R17, correlated strongly with classical marker EOMES (Fig. 2C). Immunostaining for HOPX (HOP homeobox), PTPRZ1 (protein phosphatase zeta-1), and TNC (tenasin C) proteins revealed their expression in cells with basal fibers that lacked EOMES expression, linking this molecular identity to the typical morphology of oRG cells (Fig. 2D).

To next relate this molecular identity to distinctive oRG behaviors, we performed time-lapse imaging of organotypic cortical slices (between GW15 and GW19.5) infected with green fluorescent protein (GFP)–expressing adenovirus, and then examined the expression of the most specific oRG marker, HOPX (representative example shown in Fig. 2E). We observed that cells undergoing mitotic somal translocation behavior of oRG cells can generate SOX2/HOPX double-positive daughter cells with long basal processes characteristic throughout neurogenesis. Together, these results link the molecular identity determined from scRNA-seq to the anatomical location, morphology, and behavior of oRG cells.

Beyond simply marking oRG cells, the genes we identified belong to common pathways that suggest mechanisms by which human oRG cells actively maintain the OSVZ as a neural stem cell niche (Fig. 2F). Many of these genes promote growth factor signaling, including TNC, PTPRZ1, ITGB5, SDC3, HS6ST1, IL6ST, and LIFR (Wiese et al., 2012). For example, TNC potentiates fibroblast growth factor (FGF) signaling to support the maturation of neural stem cells (Garcion et al., 2004), whereas integrin signaling along the basal fiber promotes radial glia identity (Fieltz et al., 2010). Interestingly, TNC contains epidermal growth factor (EGF)–like repeats and multiple binding domains for PTPRZ1, syndecans, integrins, and other cell-surface receptors (von Holst, 2008). Thus, TNC expression in oRG cells is able to couple key protein networks regulating growth factor signaling, migration, and self-renewal in the massively expanded human OSVZ (Fig. 2G). In addition, LIFR/STAT3 signaling is known to maintain radial glia neural stem cell identity (Bonaguidi et al., 2005), and we found that p-Y705-STAT3 signaling is necessary for normal cell-cycle progression in oRG cells but is surprisingly absent in vRG cells. Finally, we directly examined the neural stem cell properties of oRG cells using single-cell clonal lineage analysis. We found that single oRG cells could generate clones of nearly 1000 daughter cells of neuron and glial cell types, highlighting the remarkable proliferative capacity of human oRG cells compared with mouse radial glia, which typically generate only 10–100 daughter cells throughout the neurogenic period (Qian et al., 2000; Gao et al., 2014).

Conclusion

Our study identified neuronal differentiation, cell-cycle progression, and anatomical position as major sources of transcriptional variation across single cells sampled from germinal niches of the developing human cortex. Using in situ hybridization and immunostaining, we connected gene expression signatures predicted by scRNA-seq to the position, morphology, and dynamic behavior of cells in tissue. Together, this multimodal characterization establishes an integrative identity for oRG cells. These neural stem cells are characterized by the expression of markers that also appear in astrocytes, but not in vRG cells, including HOPX, TNC, and ITGB5, as well as pan-radial glia markers such as VIM, HES1, and ATP1A2; the presence of a basal, but not apical fiber; mitotic–somal translocation behavior; and extensive proliferative and neurogenic capacity. The oRG subtype is most abundant in the OSVZ stem cell niche, for which it was named, but also resides in the inner SVZ and VZ, and the transcriptional state first emerges in the VZ during early cortical neurogenesis. The oRG marker genes potentially enable the construction of molecular tools for selectively visualizing, manipulating, or purifying oRG cells in tissue and for evaluating the identity of human cortical progenitor cells generated from pluripotent stem cells (Qian et al., 2016).
addition, these genes may provide insights into the cell types affected in neurodevelopmental disorders and infectious disease (Nowakowski et al., 2016).

Sequencing single-cell mRNA while retaining cell position information provides a general method for identifying distinct subpopulations whose molecular identity possibly relates to microenvironment and functional properties. Here, we explored variation in radial glia gene expression while considering stem cell niche as a covariate. Our results reveal novel molecular features of neural stem cell populations previously distinguished only by cell behavior, morphology, and position. Together with recent findings (Fietz et al., 2012; Lui et al., 2014), these results highlight three mechanisms that may maintain the “stemness” of the expanded oRG population in the OSVZ stem cell niche: local production of trophic factors such as PTN (pleiotrophin) and BMP7 (bone morphogenetic protein-7) by radial glia, expression of extracellular matrix proteins that potentiate growth factor signaling, and activation of the LIFR/p-STAT3 signaling pathway (Figs. 2G, H). Because the oRG population is thought to be responsible for the majority of human cortical neurogenesis, and OSVZ size correlates with the evolutionary expansion of
the brain, future studies could investigate the role of these genes in neurodevelopmental disorders and cortical evolution.

Future Directions
Cell capture is no longer the bottleneck for surveying cell diversity in heterogeneous tissue. New technologies allow massively parallel single-cell capture (Klein et al., 2015; Macosko et al., 2015), and the low dimensionality of gene expression data permits exploration of cells in heterogeneous tissue at extremely low sequencing depths (Jaitin et al., 2014; Pollen et al., 2014; Heimberg et al., 2016). Meanwhile, advances in data analysis and interpretation have resulted in new methods for clustering cell types, thereby predicting patterns of sequential gene expression during signaling pathway activation and lineage progression (Trapnell, 2015). In addition, new techniques have enabled analysis of additional molecular features, including genome sequence and chromatin state (Buenrostro et al., 2015).

One promising area of future work is to use sequencing to also survey cellular phenotypes in a similarly high-throughput manner. In our study, we combined gene expression data with separate low-throughput studies of spatial position, morphology, cell behavior, and developmental lineage potential. Ideally, we could measure these cellular phenotypes and community properties alongside gene expression. In situ sequencing (Lee et al., 2014) and other techniques for spatial transcriptomics and proteomics have the potential to help capture the spatial enrichments and neighborhoods of cell types. Similarly, studies using barcoded viruses that cross synapses may be able to label the connections between cell types (Pollock et al., 2014; Kebschull et al., 2016). Finally, the activity, behavior, and lineage relationships between cells may be recorded either through direct imaging of isolated cells prior to capture or through genome modification in response to processes such as cell division or activity (McKenna et al., 2016; Shipman et al., 2016). Together, these emerging technologies may someday enable integrated high-throughput analysis of single-cell gene expression patterns combined with single-cell phenotypes, as well as the connections and interactions between cells in nervous system tissue.

Acknowledgments
This chapter was adapted with permission from Pollen AA et al. (2015) Molecular identity of human outer radial glia during cortical development, Cell 163:55–67. Copyright 2015, Elsevier.

References


