Title
Fate mapping of human cerebral cortex progenitors reveals novel patterns of neuronal and glial differentiation

Permalink
https://escholarship.org/uc/item/09s802r3

Author
Allen, Denise

Publication Date
2022

Peer reviewed|Thesis/dissertation

Fate mapping of human cerebral cortex progenitors reveals novel patterns of neuronal and glial differentiation

by
Denise Elizabeth Allen

DISSERTATION
Submitted in partial satisfaction of the requirements for degree of
DOCTOR OF PHILOSOPHY

in

Biomedical Sciences

in the

GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Approved:

DocuSigned by:

Erik Ullian

1ACF3D77A55F4F0...

Erik Ullian
Chair

DocuSigned by:

Arturo Alvarez—Buylla

DocuSigned by: ...492...

Arturo Alvarez-Buylla

DocuSigned by:

Tomasz Nowakowski

CAFC10E825A34A4...

Tomasz Nowakowski

Committee Members

This work is dedicated to Ms. Chris Read, Dr. Li Gan, Dr. Grietje Krabbe, Dr. Ellen Carpenter, Dr. Elvira Khialeeva, Dr. Bill Lowry, and Dr. Minori Ohashi. Without your trust, training, and support I would never have made it to day one of a PhD.

I also dedicate this work to my parents. Thank you for your steadfast and all-encompassing support of my dreams.

ACKNOWLEDGEMENTS

I would first like to thank my PhD advisor Dr. Tomasz Nowakowski for his unwavering support and encouragement throughout my entire PhD journey. When I first interviewed with Tom I was struck by his overwhelming and contagious passion for all things neurodevelopment. That enthusiasm has not wavered one moment during my time in the lab and has at points been critical for buoying me through tough times. Even more importantly, I feel so lucky to have never once doubted that Tom wanted the best for me. While we sometimes butted heads as you always do, I never once doubted that he wanted me to follow my passions and reach my fullest potential. I am walking away from this lab deeply inspired by his passion and deeply thankful for his support.

I would also like to thank all the members of the Nowakowski lab, past and present. I interviewed for the lab when it was still half-unpacked and only three people strong, and am now leaving behind a bustling lab of ten incredible scientists plus four alumni. They tell you not to pick a lab based on the people because the people will change, but I am so lucky that in the Nowakowski lab good people are a constant. A PhD is at times elating and at other times soul-crushing, and I could not have asked for a more supportive group of people to share this PhD journey with. I would also like to thank several members in particular who have had the largest impact on my experience. David and I started in the Nowakowski lab one week apart, and I never knew how fantastic it could be to share every step of a PhD with someone else (#DADS forever). Having someone to commiserate and celebrate with as we went through the same phases at the same time was something was one of the biggest blessings of my PhD. To Matt, thank you for being a fantastic roton, the best partner I could have asked for to finish up the *Nature* paper with, and for being the person I knew I could always 100% count on for both scientific and emotional support. Finally, thank you to Ryan Delgado for teaching me so much about science and about life. I was so inspired by your unwavering passion, positivity, and confidence, and honored to be trusted with working on and ultimately finishing the project that

iv

became the *Nature* paper. Most importantly, thank you for teaching me how to stay laser focused on the project at hand. Without that inspiration I'm not sure I would have graduated as timely as I did.

I would also like to thank the members of my thesis committee, Arturo Alvarez-Buylla and Erik Ullian. Thank you for sharing so much of your time and advice over the last four years-- I left every meeting we had feeling more inspired and more motivated, with a whole new perspective on how exciting my work really was. Thank you also to the members of the Kreigstein lab, especially Madeline Andrews and Aparna Bhaduri, for helping bring me up to speed on the world of single cell genomics. Thank you to Kevin Yamauchi, Andrew Cote, and Matt Cai at the Chan Zuckerberg Biohub for welcoming me into the world of spatial transcriptomics. Thank you to Kevin Donohue for providing the missing piece to the *Science* paper and for being such a fantastic collaborator. Finally, thank you to Ed Lein, Trygve Bakken, Rebecca Hodge, Johnathan Ting, and Nik Jorstad at the Allen Institute for your help with my adventures in astrocyte single cell RNA-sequencing analysis.

Thank you also to all of the friends who have supported me during my time at UCSF. Thank you to the "BMS Wine Night" group Camillia, Lindsey, Tara, Casey, and the occasional significant other for providing such wonderful friendship. Thank you for being there to complain, celebrate, and compare notes about lab, and most importantly for providing a safe haven for being a woman in science. Thank you also to the members of the 2017 DSCB program for welcoming me into your fantastic and tightknit group. Thank you to Karishma, Nicole, Antara, and Maxine for spicing up my life with some DSCB fun. Finally, thank you to my boyfriend Edward. We met just three months before the pandemic, and I am not sure I would have made it through graduate school during a pandemic without you. Thank you for not only providing moral support, emotional support, and hugs, but also for providing the incredibly important reminder that life continues outside of lab. Thank you for pushing me to build a vibrant life outside of science and to enjoy everything else that the world has to offer.

independence that enabled me to enter a PhD program directly after my time at UCLA. Thank you to these mentors and the many, many other informal mentors who were so generous with their time to train me into a scientist who could succeed at UCSF.

Finally, thank you to my family. I often joke that I am the "black sheep" of the family for not pursuing an engineering degree like my father and sister, but I have never actually been made to feel one. I am so incredibly lucky to have parents who have provided such unwavering support for my entire life. Thank you to my parents for valuing my education above all else and never letting me doubt that you were proud of me. Thank you in particular to my mom for supporting my earliest research career--driving me 45 minutes to the city and back every day, often spending the day in the car just waiting for me to finish. Thank you for supporting me through so many unpaid research hours, doing everything in your power to make sure that I never had to worry about anything but following my dream. Thank you also to my uncle Mark Schar and aunt Elizabeth Schar for being such important role models in my life. You have been one of my biggest inspirations that I can do anything I want at any time, with just a little bit of passion and hard work.

A PhD can be a lonely pursuit. You can be increasingly siloed into your own project and own domain until it sometimes feels like you are a lone soldier on the battle field. But this acknowledgement section highlights that there is no way to get through a PhD alone. It is the people who have supported me from my childhood to now that are the reason this thesis exists today. Thank you to the entire village to loved me, supported me, and trained me to lead up to this moment.

CONTRIBUTIONS

All work described in this dissertation was carried out under the direct supervision of Dr. Tomasz Nowakowski. Chapter 2 is adapted from a research article published in *Science*, of which Denise is the sole first author. Chapter 3 is adapted from a research article published in *Nature*, for which Denise is the second of three co-first authors. Additional author contributions for each chapter are described in the Contributions section of each chapter.

Citation for Chapter 2 manuscript:

Allen, D.E., Donohue, K.C., Cadwell, C.R., Shin, D., Keefe, M.G., Sohal, V.S., Nowakowski, T.J. Fate mapping of neural stem cell niches reveals distinct origins of human cortical astrocytes. *Science.* First Release. (2022) DOI: 10.1126/science.abm5224

Citation for Chapter 3 manuscript:

Delgado, R.N.*, **Allen, D.E.*,** Keefe, M.G.*, Mancia, W.R., Ziffra, R.S., Crouch, E.E. Alvarez-Buylla, A., Nowakowski, T.J. Individual human cortical progenitors can produce excitatory and inhibitory neurons. *Nature* **601**, 397–403 (2022). https://doi.org/10.1038/s41586-021-04230-7

*These authors contributed equally to this work

Fate mapping of human cerebral cortex progenitors reveals novel patterns of neuronal and glial differentiation

Denise Elizabeth Allen

ABSTRACT

The human cerebral cortex is composed of a vast diversity of neuronal and glial cell types. Decades of work in animal models has informed a model of cortical development where neural stem cells that reside along the lateral ventricle first give rise to excitatory neurons followed by glial cell types including astrocytes and oligodendrocytes, while inhibitory cortical interneurons are born outside of the cortex and migrate in. However, it remains unclear to what degree human cortical development follows this canonical model. Recent work has demonstrated that during midgestation in humans this canonical population of ventricular radial glia differentiates into truncated radial glia that remain along the ventricle in the ventricular zone (VZ), and outer radial glia that form a second stem cell niche in the outer subventricular zone (OSVZ). In this thesis, I first use fate mapping of progenitors in the human VZ and OSVZ to demonstrate that these two stem cell niches give rise to spatially and morphologically distinct populations of astrocytes, and then use Patch-seq single cell analysis to identify unique molecular markers of these astrocyte subtypes. Second, I use a novel high throughput lineage tracing tool in combination with single cell RNA-sequencing to demonstrate that human cortical neural stem cells can give rise to both excitatory neurons and inhibitory cortical interneurons, and that these two cell types can be derived from the same individual stem cell. Together this work revolutionizes our understanding of both human cortical neurogenesis and gliogenesis and provides a foundation for understanding how these unique features of human cortical development may be perturbed in disease.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1: Introduction

Canonical model of cortical development

The human cerebral cortex is responsible for carrying out many of our higher order cognitive functions such as perception, language, and decision making. The cortex is composed of a vast diversity of cell types including excitatory and inhibitory neurons, as well as glial cells such as astrocytes, oligodendrocyte progenitor cells, oligodendrocytes, and microglia. Decades of studies in animal models have informed a canonical model of cortical development where a single layer of neuroepithelial cells differentiates into pluripotent neural stem cells--also known as radial glia--that give rise to the majority of cortical cell types. These radial glia divide asymmetrically to give rise directly to neurons as well as to neuronal intermediate progenitor cells, which together generate excitatory neurons of the deep cortical layers followed by those in the upper layers[1]. Once neurogenesis is complete, radial glia undergo a "gliogenic switch" as they transition to generating astrocytes and then oligodendrocyte lineages via asymmetric divisions and direction transformation[2,3]. Additionally, inhibitory cortical interneurons that are critical for maintaining neural circuit function are born ventrally in the ganglionic eminences and then migrate dorsally into the cortex[4,5], while microglia are born in the yolk sac and enter the central nervous system before formation of the blood brain barrier[6]. All of these cell types then undergo further maturation during the third trimester and early childhood to create the functioning adult cerebral cortex.

Human cortical development

This general paradigm of cortical development holds true in humans, with neuroepithelial cells differentiating into radial glia that give rise to excitatory neurons first in an inside out manner followed by gliogenesis. However, recent work has highlighted a few key differences

between human and rodent corticogenesis. First, the human cerebral cortex is greatly expanded in thickness compared to traditional rodent models, much of which is due to the expansion of the supragranular layers[7]. While it appears that the cellular makeup of the cortex is broadly conserved across evolution, the cell type proportions, gene expression, chromatin states, and marker genes can vary greatly[8]. Furthermore, while neurogenesis and gliogenesis are almost entirely non-overlapping in rodents, in the human neocortex gliogenesis begins while upper layer neurogenesis is still progressing, and the two processes take place concurrently for several weeks[9,10]. This suggests that there may be further unique features of how the human cortex is built.

One possible mechanism for enabling these dramatic differences in human cortex size and development is the expansion and diversification of the human radial glia population. Early human neurogenesis proceeds from a single population of radial glia that is similar to those seen in rodent models in that they reside along the lateral ventricle in the ventricular zone (VZ) and extend a basal process to the pial surface[11,12]. However, during mid-neurogenesis at approximately gestational week 16 (GW16) this single population of ventricular radial glia differentiates into CRYAB[+] truncated radial glia (tRG) that remain in the VZ but lose their pial contact, and HOPX[+] outer radial glia (oRG) that lose their ventricular contact and form a second stem cell niche in the outer subventricular zone (OSVZ)[13]. While a small population of HOPX[+] oRG are observed in the developing mouse cortex[14], this population appears to be greatly expanded in the human. Their highly proliferative nature and both neurogenic[15,16] and gliogenic[17–19] potential have led many to hypothesize that this expanded stem cell population is at least in part responsible for the expansion of the human cortex compared to other mammals, especially of the upper layers that are formed after the emergence of oRG[20,21]. Similarly, tRG have not been observed in mice but have been observed in the gyrencephalic ferret cortex, suggesting they could also play a role in this cortical expansion and folding[22]. Unlike oRG however, the neurogenic and/or gliogenic potential of tRG has not been extensively explored[13].

Together these findings of dramatic differences between corticogenesis in humans versus animal models emphasize the need for further direct study of human cortical development.

Cortical astrocyte diversity

Historically the study of cell type diversification in the developing neocortex has focused on neuronal development, with glial cells added on as a homogenous afterthought. Yet astrocyte diversity has been appreciated since the drawings of Ramon y Cajal. The three major classes of astrocytes currently recognized in the human cortex include protoplasmic, fibrous, and interlaminar astrocytes. Fibrous astrocytes reside in the white matter where they extend a dense arbor of long, thin processes along axon tracts to contact nodes of Ranvier. This subtype strongly expresses the canonical astrocyte marker GFAP along with ID3 and CD44. Protoplasmic astrocytes reside in the gray matter where they elaborate a dense arbor of more highly branched processes and play a critical role in regulating ion balancing, neurotransmitter reuptake, ensheathing synapses, and regulating the blood brain barrier. These cells weakly express GFAP if at all, and GFAP expression is primarily restricted to their primary processes. Instead, they are more reliably marked expression of AQP4 and ALDH1L1[23]. Interlaminar astrocytes have their cell body in layer 1 of the cortex and extend many short processes upward to form the glial limitans, as well as a 1-4 long processes down into the grey matter[24]. Layer 1 astrocytes exist across mammals, but the extent to which their processes penetrate the underlying cortex greatly varies. In mice these cells rarely exit layer 1, while in humans they can penetrate 500 μm down into the cortex[25,26]. While these cells have been described extensively in terms of their morphology and gene expression across mammals, their exact function remains unknown. In addition to playing different roles in the healthy brain, these three subtypes have also been reported to be differentially affected in Autism Spectrum Disorder[27], Schizophrenia[28,29], Bipolar Disorder[28,30], Down Syndrome[31,32], and Alzheimer's Disease[33].

3

However, this classical model of three major astrocyte subtypes may soon become outdated as our understanding of astrocyte diversity grows. Recent work has suggested variations in astrocyte morphology[34,35] and gene expression[36] across the cortical layers--perhaps unsurprising as this is type of various is well established for neurons[37]. These findings suggest that there are additional levels of astrocyte diversity that have not yet been uncovered.

Development of cortical astrocytes

At E17 in mouse[38] and around GW17 in humans[39], radial glia in the developing cortex undergo a "gliogenic switch" as they stop giving rise to excitatory neurons and switch to generating astrocytes via either asymmetric division or direct transformation[17]. However, it remains unclear what exact differentiation steps occur between the radial glia stage and mature differentiated astrocytes. During neuron and oligodendrocyte development, there is a molecularly and morphologically distinct intermediate cell type in the form of neuronal intermediate progenitor cells[40] or oligodendrocyte progenitor cells[19], respectively. However, whether there exists and analogous "astrocyte progenitor cell" is hotly debated. The idea of a dual astrocyte/oligodendrocyte progenitor has been considered since early descriptions of so-called O2A or NG2 cells that were able to give rise to both oligodendrocyte and astrocytes in culture[41]. However, reports are conflicting as to whether they are also able to give rise to both cell types *in vivo*[42] or only to oligodendrocytes[43]. Several recent papers have highlighted populations of EGFR[+] cells as candidates for this dual-potent progenitor based on expression of shared markers either by immunohistochemistry[18] or single cell RNA-sequencing[19,44,45], but direct lineage tracing has yet to be demonstrated.

Even less well understood is how either the canonical astrocyte diversity--protoplasmic, fibrous, and interlaminar astrocytes--or more recent descriptions of astrocyte diversity--unique morphologies or molecular profile across cortical layers--emerges. In the spinal cord, gradients of secreted bone morphogenic protein and sonic hedgehog establish distinct progenitor

domains marked by combinatorial expression of PAX6, NKX6.1, NKX2.2, OLIG2, and SCL transcription factors, which then give rise to spatially and molecularly distinct populations of grey and white matter astrocyte, indicating that astrocyte diversity could be determined at the level of the progenitor[46]. This deterministic model is further supported by studies demonstrating that deleting the key glial transcription factor OLIG2 preferentially depletes white matter astrocytes[47] while deleting SOX9 preferentially depletes grey matter astrocytes[48]. However, whether these transcription factors truly regulate a fate decision rather than changes in morphology or GFAP expression and any upstream partners regulating this process has not been determined. Conversely, multiple studies have demonstrated that astrocytes from different brain regions can strongly affect the behavior of adjacent neurons[49,50], but the degree to which astrocyte diversification is shaped by their environment remains to be determined.

Lineage tracing methods

The prospect of being able to track the output of single progenitor cells has been the holy grail of the neurodevelopment field for decades. Early studies that used sparse labeling with colorimetric readouts to mark individual progenitors and their progeny[51–55] built the foundation of our current understanding of cortical development, but were extremely low throughput and relied on the assumption that progeny cells did not migrate far from their origin. Methods such as Brainbow[56] and MADM[57] enabled higher throughput imaging-based tracing, but still lacked high resolution. However, the advent of revolutionary molecular biology tools such as next generation sequencing, CRISPR, and single cell analysis has opened a whole new world of lineage tracing approaches. These tools have allowed the creation of new lineage tracing approaches that use either unique static[58–60] or "evolvable" barcodes that edit themselves along the lineage trajectory[61–64] to map the lineage of progenitors at the single cell level and eventually at the level of individual cell divisions in a high throughput manner.

However, most of these tools require a system that can be highly genetically modified, as they often require significant genome editing to introduce large constructs encoding fluorophores or DNA barcode sequences. This makes these tools incompatible with systems such as primary human tissue, which cannot be engineered from birth and is highly sensitive to manipulation. Some have tried to circumvent this issue by introducing ways to perform lineage tracing using endogenous somatic[65] or mitochondrial[66] mutations, but so far these techniques have also lacked the throughput or resolution needed for detailed tracing. This therefore leaves a gap in the toolbox for those seeking to study the development of delicate systems such as the human brain.

In this thesis

Together, the following thesis builds upon our existing knowledge of cortical development gleaned from decades of work in animal models and tests whether these principles are maintained in humans. In Chapter 2 we perform fate mapping of progenitors in the human VZ and OSVZ to compare the cell type output of these two distinct stem cell niches. This fate mapping work demonstrates that both niches give rise to both neurons and glia, highlighting the continued role of the VZ in midgestational corticogenesis as opposed to all late cortigogenesis occurring in the OSVZ. However, this mapping also reveals that the VZ and OSVZ give rise to spatially, morphologically, and molecularly distinct populations of astrocytes. The VZ gives rise to astrocytes that localize to the prospective grey matter of the cortical plate and subplate and that adopt the unique morphology of a dense arbor of processes with bulbous swellings along their length. In contrast, the OSVZ gives rise to astrocytes that remained in the prospective white matter of the SVZ and extend a dense arbor of long, smooth processes. Single cell analysis of these two morphotypes using Patch-seq further revealed that these two cell types have unique molecular profiles, and that the VZ-derived "dense bulbous" astrocytes are marked by expression of ITGB4 while the OSVZ-derived dense smooth astrocytes are

6

marked by expression of ANGPTL4. Together this work reveals divergent origins of human cortical astrocytes and emphasizes the role of progenitor niche in the fate of their progeny in the human brain.

In Chapter 3, I and my co-authors develop and utilize a novel high throughput lineage tracing tool to map the clonal output of individual human radial glia. My co-author Ryan Delgado developed a novel lineage tracing tool called "single-cell-RNA-sequencing-compatible tracer for identifying clonal relationships" (STICR) that consists of a highly diverse library of DNA barcodes that can be delivered to individual progenitors via lentiviral infection, passed down to their progeny, and then read out via single cell RNA-sequencing to determine both cell type identity and shared clonal barcodes. After validating this tool, we use it to label more than one thousand primary human cortical progenitors to demonstrate the emergence of both excitatory neurons and inhibitory cortical interneurons from the same cortical radial glia cell both *in vitro* and after xenograft transplantation into the mouse cortex. This finding therefore revolutionizes our model of excitatory and inhibitory neuron generation in the human cortex.

References

1.   Noctor, S. C., Flint, A. C., Weissman, T. A., Dammerman, R. S. & Kriegstein, A. R.

      Neurons derived from radial glial cells establish radial units in neocortex. *Nature* **409**,

      714–720 (2001).

2.   Gressens, P., Richelme, C., Kadhim, H. J., Gadisseux, J. F. & Evrard, P. The germinative

      zone produces the most cortical astrocytes after neuronal migration in the developing

      mammalian brain. *Neonatology* **61**, 4–24 (1992).

3.   on, J. -P, Takahashi, T. & Caviness, V. S. Ontogeny of radial and other astroglial cells in

      murine cerebral cortex. *Glia* **4**, 138–148 (1991).

4.   Gorski, J. A. *et al.* Cortical excitatory neurons and glia, but not GABAergic neurons, are

      produced in the Emx1-expressing lineage. *J. Neurosci.* **22**, 6309–6314 (2002).

5.   Anderson, S. A., Eisenstat, D. D., Shi, L. & Rubenstein, J. L. R. Interneuron migration

      from basal forebrain to neocortex: dependence on Dlx genes. *Science* **278**, 474–476

      (1997).

6.   Ginhoux, F. *et al.* Fate mapping analysis reveals that adult microglia derive from primitive

      macrophages. *Science* **330**, 841–845 (2010).

7.   Hutsler, J. J., Lee, D. G. & Porter, K. K. Comparative analysis of cortical layering and

      supragranular layer enlargement in rodent carnivore and primate species. *Brain Res.*

      1052, 71–81 (2005).

8.   Bakken, T. E. *et al.* Comparative cellular analysis of motor cortex in human, marmoset

      and mouse. *Nature* **598**, 111–119 (2021).

9.   Levitt, P., Cooper, M. L. & Rakic, P. Coexistence of neuronal and glial precursor cells in

      the cerebral ventricular zone of the fetal monkey: an ultrastructural immunoperoxidase

      analysis. *J. Neurosci.* **1**, 27–39 (1981).

10.   Malik, S. *et al.* Neurogenesis continues in the third trimester of pregnancy and is

suppressed by premature birth. *J. Neurosci.* **33**, 411–23 (2013).

11.   Haubensak, W., Attardo, A., Denk, W. & Huttner, W. B. From The Cover: Neurons arise

in the basal neuroepithelium of the early mammalian telencephalon: A major site of

neurogenesis. *Proc. Natl. Acad. Sci.* **101**, 3196–3201 (2004).

12.   Noctor, S. C., Martinez-Cerdeño, V., Ivic, L. & Kriegstein, A. R. Cortical neurons arise in

symmetric and asymmetric division zones and migrate through specific phases. *Nat.*

Neurosci. **7**, 136–144 (2004).

13.   Nowakowski, T. J., Pollen, A. A., Sandoval-Espinosa, C. & Kriegstein, A. R.

Transformation of the Radial Glia Scaffold Demarcates Two Stages of Human Cerebral

Cortex Development. *Neuron* **91**, 1219–1227 (2016).

14.   Vaid, S. *et al.* A novel population of Hopx-dependent basal radial glial cells in the

developing mouse neocortex. *Development* **145**, dev169276 (2018).

15.   Hansen, D. V., Lui, J. H., Parker, P. R. L. & Kriegstein, A. R. Neurogenic radial glia in the

outer subventricular zone of human neocortex. *Nature* **464**, 554–561 (2010).

16.   Betizeau, M. *et al.* Precursor Diversity and Complexity of Lineage Relationships in the

Outer Subventricular Zone of the Primate. *Neuron* (2013).

doi:10.1016/j.neuron.2013.09.032

17.   DeAzevedo, L. C. *et al.* Cortical radial glial cells in human fetuses: Depth-correlated

transformation into astrocytes. *J. Neurobiol.* **55**, 288–298 (2003).

18.   Rash, B. G. *et al.* Gliogenesis in the outer subventricular zone promotes enlargement and

gyrification of the primate cerebrum. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 7089–7094

(2019).

19.   Huang, W. *et al.* Origins and Proliferative States of Human Oligodendrocyte Precursor

Cells. *Cell* **182**, 594-608.e11 (2020).

20.   Lukaszewicz, A. *et al.* G1 phase regulation, area-specific cell cycle control, and

cytoarchitectonics in the primate cortex. *Neuron* **47**, 353–364 (2005).

21. Smart, I. H. M., Dehay, C., Giroud, P., Berland, M. & Kennedy, H. 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 (2002).

22. Borrell, V. & Götz, M. Role of radial glial cells in cerebral cortex folding. *Curr. Opin. Neurobiol.* **27**, 39–46 (2014).

23. Sofroniew, M. V. & Vinters, H. V. Astrocytes: Biology and pathology. *Acta Neuropathol.* **119**, 7–35 (2010).

24. Colombo, J. A. The interlaminar glia: from serendipity to hypothesis. *Brain Structure and Function* **222**, 1109–1129 (2017).

25. Falcone, C. *et al.* Cortical Interlaminar Astrocytes Are Generated Prenatally, Mature Postnatally, and Express Unique Markers in Human and Nonhuman Primates. *Cereb. Cortex* **31**, 379–395 (2021).

26. Falcone, C. *et al.* Cortical interlaminar astrocytes across the therian mammal radiation. *J. Comp. Neurol.* **527**, 1654–1674 (2019).

27. Velmeshev, D. *et al.* Single-cell genomics identifies cell type-specific molecular changes in autism. *Science* **364**, 685–689 (2019).

28. Hercher, C., Chopra, V. & Beasley, C. L. Evidence for morphological alterations in prefrontal white matter glia in schizophrenia and bipolar disorder. *J. Psychiatry Neurosci.* **39**, 376–385 (2014).

29. Williams, M. *et al.* Fibrillary astrocytes are decreased in the subgenual cingulate in schizophrenia. *Eur. Arch. Psychiatry Clin. Neurosci.* **264**, 357–362 (2014).

30. Webster, M. J., O'Grady, J., Kleinman, J. E. & Weickert, C. S. Glial fibrillary acidic protein mRNA levels in the cingulate cortex of individuals with depression, bipolar disorder and schizophrenia. *Neuroscience* **133**, 453–461 (2005).

31. Colombo, J. A., Reisin, H. D., Jones, M. & Bentham, C. Development of interlaminar

astroglial processes in the cerebral cortex of control and Down's syndrome human cases. *Exp. Neurol.* **193**, 207–217 (2005).

32. Zdaniuk, G., Wierzba-Bobrowicz, T., Szpak, G. M. & Stępień, T. Astroglia disturbances during development of the central nervous system in fetuses with Down's syndrome. *Folia Neuropathol.* **49**, 109–114 (2011).

33. Colombo, J. A., Quinn, B. & Puissant, V. Disruption of astroglial interlaminar processes in Alzheimer's disease. *Brain Res. Bull.* **58**, 235–242 (2002).

34. Lanjakornsiripan, D. *et al.* Layer-specific morphological and molecular differences in neocortical astrocytes and their dependence on neuronal layers. *Nat. Commun.* **9**, (2018).

35. Oberheim, N. A. *et al.* Uniquely hominid features of adult human astrocytes. *J. Neurosci.* **29**, 3276–3287 (2009).

36. Bayraktar, O. A. *et al.* Astrocyte layers in the mammalian cerebral cortex revealed by a single-cell in situ transcriptomic map. *Nat. Neurosci.* **23**, 500–509 (2020).

37. Lodato, S. & Arlotta, P. Generating Neuronal Diversity in the Mammalian Cerebral Cortex. *Annu. Rev. Cell Dev. Biol.* **31**, 699–720 (2015).

38. Kriegstein, A. R. & Alvarez-Buylla, A. The Glial Nature of Embryonic and Adult Neural Stem Cells. *Annu. Rev. Neurosci.* **32**, 149–184 (2009).

39. Marín-padilla, M. Prenatal development of fibrous (white matter), protoplasmic (gray matter), and layer I astrocytes in the human cerebral cortex: A Golgi study. *J. Comp. Neurol.* **357**, 554–572 (1995).

40. Noctor, S. C., Martínez-Cerdeño, V. & Kriegstein, A. R. Contribution of intermediate progenitor cells to cortical histogenesis. *Arch. Neurol.* **64**, 639–642 (2007).

41. Raff, M. C., Miller, R. H. & Noble, M. A glial progenitor cell that develops in vitro into an astrocyte or an oligodendrocyte depending on culture medium. *Nature* **303**, 390–396 (1983).

42. Zhu, X., Bergles, D. E. & Nishiyama, A. NG2 cells generate both oligodendrocytes and gray matter astrocytes. *Development* **135**, 145–57 (2008).

43. Espinosa de los Monteros, A., Zhang, M. & De Vellis, J. O2A progenitor cells transplanted into the neonatal rat brain develop into oligodendrocytes but not astrocytes. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 50–54 (1993).

44. Li, X. *et al.* Decoding Cortical Glial Cell Development. *Neurosci. Bull.* **37**, 440–460 (2021).

45. Yang, L., Li, Z., Liu, G., Li, X. & Yang, Z. Developmental Origins of Human Cortical Oligodendrocytes and Astrocytes. *Neurosci. Bull.* **38**, 47–68 (2022).

46. Tsai, H. H. *et al.* Regional astrocyte allocation regulates CNS synaptogenesis and repair. *Science (80-.).* **337**, 358–362 (2012).

47. Cai, J. *et al.* A crucial role for Olig2 in white matter astrocyte development. *Development* **134**, 1887–1899 (2007).

48. Stolt, C. C. *et al.* The Sox9 transcription factor determines glial fate choice in the developing spinal cord. *Genes Dev.* **17**, 1677–89 (2003).

49. Denis-Donini, S., Glowinski, J. & Prochiantz, A. Glial heterogeneity may define the three-dimensional shape of mouse mesencephalic dopaminergic neurones. *Nature* **307**, 641–643 (1984).

50. Petit, A., Pierret, P., Vallée, A. & Doucet, G. Astrocytes from cerebral cortex or striatum attract adult host serotoninergic axons into intrastriatal ventral mesencephalic co-grafts. *J. Neurosci.* **21**, 7182–7193 (2001).

51. Price, J., Turner, D. & Cepko, C. Lineage analysis in the vertebrate nervous system by retrovirus-mediated gene transfer. *Proc. Natl. Acad. Sci. U. S. A.* **84**, 156–60 (1987).

52. Kornack, D. R. & Rakic, P. Radial and horizontal deployment of clonally related cells in the primate neocortex: Relationship to distinct mitotic lineages. *Neuron* **15**, 311–321 (1995).

53. Mione, M. C., Cavanagh, J. F. R., Harris, B. & Parnavelas, J. G. Cell Fate Specification

and Symmetrical/Asymmetrical Divisions in the Developing Cerebral Cortex. *J. Neurosci.* **17**, 2018–29 (1997).

54. Mcconnell, S. K. Fates of Visual Cortical Neurons in the Ferret After Isochronic and Heterochronic Transplantation. *J. Neurosci.* **8**, 945–974 (1988).

55. Cepko, C. L. *et al.* Studies of cortical development using retrovirus vectors. *Cold Spring Harb. Symp. Quant. Biol.* **55**, 265–278 (1990).

56. Livet, J. *et al.* Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. *Nature* **450**, 56–62 (2007).

57. Zong, H., Espinosa, J. S., Su, H. H., Muzumdar, M. D. & Luo, L. Mosaic analysis with double markers in mice. *Cell* **121**, 479–492 (2005).

58. Wagner, D. E. *et al.* Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. *Science (80-.).* **360**, (2018).

59. Biddy, B. A. *et al.* Single-cell mapping of lineage and identity in direct reprogramming. *Nature* **564**, 219–224 (2018).

60. Kong, W. *et al.* CellTagging: combinatorial indexing to simultaneously map lineage and identity at single-cell resolution. *Nat. Protoc.* **15**, 750–772 (2020).

61. Chan, M. M. *et al.* Molecular recording of mammalian embryogenesis. *Nature* **570**, 77–82 (2019).

62. Frieda, K. L. *et al.* Synthetic recording and in situ readout of lineage information in single cells. *Nature* **541**, 107–111 (2017).

63. McKenna, A. *et al.* Whole-organism lineage tracing by combinatorial and cumulative genome editing. *Science* **353**, (2016).

64. Shipman, S. L., Nivala, J., Macklis, J. D. & Church, G. M. Molecular recordings by directed CRISPR spacer acquisition. *Science* **353**, (2016).

65. Evrony, G. D. *et al.* Cell Lineage Analysis in Human Brain Using Endogenous Retroelements. *Neuron* **85**, 49–59 (2015).

66. Ludwig, L. S. *et al.* Lineage Tracing in Humans Enabled by Mitochondrial Mutations and

   Single-Cell Genomics. *Cell* **176**, 1325-1339.e22 (2019).

CHAPTER 2

At the time of submission of this dissertation, this article was published online in Science via First Release

Fate mapping of neural stem cell niches reveals distinct origins of human cortical astrocytes

Denise E Allen[1,2,3,4], Kevin C Donohue[2,5,6,7,8], Cathryn R Cadwell[9], David Shin[1,2,3,4], Matthew G Keefe[1,2,3,4], Vikaas S Sohal[2,7,8], Tomasz J Nowakowski[1,2,3,4,7*]

[1] Department of Anatomy, The University of California San Francisco, San Francisco, USA.

[2] Department of Psychiatry and Behavioral Sciences, The University of California San Francisco, San Francisco, USA.

[3] Department of Neurological Surgery, The University of California San Francisco, San Francisco, USA.

[4] Eli and Edythe Broad Center for Regeneration Medicine and Stem Cell Research, The University of California San Francisco, San Francisco, USA.

[5] School of Medicine, The University of California San Francisco, San Francisco, USA.

[6] Center for Integrative Neuroscience, The University of California San Francisco; San Francisco, USA.

[7] Weill Institute for Neurosciences, The University of California San Francisco; San Francisco, USA.

[8] Kavli Institute for Fundamental Neuroscience, The University of California San Francisco, San Francisco, USA.

[9] Department of Pathology, The University of California San Francisco, San Francisco, USA.

*Corresponding author. Email: tomasz.nowakowski@ucsf.edu

ABSTRACT

Progenitors of the developing human neocortex reside in the ventricular and outer subventricular zones (VZ and OSVZ, respectively). However, whether cells derived from these niches have similar developmental fates is unknown. By performing fate mapping in primary human tissue, we demonstrate that astrocytes derived from these niches populate anatomically distinct layers. Cortical plate astrocytes emerge from VZ progenitors and proliferate locally, while putative white matter astrocytes are morphologically heterogeneous and emerge from both VZ and OSVZ progenitors. Furthermore, via single-cell sequencing of morphologically defined astrocyte subtypes using Patch-seq, we identify molecular distinctions between VZ-derived cortical plate astrocytes and OSVZ- derived white matter astrocytes that persist into adulthood. Together, our study highlights a complex role for cell lineage in the diversification of human neocortical astrocytes.

BACKGROUND

Radial glia serve as the neural stem cells of the developing neocortex (*1*, *2*). During midgestation in humans, radial glia can be further classified as either CRYAB$^+$ truncated radial glia (tRG) that lose their pial contact and remain in the ventricular zone (VZ) (*3*), or HOPX$^+$ outer radial glia (oRG) that lose their ventricular contact and reside in the outer subventricular zone (OSVZ) alongside neuronal intermediate progenitor cells (IPCs) (*4–6*)(*7–10*). These VZ and OSVZ stem cell niches can also be further distinguished based on their extracellular matrix components (*7*, *11*) and signaling pathway activation (*11*). Prior studies have shown that progenitor cells residing in the OSVZ contribute to cortical neurogenesis (*9*, *11*, *12*), astrogliogenesis (*13*), and oligodendrocytes (*14*), but the cells generated by human VZ progenitors are less well-characterized (*3*). tRG cells emerge at the onset of astrogliogenesis (*3*), suggesting that they could also contribute to the cortical astrocyte pool. To determine

16

whether both VZ and OSVZ progenitors contribute cortical astrocytes, we performed fate mapping of the human VZ and OSVZ. Our study revealed that VZ and OSVZ progenitors give rise to spatially, morphologically, and molecularly distinct subpopulations of astrocytes, highlighting the complexity of developmental lineage relationships in the developing human brain.

RESULTS

Cellular output of midgestation germinal zones

To fate map the human VZ and OSVZ niches, we prepared organotypic slice cultures of primary human neocortex from gestational weeks (GW) 18-23 to capture the time point shortly after tRG emergence (*3*) and right at the onset of astrogliogenesis (*13*). We labeled these niches in paired slices from the same individual using local delivery of viral vectors expressing green fluorescent protein (GFP) under the CMV/chicken β-actin (CAG) promoter (Fig. 2.1A). Two days after labeling–the earliest time point at which the GFP can be detected–GFP$^+$ cells labeled by VZ infections were found in the SOX2^{dense}, CRYAB$^+$ VZ and adjacent inner subventricular zone, while cells labeled by OSVZ infections were located in the SOX2^{dense}, CRYAB$^-$ OSVZ with a few cells in the adjacent SOX2^{sparse}, CRYAB$^-$ subplate (SP) (Fig. 2.1B, Fig. S2.1A). To determine the identity of these initially labeled cells, we stained for several markers of progenitor cell subtypes (Fig. 2.1C-F, Fig. S2.1B-C). Across slices from three individuals, 76% (SD ±10%) of the cells labeled by VZ infection were SOX2$^+$ progenitors, and 10% (SD ±2%) were CRYAB$^+$ tRG (Fig. 2.1E). After OSVZ labeling, we found that an average of 55% (SD ±8%) of OSVZ infected GFP$^+$ cells were dividing or newly born cells that incorporated 5-bromo-2'-deoxyuridine (BrdU), and that 10% (SD ±4%) were EOMES$^+$ neuronal progenitor cells and 5% (±2%) were HOPX$^+$ oRGs (Fig. 2.1F). Together, this data demonstrates that our local infections are spatially specific and label progenitor cells.

Next, we compared the distributions of VZ- and OSVZ-derived cells across the VZ, SVZ, SP, and cortical plate (CP) (Fig. S2.2A) after 8-12 days in culture (Fig. 2.1G, Fig. S2.2C-C). Consistent with prior studies (*9*, *12*, *14*, *15*), both the VZ- and OSVZ gave rise to cells with neuronal and glial morphologies (Fig. 2.1H-K). However, across 4-5 individuals OSVZ-derived glia largely remained within the SVZ, while VZ-derived glia were found throughout the cortical wall (Fig. 2.1L). To confirm that this phenotype was not induced by the serum in our culture media, we repeated this experiment in serum-free conditions and found a similar distribution (Fig. S2.3). To confirm that this observation was not the result of non-specific labeling during the initial VZ infection, we labeled the VZ using an alternative method of microdissecting out the VZ, separately bathing it in virus, and then co-culturing it with the remaining tissue slice (Fig. S2.4A). After thirteen days in culture, the transplanted VZ gave rise to GFP$^+$ cells with both neuronal and glial morphologies (Fig. S2.4B-I), and the VZ-derived glial cell population was evenly distributed across all laminae (Fig. S2.4J). Together, our data suggest that the VZ and OSVZ give rise to distinct subpopulations of glial cells with distinct localizations within the developing cortex (Fig. 2.1N).

Differential distribution of astrocytes and oligodendrocytes

Next, we used co-immunostaining to determine whether this differential glial distribution could be attributed to molecularly defined astrocytes (SOX9$^+$/OLIG2$^-$), oligodendrocyte progenitor cells (OPCs, SOX9$^+$/OLIG2$^+$), and/or oligodendrocytes (SOX9$^-$/OLIG2$^+$) (Fig. 2.2A). We found that all three glial cell types derived from the same niche shared comparable distribution patterns (Fig. 2.2B), indicating that this differential glial output is a broad feature of midgestation gliogenesis. These distribution patterns were also different from those of interneurons (Fig. S2.5C), excitatory neurons (Fig. S2.5C), or all cells (Fig. S2.5E). However, we also found a unique enrichment of astrocytes in the CP of VZ-labeled slices compared to OSVZ-labeled that did not apply to OPCs or oligodendrocytes (Fig. 2.2C, Fig. S2.5D). Thus, while the differential glial

output of VZ and OSVZ niches extends to all macroglial cell types, there is a unique enrichment of VZ-derived astrocytes in the CP (Fig. 2.2D).

To further confirm that the lack of OSVZ-derived glial migration was not due to a broader disruption of normal radial migration in our slice culture model, we determined the distribution of neuronal subtypes by staining for the glutamatergic neuron markers NEUROD2 and TBR1 and the GABAergic neuron marker DLX2 (Fig. S2.5A). We found that DLX2$^+$ cortical interneurons composed a relatively small proportion of GFP$^+$ cells in both VZ and OSVZ infections (Fig. S2.5B) but that they successfully migrated radially from their niche of origin (Fig. S2.5C). Culture in the presence of BrdU revealed BrdU$^+$ GFP$^+$/DLX2$^+$ interneurons derived from both VZ and OSVZ labeling (Fig. S2.5F), consistent with their dual origin in cortical and subcortical germinal zones in humans (*16*, *17*). By contrast, the majority of GFP$^+$ cells derived from VZ or OSVZ labeling were NEUROD2$^+$ glutamatergic neurons (Fig. S2.5B) that migrated radially outward from their niche of origin (Fig. S2.5C), and included both BrdU$^-$ and newborn BrdU$^+$ GFP$^+$/TBR1$^+$ glutamatergic neurons (Fig. S2.5G). Taken together, these findings demonstrates that extensive neuronal migration in our slice cultures are consistent with the expected distributions, indicating that the slice culture model can support normal patterns of cellular migration.

Morphological analysis of glial cell types

Radial glia and astrocytes undergo morphological transitions across their maturation trajectory, and morphology is a critical feature used to identify astrocyte subtypes in the absence of robust molecular markers (*18*). To determine whether VZ and OSVZ derived glia are morphologically heterogeneous, we analyzed the morphology of 524 randomly selected cells with glial morphology across four replicates of paired VZ- and OSVZ-labeled slices from GW20-GW23. We classified these cells into 'morphotypes' based on three previously described criteria: 1) the length of their primary processes, 2) the density of their primary processes, and 3) the types of

varicosities along their processes (Table S1). This morphometric analysis resulted in the identification of 12 glial morphotypes (Fig. 3.3A, Fig. S2.6) that are consistent with previous morphology annotations at similar ages (*18–20*).

We observed expected radial glial morphotypes including radial glia with a single apical process (*21*) (Fig. 3.3A, Fig. S2.6: "radial glia-no radial processes") that were derived from and resided in the VZ of younger samples (Fig. S2.7A-C), as well as radial glia at various stages of transforming into astrocytes (Fig. 3.3A, Fig. S2.6: "radial glia-short radial processes" and "radial glia-long radial processes") and oRG-like cells with both apical and basal fibers (Fig. 3.3A, Fig. S2.6: "radial glia-two long processes") derived from and residing in the SVZ of older samples (Fig. S2.7A-C) (*12*). We also observed SOX9$^+$/OLIG2$^+$/PDGFRA$^+$ OPCs with few, smooth processes (Fig. 3.3A, Fig. S2.6, Fig. S2.8A: "sparse smooth processes") (*22*), and SOX9$^-$/OLIG2$^+$/SOX10$^+$ immature oligodendrocytes with many short, branched processes (Fig. 3.3A, Fig. S2.6, Fig. S2.8B: "bushy processes") (*22*). A subset of cells with "bushy processes" were also beginning to express the mature oligodendrocyte marker MBP (Fig. S2.8C). These putative oligodendrocyte lineage cells were derived from both the VZ and OSVZ and found across the cortical wall of all ages studied (Fig. S2.7A-C).

Lineage divergence of morphologically distinct astrocytes

The remaining morphotypes that did not possess a primary fiber and did not express OLIG2 represent our putative astrocyte subtypes. These morphotypes included several categories with a sparser arbor of primary processes: those with "short processes", those with "sparse beady processes" defined by long thin processes with small regularly-sized and spaced swellings, and those with "sparse processes with irregular varicosities" defined by larger and more irregularly-sized and spaced varicosities (Fig 3A, Fig S6). These morphotypes were derived from both VZ and OSVZ labeling and found in multiple layers of the cortical wall (Fig. S2.7B-C).

Cells with a dense arbor of primary processes fell into three subcategories: "dense bulbous processes" with large bulbous varicosities along their processses, "dense smooth processes" with few varicosities, and "dense beady processes" with regularly spaced beady varicosities (Fig. 2.3A, Fig. S2.6). The "dense bulbous processes" and "dense smooth processes" morphotypes were unique in that they demonstrated a bias towards a single niche of origin and laminar localization. Across four individuals, 82% (SD ±13%) of astrocytes with "dense smooth processes" were derived from OSVZ labeling (Fig. 2.3B) and 78% (SD ±10%) remained in the SVZ (Fig. 2.3C) where they represented the dominant astrocyte morphotype (Fig. 2.3D, Fig. S2.7D). Gamma retrovirus labeling (Fig. S2.10) and BrdU incorporation experiments (Fig. 2.3E) demonstrated that the majority (70%, SD±13%) of these cells were born during the culture period (Fig. S2.7E), indicating that their lack of migration out of the SVZ was not due to labeling of mature cells. By contrast, 100% of "dense bulbous" astrocytes were derived from VZ labeling and BrdU incorporation confirmed they were born in this niche (Fig. 2.3B). The majority of these cells (70%, SD ±12%) were located in the CP (Fig. 2.3C) where they also composed the dominant morphotype (Fig. 2.3D), in addition to 29% (SD ±12%) in the SP (Fig. 2.3C). A subset of cells of this morphotype expressed AQP4 consistent with their astrocytic identity, and a subset were positive for Ki67 (Fig. S2.9), indicating that they could act as a locally dividing progenitor population (23). Finally, we confirmed that neither of these morphotypes were induced by our culture conditions by replicating our findings in serum-free media (Fig. S2.3). Together, this analysis demonstrates that during midgestation the VZ- and OSVZ give rise to astrocytes with distinct morphologies and laminar positions within the developing cerebral cortex (Fig. 2.3F).

Molecular characterization of astrocyte morphotypes

Finally, we sought to determine the molecular distinctions between VZ-derived "dense bulbous" and OSVZ-derived "dense smooth" astrocytes. To make this comparison we used a modified

version of the recently published Patch-seq protocol (22) to use a micropipette to collect the mRNA from morphologically defined cells followed by whole transcriptome single cell sequencing (Fig 4A, Fig. S2.11A). Across 10 individuals, we generated Patch-seq data for 28 "dense bulbous" astrocytes, 20 "dense smooth" astrocytes, and 22 neurons to serve as an outgroup. After quality control filtering (see methods) (Fig. S2.11B-D), we refined our dataset down to 12 dense bulbous astrocytes, 13 dense smooth astrocytes, 3 dividing astrocytes, 6 excitatory neurons, and 8 inhibitory neurons of high quality (Fig. 2.4B). We then performed differential gene expression analysis to identify 445 genes differentially enriched in "dense bulbous" astrocytes or "dense smooth" astrocytes (Fig. 2.4C, Fig. S2.11E). To identify the genes that most likely mark these cell types across their lifespan, we cross referenced our differentially expressed gene list with their expression in published single cell RNA-sequencing from either the developing (*24*) or adult (*25*) human neocortex (Fig. S2.11F-G). This analysis highlighted ITGB4 and ANGPTL4 as candidate markers of dense bulbous and dense smooth astrocytes, respectively, that were also enriched in astrocytes in both prenatal and adult cortex (Fig. 2.4D). Finally, we confirmed this prediction by performing *in situ* hybridization on cultured VZ- or OSVZ-labeled slices. This work demonstrated that *ITGB4* mRNA indeed localizes to VZ-derived dense bulbous astrocytes in the CP, and *ANGPTL4* mRNA localizes to OSVZ-derived dense smooth astrocytes in the OSVZ (Fig. 2.4E, Fig. S2.12A-C). Together, these experiments demonstrate that the two neural stem cell niches of the developing human cerebral cortex give rise to spatially, morphologically, and molecularly unique astrocyte subtypes (Fig. 2.4H).

DISCUSSION

Our study illuminates early developmental events that underlie the emergence of human neocortical astrocyte diversity. We find that during midgestation, OSVZ progenitor cells give rise to white matter astrocytes, while VZ progenitors give rise to more superficial gray matter astrocytes. This differentiation pattern differs from neurogenesis, where OSVZ progenitors give

22

rise to neurons destined for the upper cortical layers. Because neurogenesis and gliogenesis overlap extensively in humans (*26*), unlike in rodents (*27*), further studies will be needed to determine if individual human progenitors can give rise to both neuronal and glial cells during this time period or follow a classical neurogenic to gliogenic 'switch', and whether these patterns of differentiation are maintained throughout development. Our study also implicates truncated radial glia (*3*), which have not been described in mice, in the generation of a distinct subtype of cortical "dense bulbous" astrocytes that have been previously observed in humans (*20, 28*), but not in rodents. Additional studies are needed to determine the developmental origin of other astrocyte subtypes thought to be enriched in humans, such as varicose projection and interlaminar astrocytes (*29, 30*).

Astrocytes have been increasingly implicated in neurological disease (*31, 32*), with distinct astrocyte subtypes being implicated in schizophrenia, bipolar disorder (*33*), and autism spectrum disorders (*25*). Molecular distinctions between white matter and gray matter astrocytes identified in this study may enable insights into their unique developmental features and disease implications. For example, many genes enriched in "dense bulbous" astrocytes have been implicated in glioblastoma cell proliferation and invasion, such as *ITGB4 (34)*, *TMEM15*8 (*35*), *MGMT (36)*, and *CELSR1 (37)*, which could explain their remarkable migratory behavior. Detailed mapping of the developmental trajectories of these early astrocyte subtypes through the third trimester and into adulthood will be important to determine whether this astrocyte diversity is encoded at the level of progenitors or dynamically influenced by environmental cues present in distinct brain regions or cellular neighborhoods. In turn, these studies may reveal important subtype-specific vulnerabilities to environmental or genetic insults that underlie this diverse role for astrocytes in neurological disease.

FIGURES
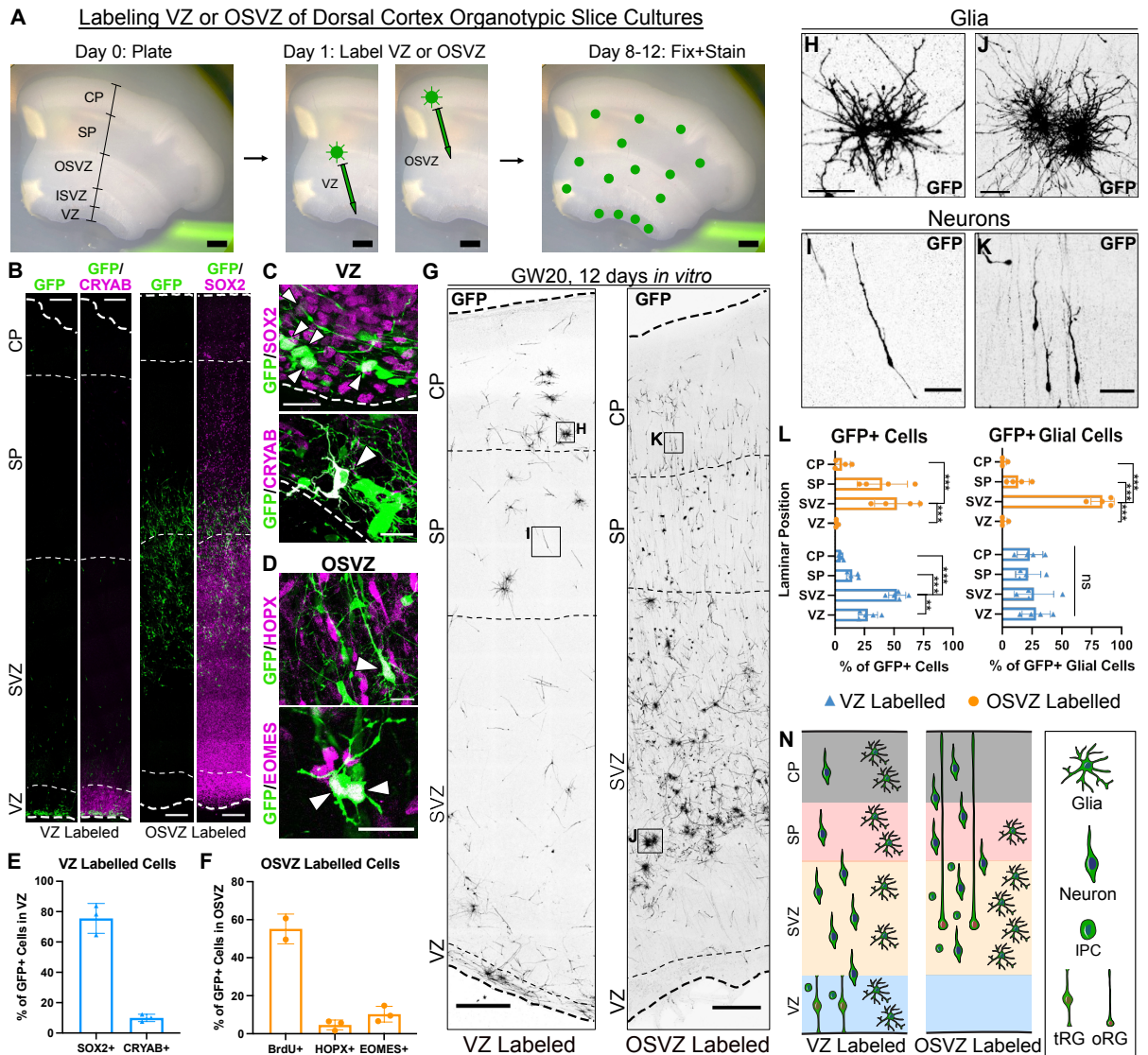
Figure 2.1: VZ and OSVZ niches both give rise to neurons but contribute distinct populations of glia. (**A**) Schematic of experimental design. VZ and OSVZ are visualized on a dissection microscope, virus is applied to the VZ or OSVZ using a glass needle, and after 8-12 days labeled GFP$^+$ cells (green dots) have migrated throughout the slice. Scale bar = 1mm. (**B**) GW19 sample 2 days after VZ or OSVZ labeling co-immunostained for GFP and CRYAB or SOX2. Thick dashed lines indicate ventricular and pial edges, thin dashed lines indicate borders between laminae. Scale bars = 250μm. (**C**) GFP$^+$/SOX2$^+$ or GFP$^+$/CRYAB$^+$ cells in the VZ 2 days after VZ labeling. White arrowheads indicate double-positive cells. Scale bar = 25μm. **D)** GFP$^+$/HOPX$^+$ or GFP$^+$/EOMES$^+$ cells in the OSVZ 2 days after OSVZ labeling. White arrowheads indicate double-positive cells. Scale bar = 25μm. **E)** Quantification (mean±SD) of the percentage of GFP$^+$ cells in the VZ that are SOX2$^+$ or CRYAB$^+$ 2 days after VZ labeling. N=3 independent samples. **F)** Quantification (mean±SD) of the percentage of GFP$^+$ cells in the OSVZ that are BrdU$^+$, HOPX$^+$, or EOMES$^+$ 2 days after OSVZ labeling. N=2-3 independent samples. **G)** GFP immunostaining of GW20 sample 12 days after VZ or OSVZ labeling. Dashed lines as in (B). Scale bar = 500um. (**H-K**) Insets from (G) showing cells with glial morphology (H,

J) and neuronal morphology (I, K). Scale bars = 50µm. (**L**) Quantification (mean±SD) of the percentage of total VZ- or OSVZ-derived GFP$^+$ cells or GFP$^+$ cells with glial morphology located in each lamina of the developing cortex at the end of culture (n=4-5 independent samples). Laminar borders were determined using DAPI density (Fig. S2.2A). Statistical significance assessed with one-way analysis of variance (ANOVA) with Tukey's multiple comparisons correction. *: q < 0.033, **: q < 0.002, ***: q < 0.001 ns: not significant. (**N**) Schematic summarizing distribution of VZ- or OSVZ-derived cells.
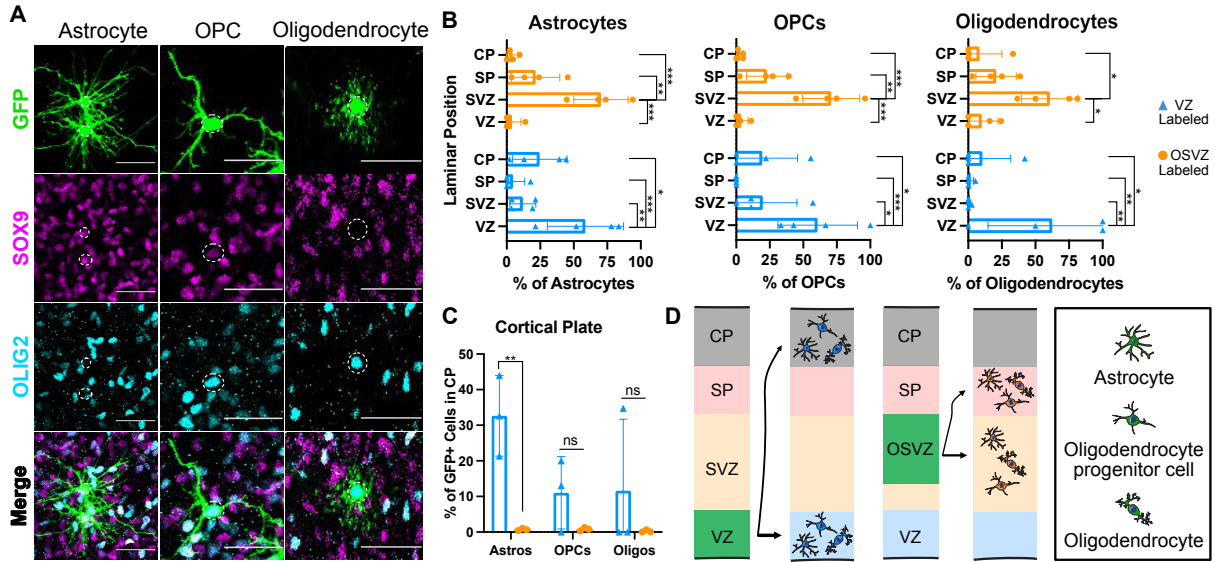
Figure 2.2: VZ and OSVZ give rise to spatially distinct populations of astrocytes, OPCs, and oligodendrocytes. (**A**) Co-immunostaining showing representative GFP$^+$/SOX9$^+$/OLIG2$^-$ astrocytes, GFP$^+$/SOX9$^+$/OLIG2$^+$ OPC, and GFP$^+$/SOX9$^-$/OLIG2$^+$ oligodendrocyte. All scale bars = 50µm. (**B**) Quantification (mean±SD) of the percentage of VZ- or OSVZ-derived astrocytes, OPCs, or oligodendrocytes that were located in the VZ, SVZ, SP, or CP at the end of culture. N=4 independent samples for all. (**C**) Quantification (mean±SD) of the percentage of total GFP$^+$ cells in the CP of VZ- or OSVZ-labeled slices that were astrocytes (Astros), OPCs, or oligodendrocytes (Oligos). N=3 independent samples. All statistical significance assessed with two- way ANOVA with Tukey's multiple comparison test correction. *: q < 0.033, **: q < 0.002, ***: q < 0.001 ns: not significant. (**D**) Schematic summarizing the distinct distribution of VZ versus OSVZ-derived astrocytes, OPCs, and oligodendrocytes. Green color indicates the niche of origin. Arrow line weight corresponds to relative abundance in the target layer.
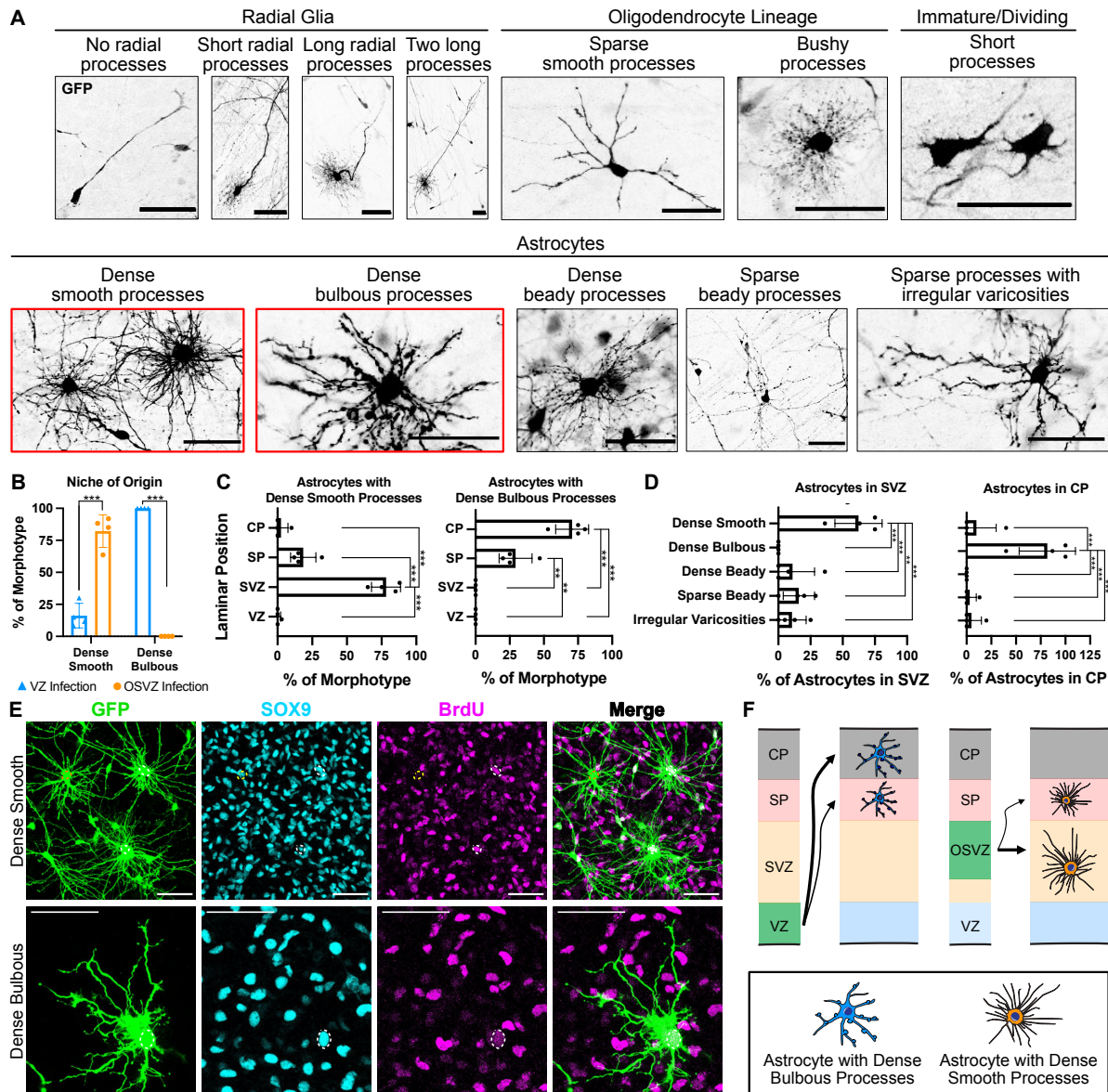
Figure 2.3: VZ and OSVZ give rise to morphologically distinct astrocyte subtypes. (**A**) GFP immunostaining showing representative images of the twelve glial morphotypes identified in this study. Scale bars = 50µm. Red boxes outline morphotypes analyzed later in the figure. (**B**) Quantification (mean±SD) of the percentage of astrocytes with dense smooth or dense bulbous processes derived from VZ or OSVZ labeling. (**C**) Quantification (mean±SD) of the percentage of astrocytes with dense smooth or dense bulbous processes located in each lamina at the end of culture. (**D**) Quantification (mean±SD) of the percentage of GFP+ astrocytes within the SVZ or CP that adopted each of our five identified astrocyte morphologies. Statistical significance assessed with one or two-way ANOVA with Tukey's multiple comparison correction. *: q < 0.033, **: q < 0.002, ***: q < 0.001 ns: not significant. N=4 independent pairs of VZ and OSVZ infections for (B-D). (**E**) Representative images of GFP+/SOX9+ "dense smooth" or "dense bulbous" astrocytes that are BrdU+ (white dashed outline) or BrdU- (yellow dashed outline). Scale bar = 50µm. (**F**) Schematic summarizing the differential contribution of the VZ and OSVZ to astrocyte morphotypes. Arrow line weight corresponds to relative contribution to the target layer of the cortical wall.
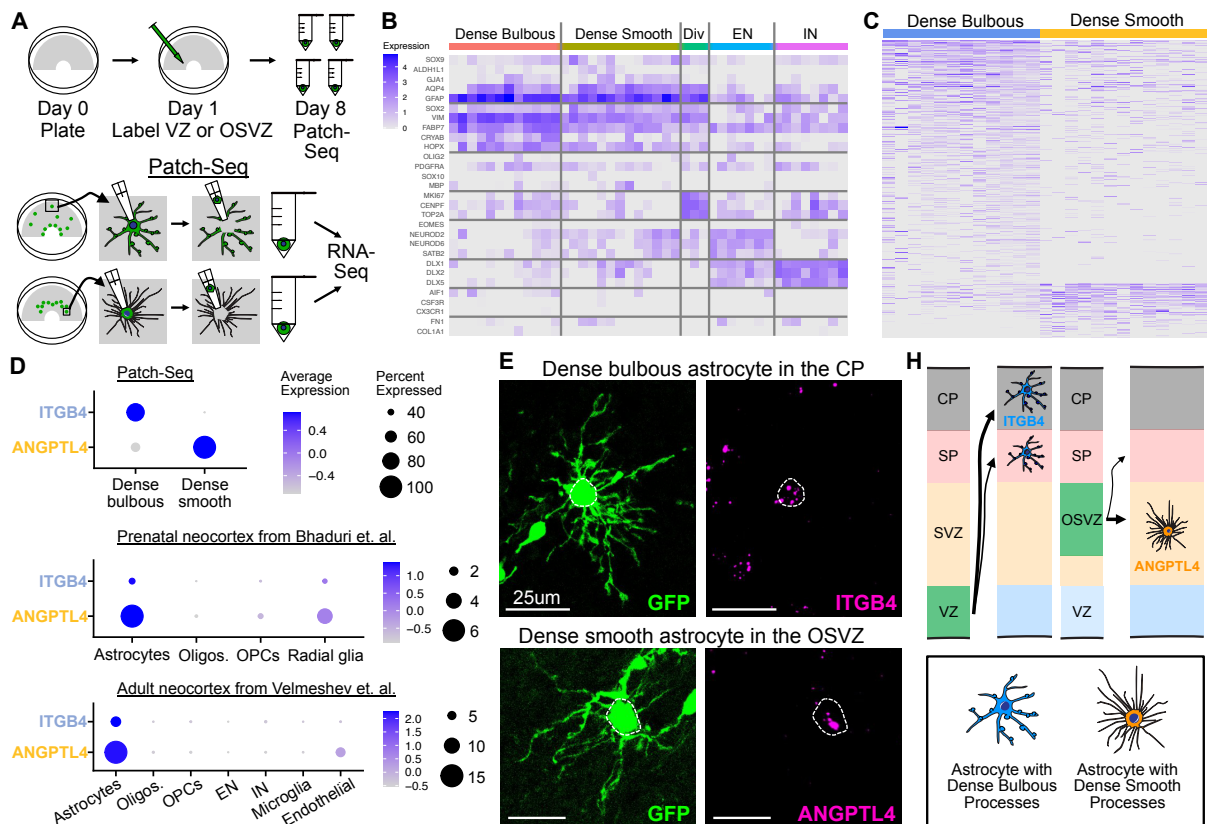
Figure 2.4: "Dense bulbous" and "dense smooth" astrocytes have unique molecular profiles. (**A**) Schematic depicting the collection of mRNA from morphologically-defined astrocyte morphotypes using Patch-seq. (**B**) Heatmap depicting expression of major cell type marker genes across the 45 cells collected by Patch-seq and used for further analysis. Horizontal gray lines indicate the borders between expression of markers of astrocytes, radial glia, oligodendrocyte lineage cells, dividing cells, excitatory neurons, inhibitory neurons, microglia, and endothelial cells. Div: Dividing cells. EN: excitatory neurons. IN: inhibitory neurons. (**C**) Heatmap depicting expression of the 445 genes differentially expressed between "dense bulbous" and "dense smooth" astrocytes (rows), plotted across the 25 astrocytes collected by Patch-seq (columns). (**D**) Dotplots depicting the expression of ITGB4 and ANGPTL4 across astrocyte morphotypes collected by patch-seq and across cell types in the prenatal or adult cortex Oligos: Oligodendrocytes; OPCs: oligodendrocyte progenitor cells; EN: excitatory neurons; IN: inhibitory neurons. (**E**) *In situ* hybridization demonstrating expression of ITGB4 in a VZ-derived dense bulbous astrocyte in the CP and ANGPTL4 expression in an OSVZ-derived dense smooth astrocyte in the OSVZ. (**H**) Schematic summarizing the contribution of VZ and OSVZ to ITGB4+ "dense bulbous" astrocytes and ANGPTL4+ "dense smooth" astrocytes.

28
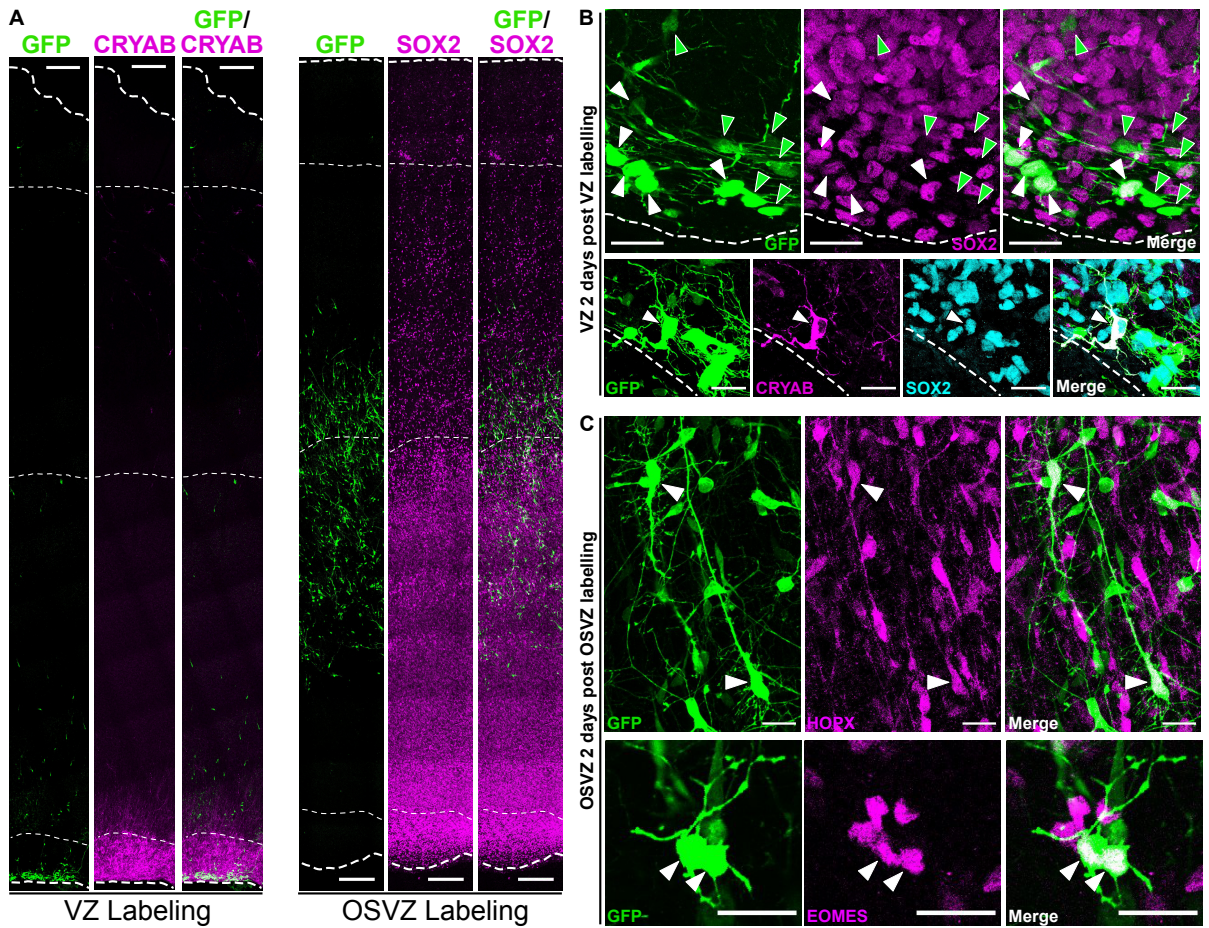
Supplemental Figure 2.1: Individual channel images for Figure 1 panels B-D. (**A**) Separate GFP, CRYAB, and GFP/CRYAB merge images for images shown in Figure 1B. (**B**) Same as (A) for Figure 1C. White arrowheads indicate double- or triple-positive cells. Green arrowheads indicate GFP[+] only cells. (**C**) Same as (A) for Figure 1D. White arrowheads indicate double- or triple-positive cells.
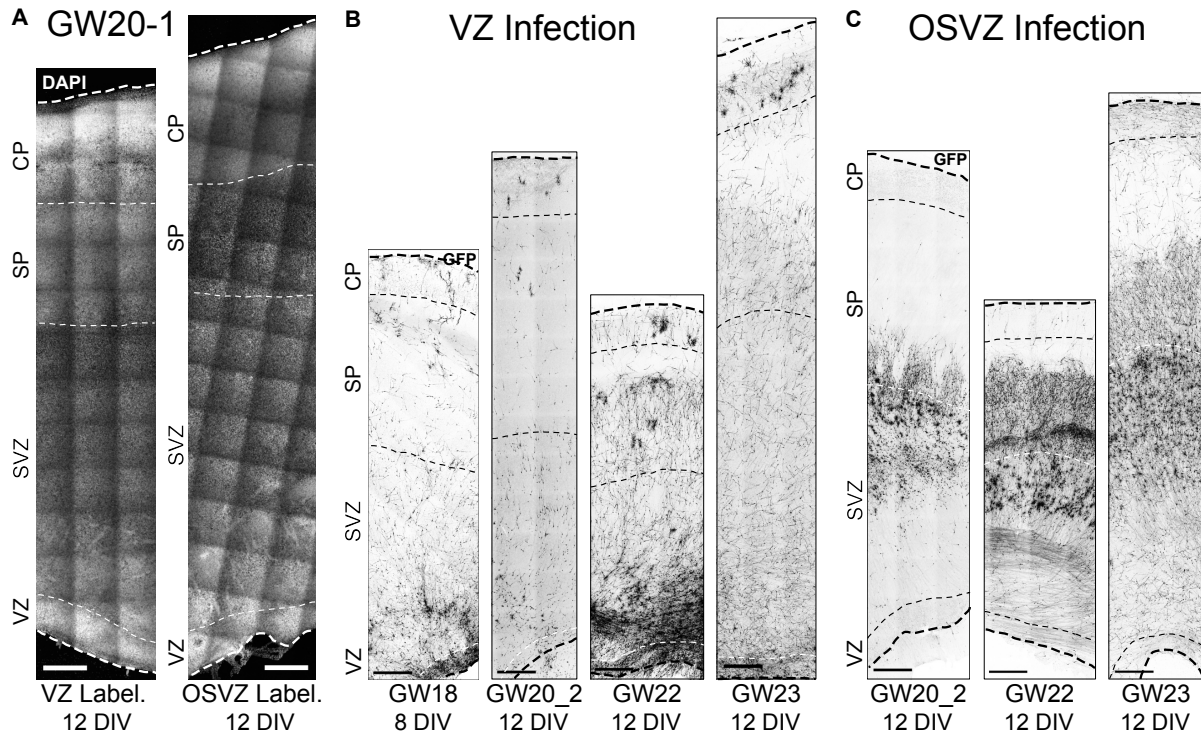
Supplemental Figure 2.2: VZ and OSVZ labeling of additional samples. (**A**) GFP immunostaining showing the four additional biological replicates quantified in Figure 1 showing the distribution of GFP$^+$ VZ-derived cells 8-12 days after labeling. Sample age and number of days cultured in vitro (DIV) under each sample. Thick dashed lines indicate ventricular and pial edges of the slices, thin dashed lines indicate boundaries between laminae. Scale bars = 500µm. (**B**) Same as (A) with initial OSVZ labeling. (**C**) DAPI immunostaining of the GW20 sample shown in Figure 1 demonstrating how boundaries were drawn between laminae based on DAPI density for quantifying the percent of cells in the VZ, SVZ, SP, and CP at the end of culture. Representative of delineations used for quantification of all samples. Scale bars = 500µm.
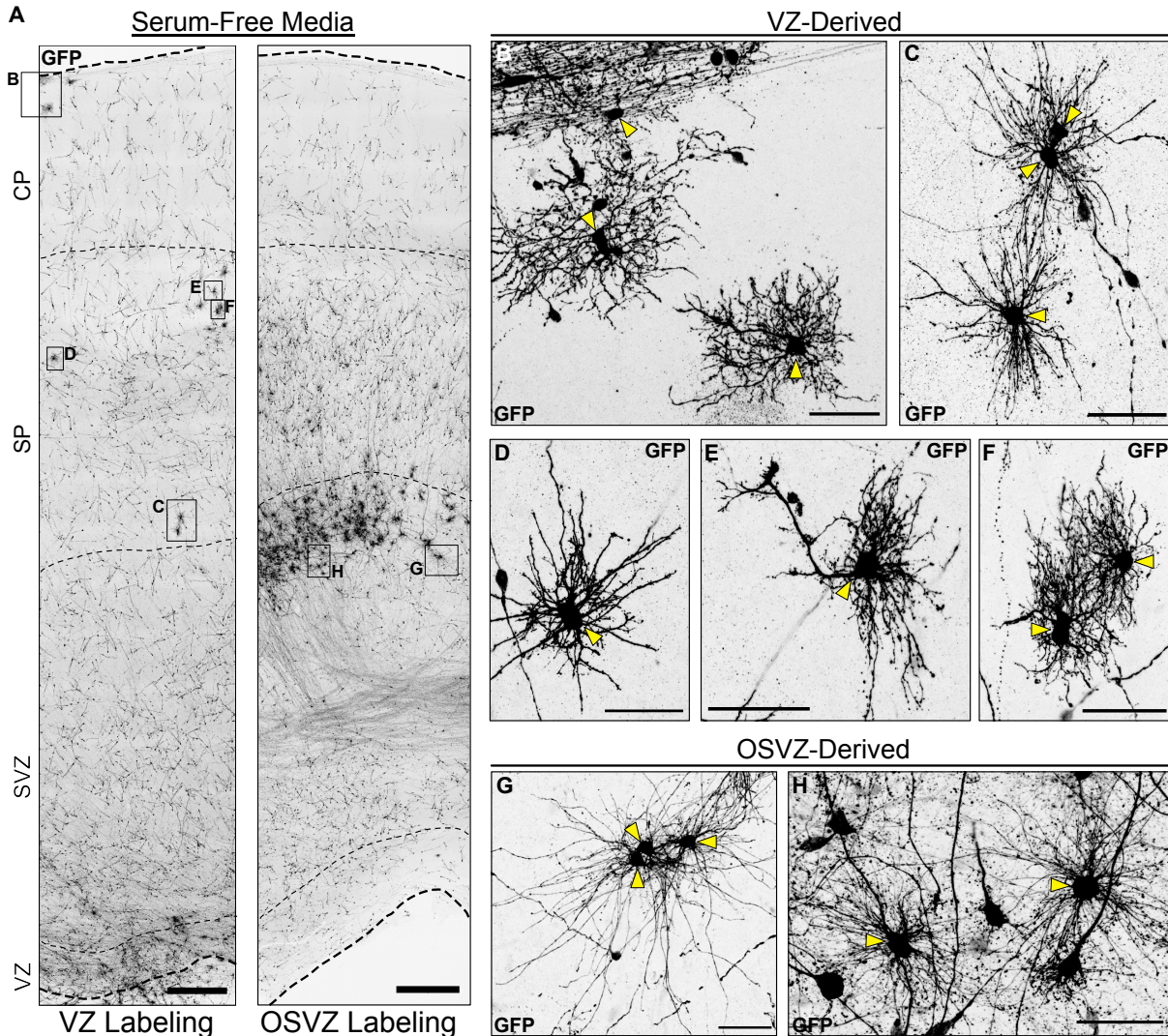
Supplemental Figure 2.3: VZ and OSVZ labeling in serum-free media replicates initial findings. (**A**) GFP immunostaining of VZ- or OSVZ-labeled GW23 sample cultured in media where heat-inactivated fetal bovine serum was replaced with B27 supplement. Thick dashed lines indicate ventricular and pial edges of the slices, thin dashed lines indicate boundaries between laminae. Scale bars = 500μm. (**B-G**) Insets from (**A**) showing astrocytes derived from both niches. Yellow arrowheads indicate the soma of each astrocyte. VZ-derived astrocytes in the SP and CP (**B-F**) have "dense bulbous" astrocyte morphology and OSVZ-derived astrocytes in the OSVZ (**G-H**) have "dense smooth" astrocyte morphology. Scale bars = 50μm.
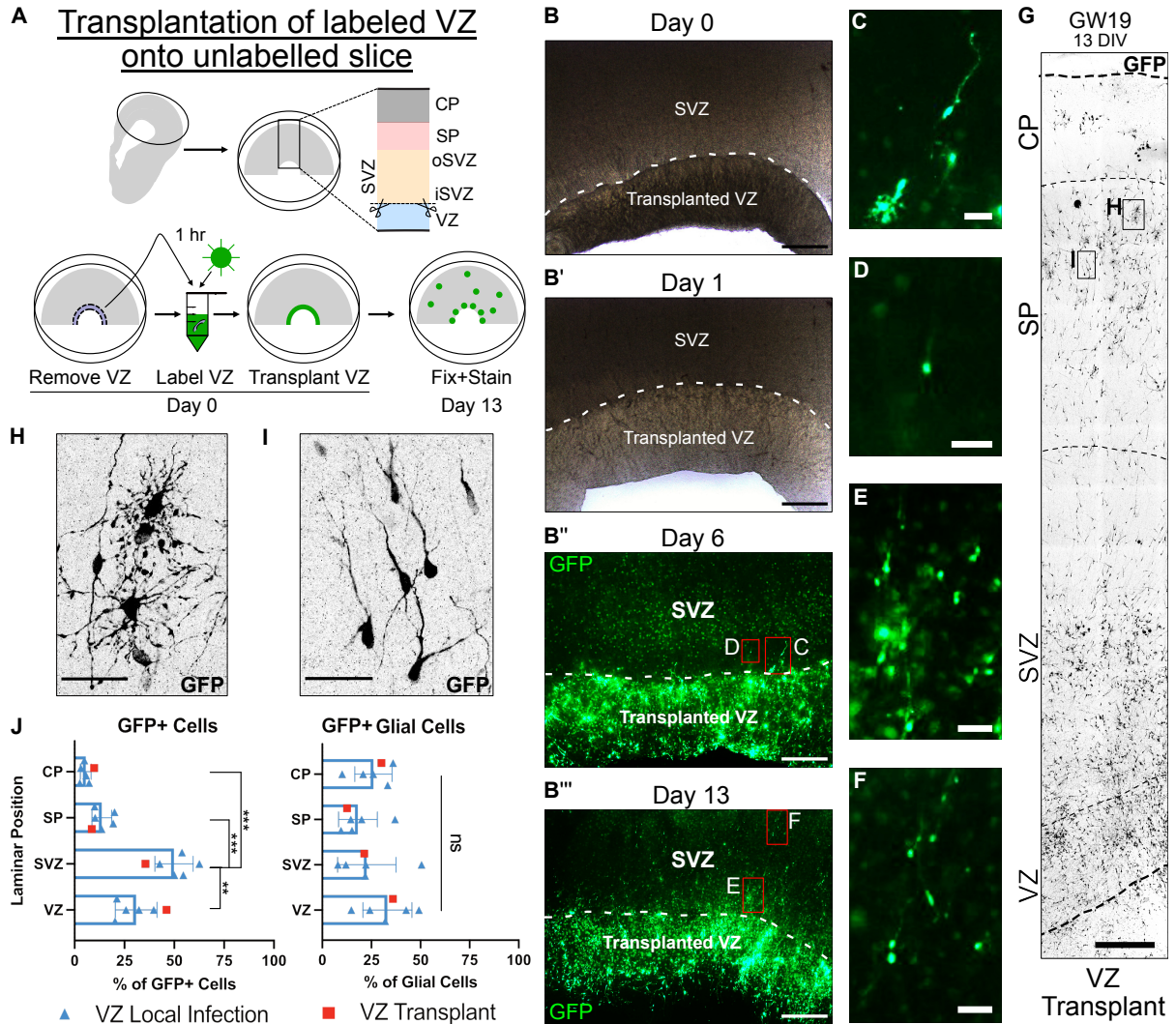
Supplemental Figure 2.4: Transplanted VZ gives rise to cells with neuronal and glial morphology. (**A**) Schematic showing procedure for labeling and transplanting VZ. (**B-B'**) Brightfield images showing placement of transplanted VZ on the day of plating (B), and VZ integration with the unlabeled slice by the next day (B'). Dashed white lines indicate the border between VZ and SVZ. Scale bar = 1mm. (**B''-B'''**) Images showing GFP⁺ cells migrating out of the VZ into the SVZ on day 6 (B'') and 13 (B''') of culture. Dashed white lines indicate the border between VZ and SVZ. Scale bars = 1mm. (**C-F**) Insets from B'' and B''' showing migrating GFP⁺ cells. Scale bars = 100µm. (**G**) GFP immunostaining showing distribution of GFP⁺ cells after 13 days in vitro (DIV). Scale bar = 500µm. (**H-I**) Insets from (G) showing cells with glial (H) and neuronal (I) morphology. Scale bar = 50µm. (**J**) Quantification of VZ-derived cells from Figure 1L with data from GW19 VZ transplant sample added as red squares. Statistical significance assessed with one way ANOVA with Tukey's multiple test correction. *: q < 0.033, **: q < 0.002, ***: q < 0.001 ns: not significant.
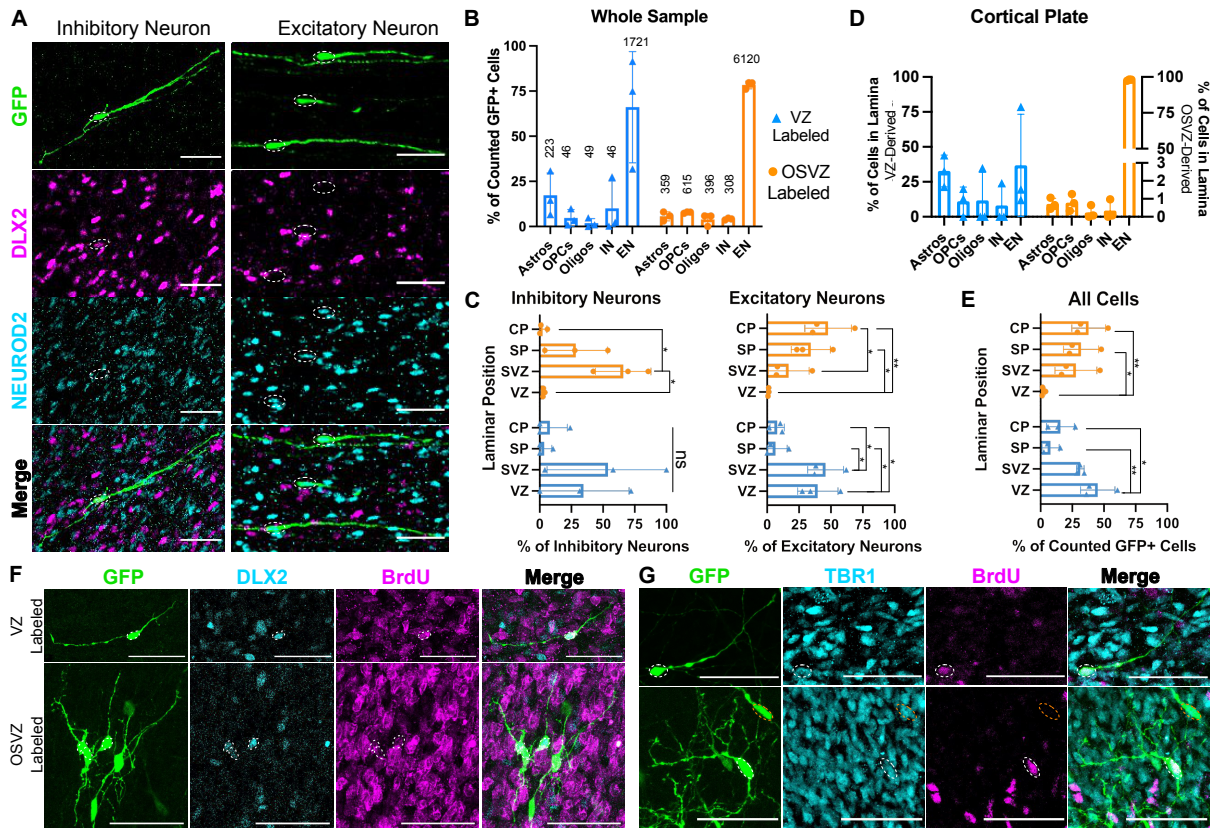
Supplemental Figure 2.5: Both VZ and OSVZ give rise to excitatory and inhibitory neurons. (**A**) Co- immunostaining showing a representative GFP$^+$/DLX2$^+$ inhibitory neuron and GFP$^+$/NEUROD2$^+$ excitatory neurons. Scale bar = 50μm. (**B**) Quantification (mean±SD) of the percentage of all analyzed GFP$^+$ cells derived from VZ or OSVZ labeling that were astrocytes (Astros), OPCs, oligodendrocytes (Oligos), inhibitory neurons (IN), or excitatory neurons (EN). Number of total cells counted in each category written above each bar. (**C**) Quantification (mean±SD) of the percentage of VZ- or OSVZ-derived inhibitory neurons (left) or excitatory neurons (right) that were located in the CP, SP, SVZ, or VZ at the end of culture. (**D**) Quantification (mean±SD) of the percentage of VZ- or OSVZ-derived GFP$^+$ cells in the CP that were astrocytes, OPCs, oligodendrocytes, inhibitory neurons (IN) or excitatory neurons (EN) at the end of culture. Astrocyte, OPC, and oligodendrocyte bars are the same as are plotted in Figure 2C. (**E**) Quantification (mean±SD) of the percentage of all analyzed VZ- or OSVZ-derived GFP$^+$ cells that were located in the CP, SP, SVZ, or VZ. Statistical significance for all comparisons was assessed with two way ANOVA with Tukey's multiple test correction. *: q < 0.033, **: q < 0.002, ***: q < 0.001 ns: not significant. N=3 independent samples for all comparisons. (**F**) Representative images of GFP$^+$/TBR1$^+$/Brdu$^+$ excitatory neurons derived from VZ labeling in GW18 sample. (**G**) Representative images of GFP$^+$/DLX2$^+$/BrdU$^+$ interneurons derived from VZ labeling (left) and OSVZ labeling (right) of GW19 sample.
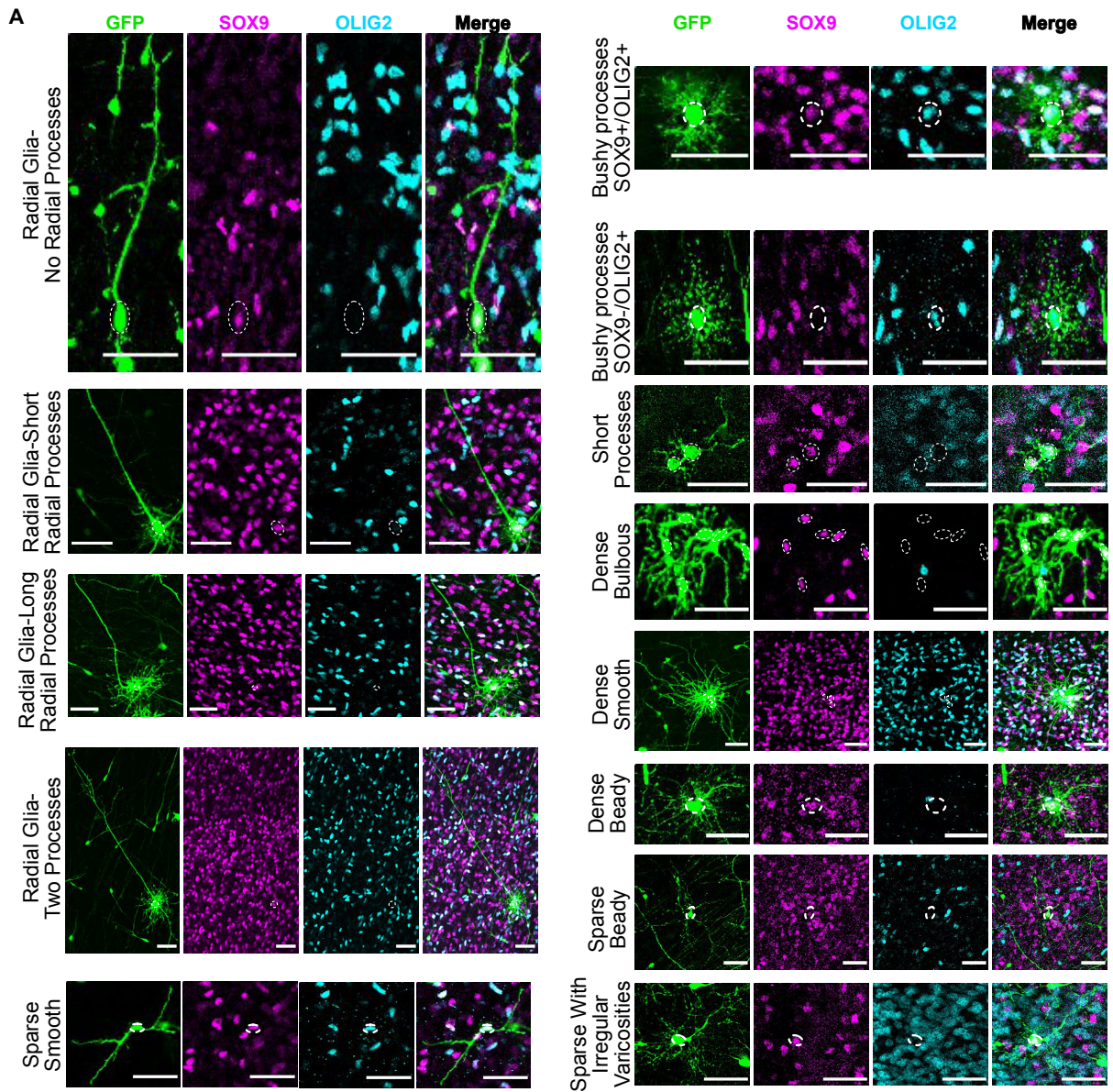
Supplemental Figure 2.6: All identified glial morphotypes express molecular markers of glia.(**A**) Co-immunostaining showing representative images of each glial morphotype that are positive for the pan- glial marker SOX9 and/or the oligodendrocyte lineage marker OLIG2. All scale bars = 50μm.
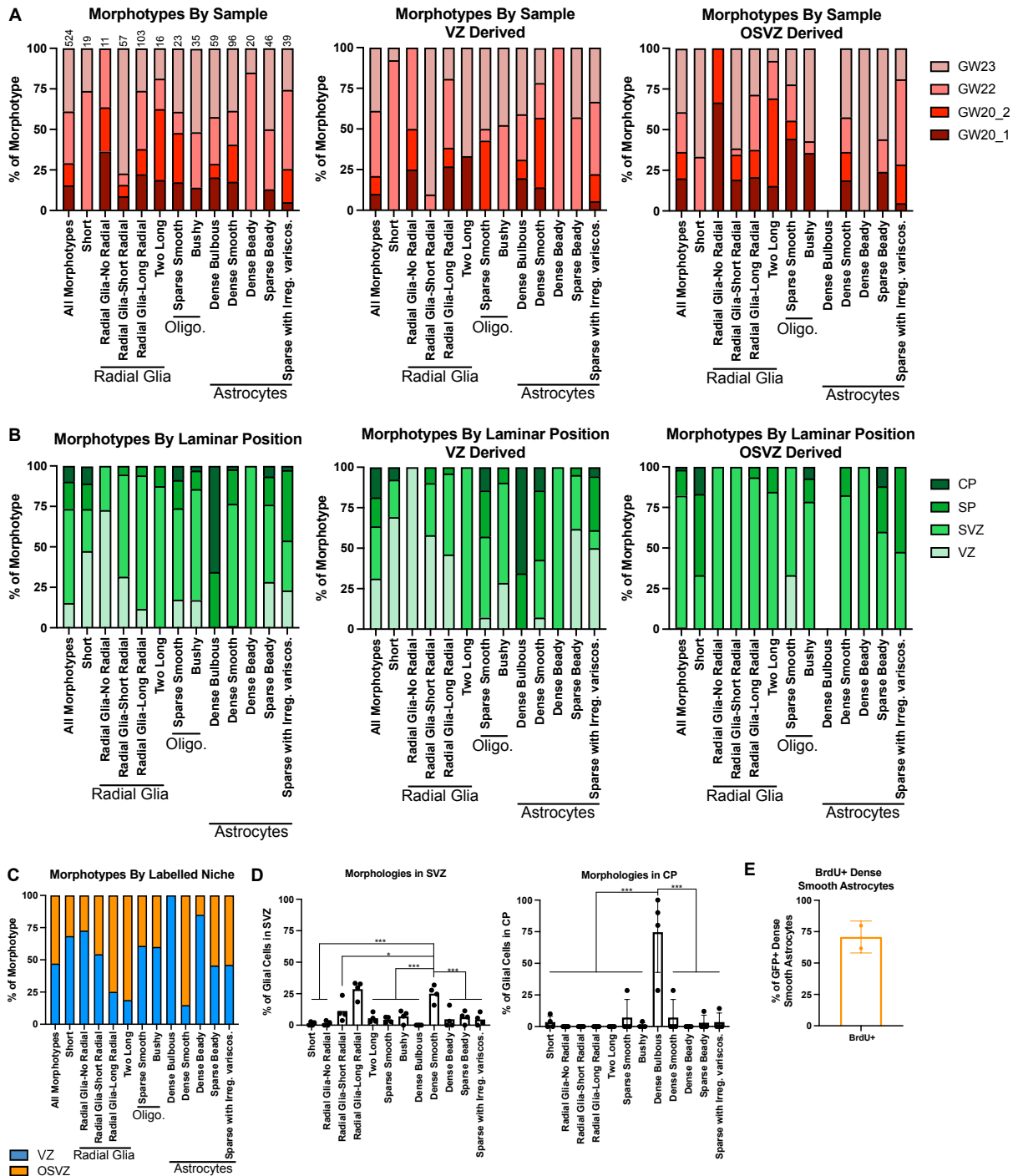
Supplemental Figure 2.7: Metadata for glial morphotype analysis. (**A**) Stacked bar plot showing the percentage of total cells of each glial morphotype derived from each of four independent biological samples, with the total number of cells in each morphotype category above each bar (left). Same analysis split by VZ- derived cells (middle) and OSVZ-derived cells (right). (**B**) Stacked bar plot showing the percentage of cells of each glial morphotype that were located in the CP, SP, SVZ, or VZ at the end of culture (left), and the same analysis split by VZ-derived cells (middle) and OSVZ-derived cells (right). (**C**) Stacked bar plot showing the percentage of cells of each glial morphotype that were derived from VZ or OSVZ labeling. (**D**)

Quantification (mean±SD) of the percentage of analyzed glial cells in the SVZ (left) or CP (right) that adopted each of our twelve glial morphotypes. Statistical significance for all comparisons was assessed with two-way ANOVA with Tukey's multiple test correction. *: q < 0.033, **: q < 0.002, ***: q < 0.001 ns: not significant. N=4 independent samples for all comparisons. (**E**) Quantification (mean±SD) of the percentage of GFP$^+$/SOX9$^+$ "dense smooth" astrocytes that were Brdu$^+$ at the end of culture. Data generated from tissue from two donors.
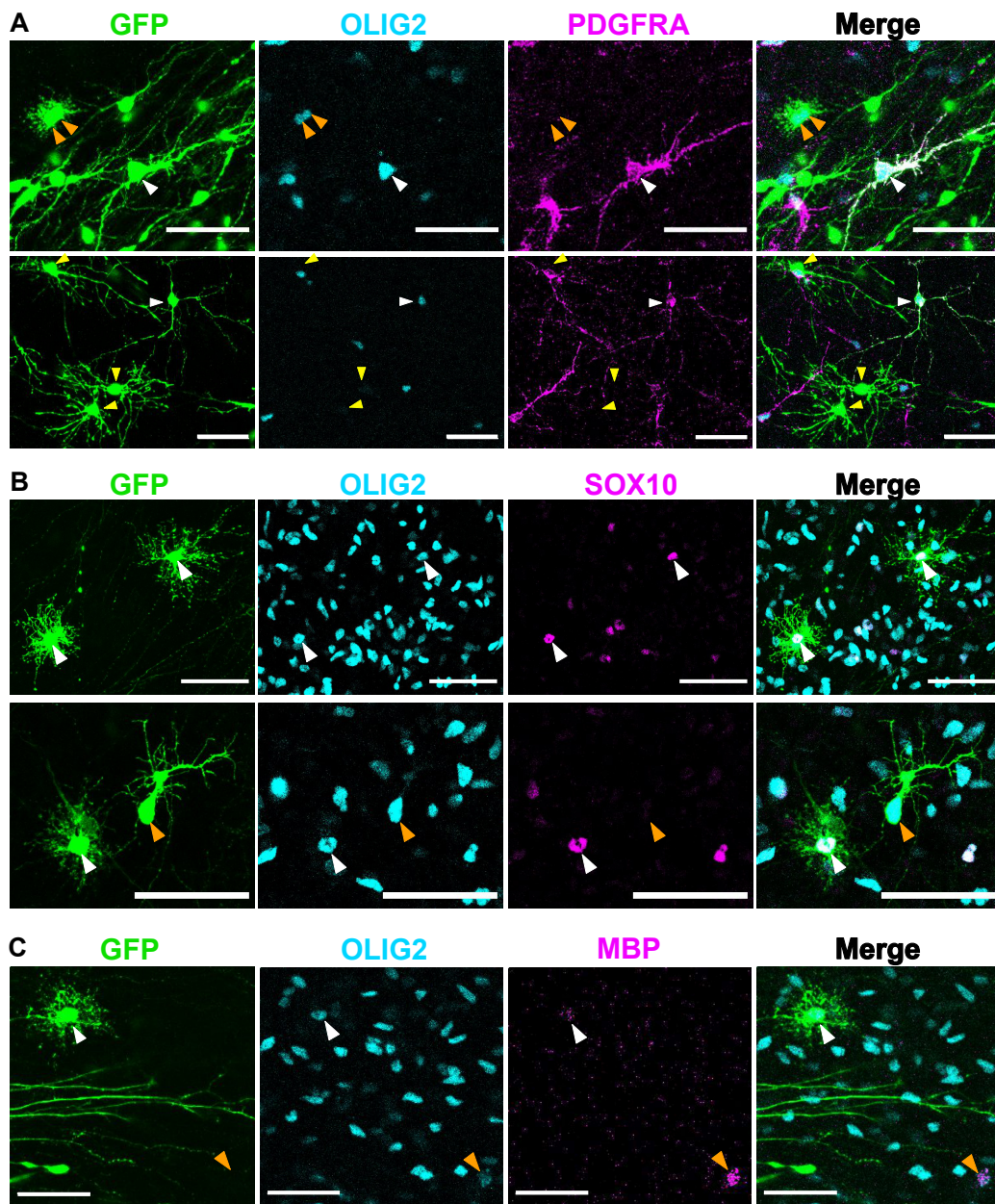
Supplemental Figure 2.8: Additional characterization of oligodendrocyte lineage morphotypes. (**A**) Co- immunostaining demonstrating cells with "sparse smooth processes" that are GFP$^+$/OLIG2$^+$/PDGFRA$^+$ (white arrowheads), consistent with OPC identity. Orange arrowheads indicate GFP$^+$/OLIG2$^+$/PDGFRA$^-$ cells with "bushy" morphology, consistent with oligodendrocyte identity. Yellow arrowheads indicate GFP$^+$/OLIG2$^-$/PDGFRA$^-$ cells with astrocytic morphology. (**B**) Co-immunostaining demonstrating cells with "bushy" processes that are GFP$^+$/OLIG2$^+$/SOX10$^+$ (white arrowheads), consistent with oligodendrocyte identity. Orange arrowheads indicate a GFP$^+$/OLIG2$^+$/SOX10$^-$ cells with "sparse smooth processes", consistent with OPC identity. (**C**) Co-immunostaining for GFP, OLIG2, and the mature oligodendrocyte marker MBP. Orange arrowheads indicate a GFP$^-$/OLIG2$^+$/MBP$^+$ maturing oligodendrocyte and white arrowheads indicate a GFP$^+$/OLIG2$^+$ oligodendrocyte with "bushy" processes that is just beginning to express MBP. All scale bars = 50μm.
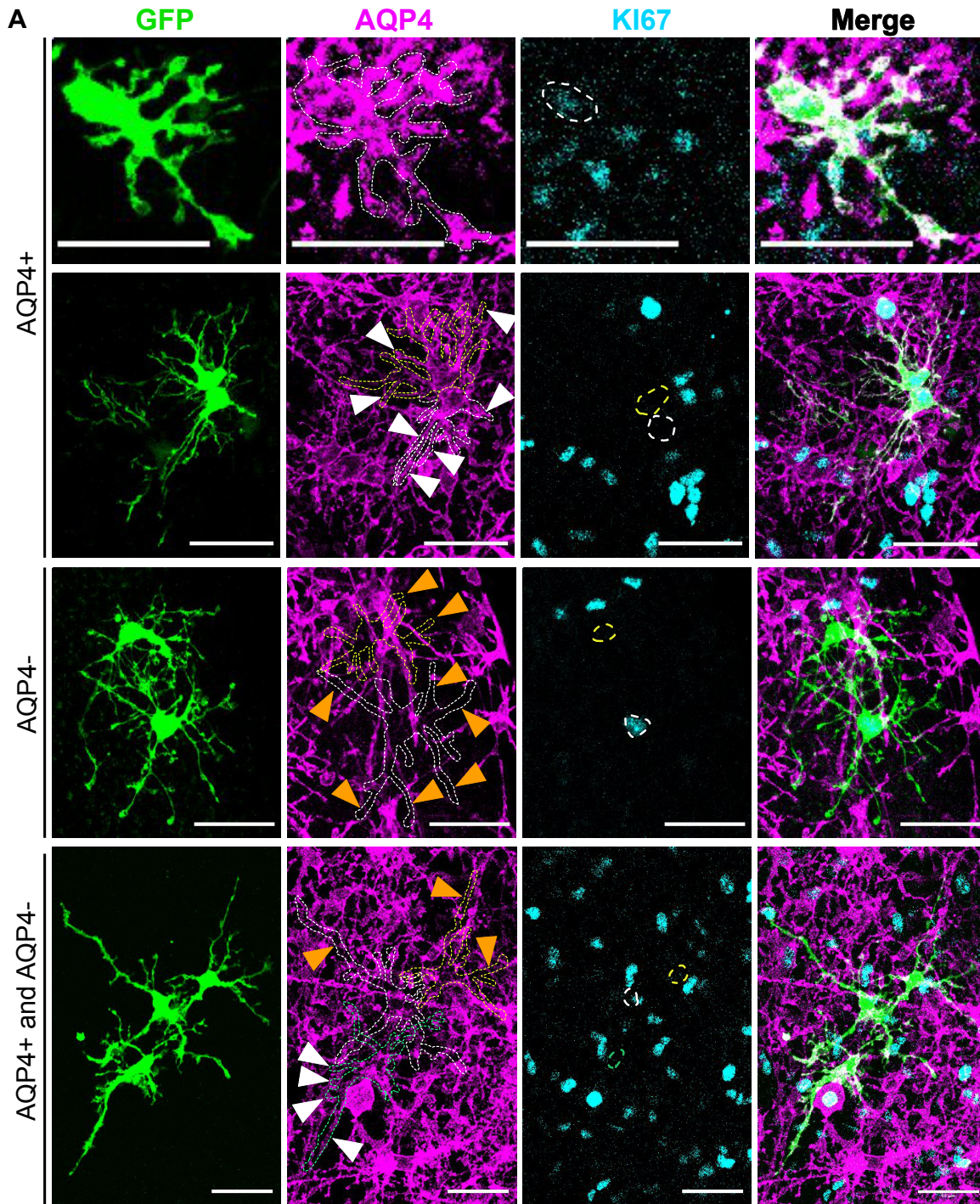
Supplemental Figure 2.9: Dense bulbous astrocytes express AQP4 and KI67. (**A**) Co-immunostaining for GFP, AQP4, and KI67 in astrocytes with dense bulbous processes. Individual cells are outlined in different colored dashed lines and overlaid on AQP4 staining. White arrowheads indicate AQP4$^+$ processes and orange arrowheads indicate AQP4$^-$ processes. The same color dashes lines are used to circle the corresponding nuclei in the KI67 channel. All scale bars = 50µm.
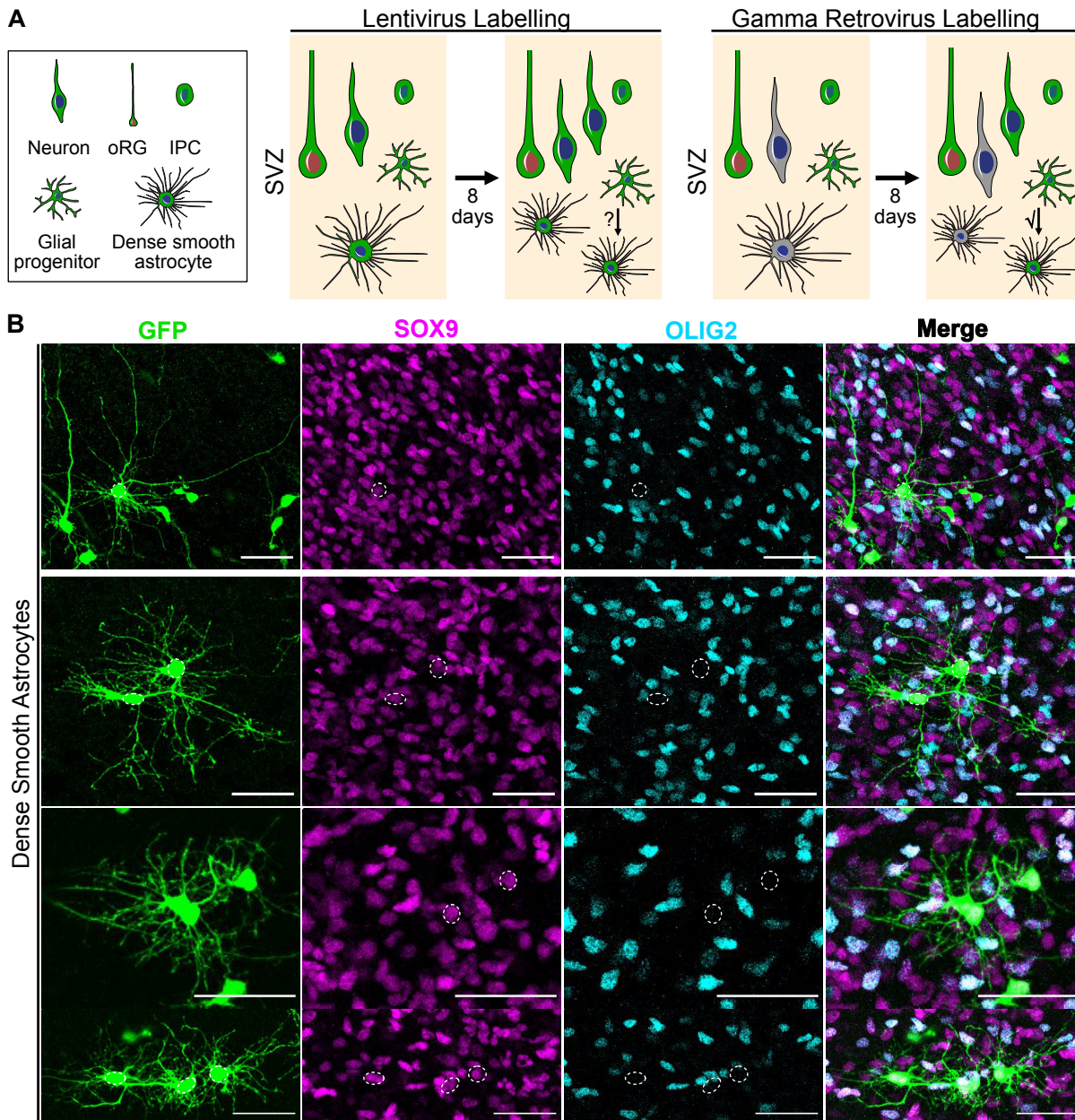
Supplemental Figure 2.10: Gamma retrovirus-labeled cells give rise to astrocytes with dense smooth processes. (**A**) Schematic depicting the difference between lentiviral and gamma retroviral labeling. Lentivirus can infect mitotic and post-mitotic cells while gamma retrovirus can only infect dividing cells. Therefore, any cells labeled by gamma retrovirus are progenitors or their daughter cells. (**B**) Representative images of pNIT- GFP$^+$/SOX9$^+$/OLIG2$^-$ dense smooth astrocytes in the OSVZ of a GW23 sample 7 days after gamma retroviral labeling. Scale bars = 50μm.

A Confocal — GFP; Upright microscope — GFP; Pipet

B Negative Control cDNA; Cell 66 cDNA

C nCount_RNA

D nFeature_RNA

Legend: GW24_1, GW21_1, GW20, GW19_1, GW19_2, GW24_2, GW19_3, GW24_3, GW21_2, GW23

F Prenatal neocortex from Bhaduri et. al.

G Adult neocortex from Velmeshev et. al.

E Expression; Dense Bulbous; Dense Smooth; Continued...
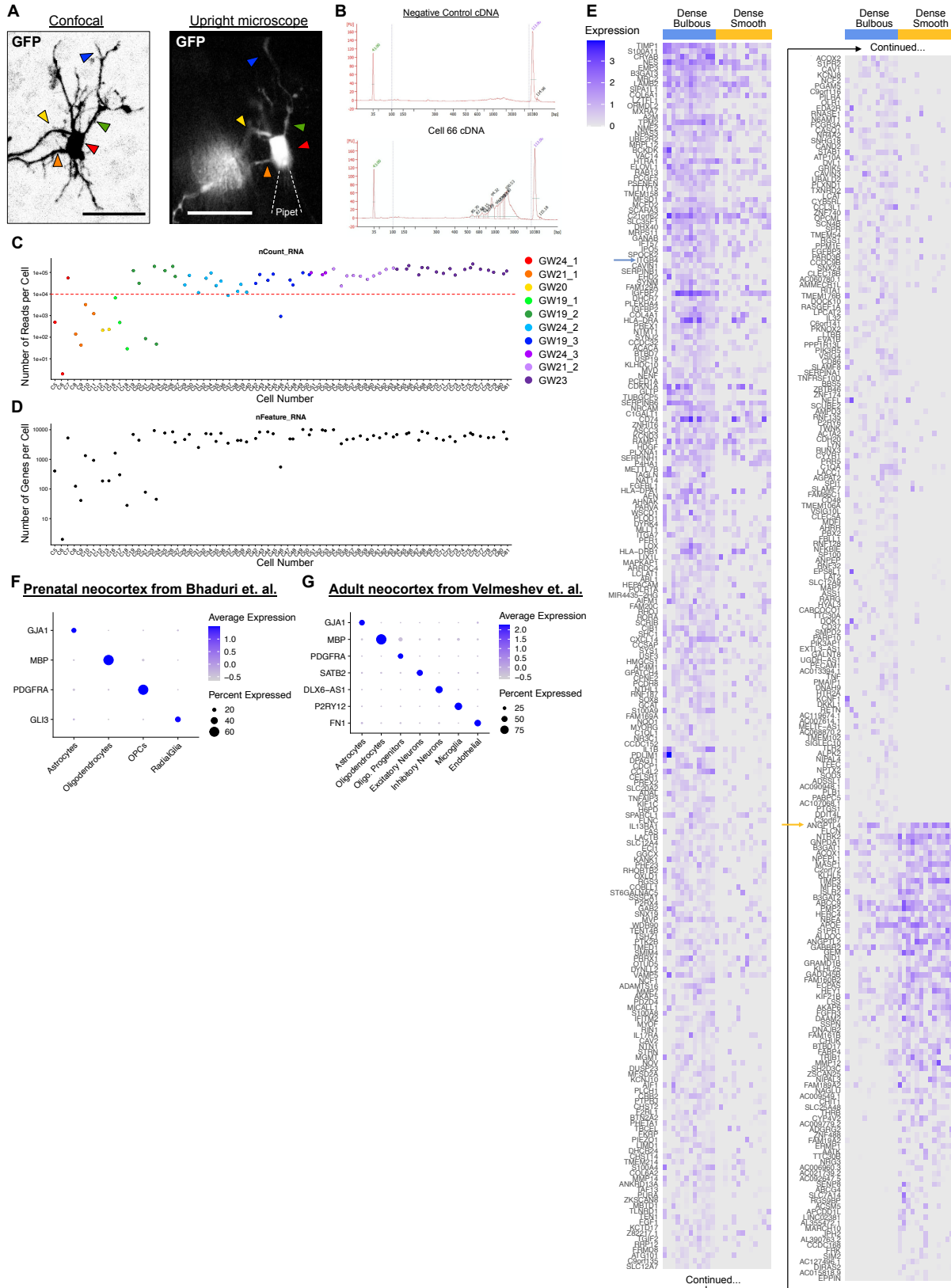
Supplemental Figure 2.11: Supporting data for Patch-seq analysis. (**A**) Confocal image of a representative GFP⁺ astrocyte with dense bulbous processes (left) and image of the same cell visualized on the upright microscope during patching (right). Colored arrows indicate corresponding morphological features. Shadow of the patching pipette indicated by white dashed lines in the right image. Scale bars = 50µm. (**B**) Bioanalyzer traces of cDNA amplified from negative control sample (left) and cell #66 (right). (**C**) Plot of the number of reads detected for all cells collected by Patch-seq. Cells with fewer than 1,000 reads (red dashed line) were discarded. Cells from different biological samples indicated by dot colors. (**D**) Plot of the number of genes detected per cell. (**E**) Same heatmap of genes differential expressed between "dense bulbous" and "dense smooth" astrocytes as in Figure 4C with gene names labeled. Arrows indicate the positions of ITGB4 and ANGPTL4. (**F-G**) Dotplots depicting expression of cell type markers in the correct annotated cell types in Bhaduri et. al. (F) and Velmeshev et. al. (G).
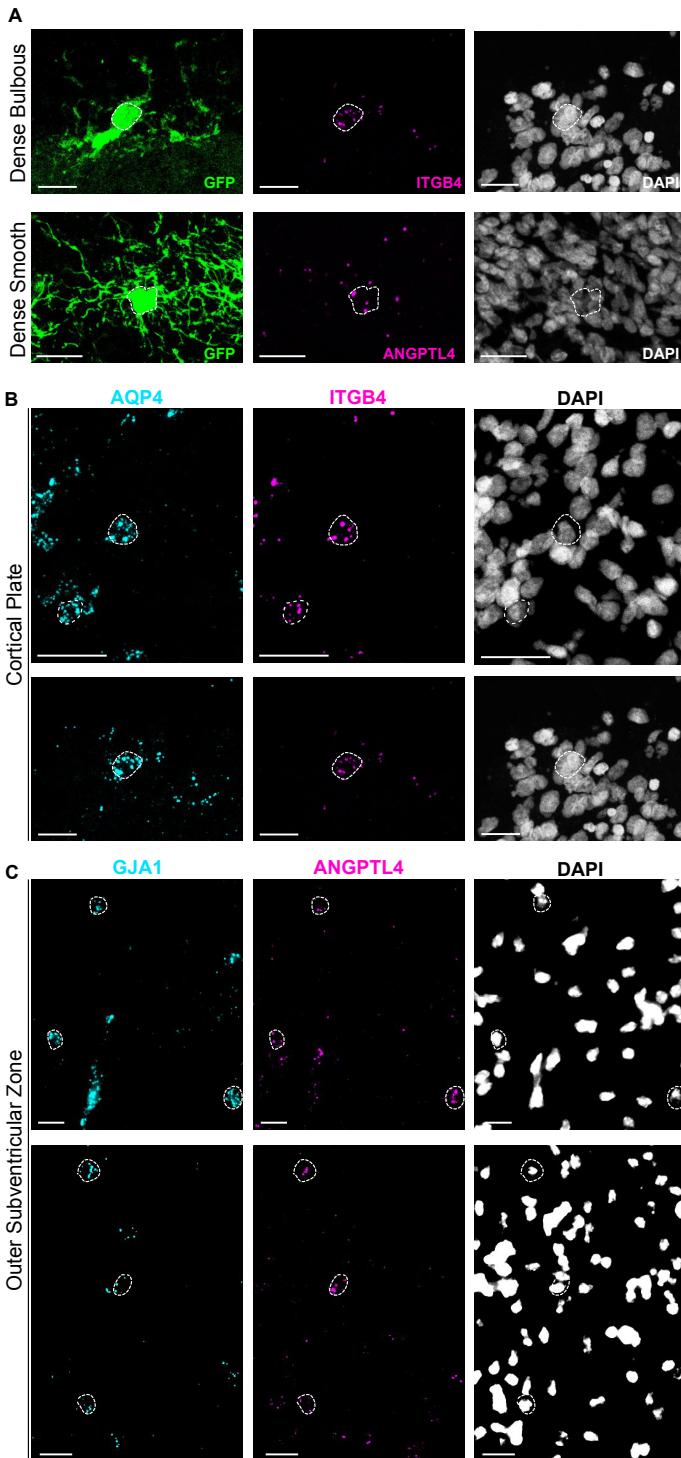
Supplemental Figure 2.12: Expression of ITGB4 and ANGPTL4 in CP and OSVZ astrocytes. (**A**) Additional images of an ITGB4[+] VZ-derived dense bulbous astrocyte in the CP (top) and an ANGPTL4[+] OSVZ-derived dense smooth astrocyte in the OSVZ (bottom). (**B**) Multiplexed *in situ* hybridization demonstrating expression of ITGB4 in AQP4[+] astrocytes in the CP. (**C**) Multiplexed in situ hybridization demonstrating expression of ANGPTL4 in GJA1[+] astrocytes in the OSVZ.

MATERIALS AND METHODS

Organotypic slice culture of prenatal brain slices

Deidentified primary tissue samples were collected with previous patient consent in strict observance of the legal and institutional ethical regulations. Cortical brain tissue was immediately placed in a sterile conical tube filled with oxygenated artificial spinal fluid (aSCF) containing 125 mM NaCl, 2.5 mM KCl, 1mM $MgCl_2$, 1 mM $CaCl_2$, and 1.25 mM NaH_2PO_4 bubbled with carbogen (95% O_2/5% CO_2). Only tissue that has a visibly intact ventricular surface was selected for use. Blood vessels and meninges were removed from the cortical tissue, and then the tissue block was embedded in 3.5% low-melting-point agarose (Thermo Fisher, BP165-25) and sectioned perpendicular to the ventricle to 300 µm using a Leica VT1200S vibrating blade microtome in aCSF. Slices were transferred to slice culture inserts (Millicell, PICM03050) on six-well culture plates (Corning) and cultured in a 37 °C incubator at 5% CO_2, 8% O_2 at the liquid-air interface created by the cell-culture insert. The prenatal brain slice culture medium was composed of 66% (vol/vol) Basal Medium Eagle's (UCSF Stem Cell Core, CCFZ022), 25% (vol/vol) Hank's Balanced Salt Solution (Lonza, 10-527F), 5% (vol/vol) heat inactivated FBS (Hyclone, SH30071.03HI), 1% N2 supplement (Thermo Scientific, 17502-048), 1% penicillin/ streptomycin (Gibco, 15140122), and 1% GlutaMax (Gibco, 35050-061). For serum-free cultures, the heat inactivated FBS was replaced with 1% B27 (Invitrogen, 12587-010). For BrdU experiments, media was supplemented with 10 µM BrdU (Thermo Fisher, B23151) for either the first four days post-VZ labeling, or for all seven days of culture post-OSVZ labeling. Media was half-changed every other day until the experimental endpoint. If the slice cultures started to disintegrate or accumulate large numbers of dead cells early in culture the experiment was terminated and not used.

Local VZ and OSVZ Viral Labeling

The morning after plating, lentivirus encoding CAG-EGFP was diluted in slice culture media and approximately 1-2ul of virus was applied to the VZ or OSVZ using a fine glass needle. We also chose to use the hybrid CMV/chicken β-actin (CAG) promoter as this promoter maintains strong expression in postmitotic neurons, in contrast to the cytomegalovirus (CMV) promoter which becomes inactive in neurons. Using a Leica MZ10F dissecting microscope, VZ labeling was targeted to the ventricular surface of each slice and OSVZ labeling was targeted to the OSVZ by targeting just inside of the basal most edge of the fiber tracts running through the OSVZ. Virus was targeted to the center of each germinal zone to minimize leakage of virus around the edges of the slices. For labeling with gamma retrovirus, the same technique was used with undiluted pNIT-GFP Moloney Murine Leukemia Virus (MMLV) (gift from Fred Gage at the Salk Institute). Slices were then transferred back to the incubator and media was half-changed every other day until experimental endpoint.

VZ Transplantation

Slices of prenatal brain tissue were prepared as described above. Prior to plating, slices were transferred to a dish of ice cold PBS and visualized under a Leica MZ10F dissecting microscope to identify and remove the VZ using a microscalpel (Surgical Specialties, 72-2201). The VZs were then transferred to a 1.5 ml tube containing CAG-GFP lentivirus diluted in PBS using a micropipette with wide bore tip (VWR, 89049-168). The VZs were incubated in the diluted virus for 1 hour in a 37 °C incubator at 5% CO2, 8% O2 and then gently washed 3x in ice cold PBS. Meanwhile, the remaining slices were plated on slice culture inserts as described above. After labeling and washing, the VZs were placed adjacent to the cut edge of the SVZ of the plated slices from the same individual using a p20 tip. By the next day the transplanted VZ fused with the adjacent unlabeled slice (Fig. S2.4). Brightfield images of the VZ transplant sample were collected using an Olympus CKX53-T inverted microscope with 4x air objective. Images of the

CAG-GFP labeling were collected using the same microscope and objective with illumination from a CoolLED pE-300 illumination source and GFP filter cube.

Immunofluorescence of prenatal brain slices

Prenatal brain slices were fixed in 4% PFA (Fisher, #50-980-487) in PBS for 1 hour at 4oC with gentle agitation. When staining for BrdU, antigen retrieval was performed by incubating slices in 1N HCl for 1 hour followed by neutralization in 1M Boric Acid for 30 minutes. Blocking and permeabilization were performed in a blocking solution consisting of 10% normal donkey serum and 1% Triton X-100 in PBS for >1 hour. Primary antibodies were diluted in the same blocking solution and slices were incubated with primary antibodies overnight at 4oC with gentle agitation. Slices were then washed 5x 30 minutes in PBS + 1% Triton-X (1% PBST) and incubated in secondary antibodies diluted in blocking solution overnight at 4oC with gentle agitation. Slices were washed 5x 30 minutes in 1% PBST and mounted on glass slides with ProLong Gold Antifade Mountant (Invitrogen, P36930) and #1.5 coverslips (Azer Scientific, 1152460).

The antibodies used in this study include: chicken anti-GFP (Aves, GFP-1020, 1:1000), rabbit anti-NEUROD2 (Abcam, ab104430, 1:500), mouse anti-DLX2 (Santa Cruz Biotechnology, SC-393879), rabbit anti-SOX9 (Abcam, ab104430, 1:250), mouse anti-OLIG2 (Millipore Sigma, MABN50, 1:200), ms anti-CRYAB (Abcam, AB13496), goat anti-SOX2 (Santa Cruz Biotechnologies, sc-17320), rabbit anti-HOPX (Santa Cruz Biotechnologies, SC-30216, 1:500), mouse anti-EOMES (eBioscience, 14-4877-82, 1:500), rabbit anti- PDGFRA (Cell Signaling, 3174S, 1:200), rabbit anti-MBP (Abcam, ab40390, 1:100), rabbit anti-AQP4 (Atlas Antibodies, HPA014784, 1:200), mouse anti-KI67 (Dako, M7240, 1:200), rabbit anti-SOX10 (Abcam, ab155279, 1:200), rabbit anti-TBR1 (Abcam, AB31940, 1:200), and rat anti-BrdU (Abcam, ab6326, 1:200). Secondaries used include donkey anti-chicken 488 (Jackson Immunoresearch, 703-545-155, 1:500) and AlexaFluor goat anti-chicken 488, donkey anti-rabbit 594, donkey anti-

mouse 594, donkey anti-rat 594, donkey anti-goat 594, donkey anti-rabbit 647, and donkey anti-mouse 647 (Thermo Fisher Scientific, 1:1000).

Immunostaining quantification

Four independent biological replicates (two GW20s, one GW22, and one GW23) of paired VZ and OSVZ infections were used for glial morphology analysis in Figure 4. Tile scans from the VZ to the pia of each sample were collected using a Leica SP8 confocal with 20x air objective and 5μm optical sections. All tile scans were cropped to a similar width using ImageJ/Fiji and the VZ, SVZ, SP and CP lamina were defined using DAPI nuclear density (Fig. S2.2C). We defined cells with glial morphology as cells that had multiple primary processes and no axon, and we only considered glial cells that did not significantly overlap with adjacent cells so that their entire morphology could be visualized. We randomly selected at least 20 cells from each sample, z-projected the z planes that contained the target cell, converted the image to grayscale, and inverted the colors. These cell images were then manually grouped and split by DA based on 1) primary process density, 2) primary process length, and 3) types of process varicosities as been described in previous studies (*18*, *19*, *38*, *39*). We also used co-immunostaining to confirm that cells from each morphotype category were positive for the glial markers SOX9 and/or OLIG2 (Fig. S2.6), and required that each category be composed of cells from at least two samples (though most categories include cells from all four samples; Fig. S2.7A).

Four independent samples (one GW19, two GW21, and one GW23) were used to determine the percentage of SOX2$^+$, CRYAB$^+$, HOPX$^+$, or EOMES$^+$ GFP cells after initial labeling in Figure 1E-F. For quantification of SOX2$^+$ and CRYAB$^+$ cells, tile scans from the VZ to the OSVZ of VZ-labeled slices were acquired as above but with 63x magnification to better resolve the CRYAB staining. These images were then adjusted to the same brightness/contrast settings, cropped to a similar width, and cropped to just the VZ based on SOX2 and CRYAB staining density. For quantification of HOPX+ and EOMES+ cells, tilescans from the VZ to the SP of OSVZ-labeled

slices were acquired as above. These images were then also adjusted to the same brightness/contrast settings across samples, cropped to a similar width, and cropped to just the OSVZ based on HOPX and EOMES staining density. The number of GFP+ and GFP+/marker+ cells within each compartment were then counted manually using the CellCounter plugin for ImageJ/Fiji.

The same pairs of VZ and OSVZ infections (two GW20, one GW22, and one GW23) used for the morphology analysis above, in addition to a fifth GW18 VZ infection, were used to determine GFP+ cell distribution in Figure 1L. The same VZ, SVZ, SP, and CP borders were also used. The distribution of the glial cells used for the morphology analysis were used to determine the distribution of GFP+ glial cells in Figure 1L. The distribution of all GFP+ cells was determined by manually counting with the CellCounter plugin in ImageJ/Fiji.

Tissue from four independent donors (one GW19, one GW20, and two GW21s) were used to quantify the distribution of molecularly defined astrocytes, OPCs, and oligodendrocytes, and three independent samples (the same GW19, GW20, and one GW21) were used for neuronal quantification. Tilescans from the VZ to the pia of VZ- or OSVZ-labeled slices were collected as above, and borders between the VZ, SVZ, SP, and CP were determined based on DAPI density as above. The images were then adjusted to the same brightness/contrast settings and cropped to a similar width. The number of GFP+/marker+ cells were counted manually using the CellCounter plugin for ImageJ.

All data is plotted as mean±standard deviation in Graphpad Prism (version 9.2.0 for MacOS). Statistical significance was assessed using one- or two-way ANOVA with Tukey's multiple test correction in Graphpad Prism.

Patch-seq

Human neocortical organotypic slice cultures were prepared, VZ or OSVZ labeled with CAG-GFP lentivirus, and cultured for 10 days as described above. After 10 days in culture, cells with

dense bulbous processes were imaged using a Leica SP8 confocal system at 20x

magnification. Imaging was performed in an environmental chamber at 37 °C and humidified

with 5% CO2, 5% O2, balance % N2 to maintain slice health. After imaging, slices were

transferred to a room temperature bath of continually carbogenated artificial cerebrospinal fluid

(aCSF) containing: 124 mM NaCl, 24 mM NaHCO3, 12.5 mM glucose, 5 mM HEPES, 2.5

mM KCl, 2 mM CaCl2·2H2O, 2 mM MgSO4·2H2O, and 1.25 mM NaH2PO4 (pH 7.3-7.4). After

recovering for at least 45 minutes in the bath, slices were then transferred to a continually

perfused, carbogenated aCSF bath, heated to 32.5°C, and mounted on an upright microscope

(BX51WI, Olympus). Target dense bulbous astrocytes were identified by matching their

appearance in the images taken on the Leica SP8 to the fluorescence signal on the upright

microscope, and astrocytes with dense smooth processes were directly identified on the upright

microscope. Cells were visualized for patching with differential interference contrast microscopy.

Patching was performed according to (40). Briefly, patch electrodes (tip resistance 3-6 MOhms)

were filled with 0.2 to 0.5 µl of (in mM): 111 potassium gluconate, 13.4 biocytin, 10 HEPES, 5

phosphocreatine, 4 KCl, 4 MgATP, 0.3 Na3GTP, 0.2 mM EGTA, and 1 U/µl RNase inhibitor

(Takara, 2313B). Giga seal or semi-loose seal access was obtained on the target cell, then the

cell was plucked from the tissue by drawing the pipette out of the bath similar to Bardy et al. (41)

and transferred using an *ad hoc* positive pressure delivery device into a lysis buffer containing 4

µl of lysis buffer containing 5 mM dNTP, 2.5 µM Oligo- dT30 VN, 1 U/µl RNase inhibitor, 0.1%

vol/vol Triton X-100, and 20 µg/mL glycogen (40) . Negative control samples were collected by

inserting the pipette into the tissue without targeting a specific cell. Specimens were stored for

up to 1 month at -80°C until sequencing library generation. Reverse transcription and cDNA

amplification were performed as described in (40). Sequencing libraries were prepared using

the Nextera XT DNA Library Preparation Kit (Illumina, FC-131-1096) and Nextera Index Kit v2

Set A (Illumina, FC-131-2001) with the following modifications to make it compatible with Patch-

Seq: reactions were run at half volume (10ul), 1ul of cDNA input was replaced with 1ul of 40%

PEG-8000 (Sigma, P1458-25ML), 500ng of cDNA was used as input for each sample, and

tagmentation was run for 6 minutes. Libraries were then combined and cleaned as in (*40*) and

sequenced on a MiSeq using MiSeq Reagent Kit v2 (300-cycles) (Illumina, MS-102-2002) to an

average depth of 74,022 reads per cell.

Patch-Seq data analysis

MiSeq reads were demultiplexed using Illumina's Generate FASTQ program. Reads were

aligned to the human genome (GRCh38-3.0.0) using STAR 2.7.9a with the --soloType

SmartSeq --soloUMIdedup Exact -- soloStrand Unstranded flags for Smart-seq2 data. The cell

by gene matrix generated in Solo.out was then loaded into Seurat 4.0.4 in R 4.1.0 using the

ReadMtx command with argument min.cells = 1 to filter out genes that were not detected in any

cells.

Cells with fewer than 10,000 reads per cell were removed. The standard Seurat workflow was

then used to log- normalize and scale the data and perform principal component analysis. The

DoHeatmap command was used to determine the expression of radial glia, astrocyte,

oligodendrocyte, dividing, excitatory neuron, inhibitory neuron, microglia, and endothelial cell

markers across all cells. In accordance with recommendations by Tripathy et. al. (*42*), 11

additional cells were removed for abnormal gene expression likely caused by collection of cell

fragments from adjacent cells during target cell patching. These cells included 8 neurons that

expressed high levels of glial genes and 3 "dense bulbous" astrocytes that expressed high

levels of interneuron genes. We also removed 4 cells that had poor gene detection across all

markers. Astrocytes with strong expression of dividing cell markers were segregated into a

"dividing" category. Neuronal cells were segregated into excitatory and inhibitory categories.

The data was then re-processed using the standard Seurat workflow.

The FindAllMarkers command was then used to identify differentially expressed genes between

dense bulbous and dense smooth astrocytes. These markers were filtered to genes with low

pct.2 to identify the most specific markers. Expression of these most specific markers was then projected onto glial cells from Bhaduri et. al. (see below) to identify markers that were also expressed in prenatal astrocytes, and onto Velmeshev et. al. to identify markers that were specific to astrocytes in adult cortex. The plots in Fig. 2.4 and Fig. S2.11 were generated using Seurat visualization tools.

Analysis of scRNA-seq from Bhaduri et. al. and Velmeshev et. al.

Data for all cells from the GW18, GW19, GW20, GW22, and GW25 samples from Bhaduri et. al. (*24*) was obtained directly from Dr. Aparna Bhaduri. Seurat's FindAllMarkers command was used to identify top markers of each cluster, and clusters whose top markers included the radial glia markers GLI3, GFAP, SOX2, CRYAB, or HOPX, the astrocyte markers GJA1, AQP4, or SPARCL1, or the oligodendrocyte markers PDGFRA, OLIG2, or MBP were kept. These cells were then iteratively clustered using the standard Seurat v3 pipeline and clusters without these top marker genes were removed until only radial glia, astrocytes, and oligodendrocyte lineage cells remained.

For cross referencing with the Velmeshev et. al. data, cells with the diagnosis of "ASD" were removed and the remaining Control cells were re-processed through the standard Seurat workflow prior to analysis of candidate dense bulbous and dense smooth astrocyte markers.

***In situ* hybridization for ITGB4 and ANGPTL4**

VZ- or OSVZ-labeled slices were prepared as above from a GW20 sample. The slices were cultured for 10 days before being fixed in 4% PFA for 1 hour at 4oC, cryopreserved in 30% sucrose overnight at 4oC, and then embedded in a 50:50 ratio of 30% sucrose and OCT (VWR, 25608-930) and frozen solid. The samples were then sectioned to 20μm using a CryoStar NX50 cryostat (Thermo Scientific), mounted on glass slides, and frozen at -80C. Sections containing target astrocyte morphotypes were counterstained with DAPI and coverslips were mounted with

Prolong Gold Antifade Mountant. GFP+ VZ-derived dense bulbous astrocytes in the CP and OSVZ-derived dense smooth astrocytes in the OSVZ were imaged using a Leica SP8 confocal microscope with 20x air objective+2x optical zoom and 1μm optical sections. The imaging positions were saved to be re-imaged after *in situ* hybridization. The DAPI channel was also imaged to aid with image registration. The coverslips were then removed by soaking the slides in 5x SSC for 1-3 hours. *In situ* hybridization for AQP4, GJA1, ITGB4, and ANGPTL4 was performed on these slides using the RNAscope Multiplex Fluorescent v2 kit (Advanced Cell Diagnostics, 323100) according to the manufacturer's instructions. The same positions as above were then reimaged and the DAPI channels used for image registration to align GFP+ cells with *in situ* AQP4, GJA1, ITGB4, and/or ANGPTL4 signal.

ACKNOWLEDGEMENTS

AUTHOR CONTRIBUTIONS

Conceptualization: DEA, TJN. Methodology: DEA, KCD, CC, DS, MGK, TJN. Investigation: DEA, KCD, CC, DS, MGK. Visualization: DEA, MGK, TJN. Funding acquisition: DEA, KCD, DS, VSS, TJN. Project administration: DEA, TJN. Supervision: VSS, TJN. Writing -- original draft: DEA, TJN. Writing -- review and editing: DEA, KCD, CC, DS, MGK, TJN.

COMPETING INTERESTS

The authors declare no competing interests.

DATA AND MATERIALS AVAILABILITY

Morphology classification data is provided in Supplementary Table 1. Patch-seq counts are

provided in Supplementary Table 2 and corresponding metadata are provided in Supplementary

Table 3. Code used to generate the data is available at the Github repository:

https://zenodo.org/badge/latestdoi/471078301. Materials requests should be addressed to TJN.

REFERENCES

1. C. R. Cadwell, A. Bhaduri, M. A. Mostajo-Radji, M. G. Keefe, T. J. Nowakowski, Development and Arealization of the Cerebral Cortex. *Neuron*. **103**, 980–1004 (2019).

2. B. Libé-Philippot, P. Vanderhaeghen, Cellular and Molecular Mechanisms Linking Human Cortical Development and Evolution. *Annu. Rev. Genet.* **55**, 555–581 (2021).

3. T. J. Nowakowski, A. A. Pollen, C. Sandoval-Espinosa, A. R. Kriegstein, Transformation of the Radial Glia Scaffold Demarcates Two Stages of Human Cerebral Cortex Development. *Neuron*. **91**, 1219–1227 (2016).

4. E. Taverna, M. Götz, W. B. Huttner, The cell biology of neurogenesis: toward an understanding of the development and evolution of the neocortex. *Annu. Rev. Cell Dev. Biol.* **30**, 465–502 (2014).

5. N. Kalebic, C. Gilardi, B. Stepien, M. Wilsch-Bräuninger, K. R. Long, T. Namba, M. Florio, B. Langen, B. Lombardot, A. Shevchenko, M. W. Kilimann, H. Kawasaki, P. Wimberger, W. B. Huttner, Neocortical Expansion Due to Increased Proliferation of Basal Progenitors Is Linked to Changes in Their Morphology. *Cell Stem Cell*. **24**, 535–550.e9 (2019).

6. N. Kalebic, W. B. Huttner, Basal Progenitor Morphology and Neocortex Evolution. *Trends Neurosci.* **43**, 843–853 (2020).

7. S. A. Fietz, I. Kelava, J. Vogt, M. Wilsch-Bräuninger, D. Stenzel, J. L. Fish, D. Corbeil, A. Riehn, W. Distler, R. Nitsch, W. B. Huttner, OSVZ progenitors of human and ferret neocortex are epithelial-like and expand by integrin signaling. *Nat. Neurosci.* **13**, 690–699 (2010).

8. I. H. M. Smart, C. Dehay, P. Giroud, M. Berland, H. Kennedy, 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 (2002).

9. D. V. Hansen, J. H. Lui, P. R. L. Parker, A. R. Kriegstein, Neurogenic radial glia in the outer subventricular zone of human neocortex. *Nature*. **464**, 554–561 (2010).

10. I. Reillo, C. de Juan Romero, M. Á. García-Cabezas, V. Borrell, A role for intermediate radial glia in the tangential expansion of the mammalian cerebral cortex. *Cereb. Cortex*. **21**, 1674–1694 (2011).

11. A. A. Pollen, T. J. Nowakowski, J. Chen, H. Retallack, C. Sandoval-Espinosa, C. R. Nicholas, J. Shuga, S. J. Liu, M. C. Oldham, A. Diaz, D. A. Lim, A. A. Leyrat, J. A. West, A. R. Kriegstein, Molecular identity of human outer radial glia during cortical development. *Cell*. **163**, 55–67 (2015).

12. M. Betizeau, V. Cortay, D. Patti, S. Pfister, E. Gautier, A. Bellemin-Ménard, M. Afanassieff, C. Huissoud, R. J. Douglas, H. Kennedy, C. Dehay, Precursor diversity and complexity of lineage relationships in the outer subventricular zone of the primate. *Neuron*. **80**, 442–457 (2013).

13. L. C. deAzevedo, C. Fallet, V. Moura-Neto, C. Daumas-Duport, C. Hedin-Pereira, R. Lent, Cortical radial glial cells in human fetuses: depth-correlated transformation into astrocytes. *J. Neurobiol.* **55**, 288–298 (2003).

14. B. G. Rash, A. Duque, Y. M. Morozov, J. I. Arellano, N. Micali, P. Rakic, Gliogenesis in the outer subventricular zone promotes enlargement and gyrification of the primate cerebrum. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 7089–7094 (2019).

15. W. Huang, A. Bhaduri, D. Velmeshev, S. Wang, L. Wang, C. A. Rottkamp, A. Alvarez-Buylla, D. H. Rowitch, A. R. Kriegstein, Origins and Proliferative States of Human

Oligodendrocyte Precursor Cells. *Cell*. **182**, 594–608.e11 (2020).

16. O. Marín, J. L. Rubenstein, A long, remarkable journey: tangential migration in the telencephalon. *Nat. Rev. Neurosci.* **2**, 780–790 (2001).

17. R. N. Delgado, D. E. Allen, M. G. Keefe, W. R. Mancia Leon, R. S. Ziffra, E. E. Crouch, A. Alvarez-Buylla, T. J. Nowakowski, Individual human cortical progenitors can produce excitatory and inhibitory neurons. *Nature*. **601**, 397–403 (2022).

18. M. Marín-Padilla, Prenatal development of fibrous (white matter), protoplasmic (gray matter), and layer I astrocytes in the human cerebral cortex: a Golgi study. *J. Comp. Neurol.* **357**, 554–572 (1995).

19. C. Falcone, E. Penna, T. Hong, A. F. Tarantal, P. R. Hof, W. D. Hopkins, C. C. Sherwood, S. C. Noctor, V. Martínez-Cerdeño, Cortical Interlaminar Astrocytes Are Generated Prenatally, Mature Postnatally, and Express Unique Markers in Human and Nonhuman Primates. *Cereb. Cortex*. **31**, 379–395 (2021).

20. H. Akiyama, I. Tooyama, T. Kawamata, K. Ikeda, P. L. McGeer, Morphological diversities of CD44 positive astrocytes in the cerebral cortex of normal subjects and patients with Alzheimer's disease. *Brain Res.* **632**, 249–259 (1993).

21. D. E. Schmechel, P. Rakic, A Golgi study of radial glial cells in developing monkey telencephalon: morphogenesis and transformation into astrocytes. *Anat. Embryol.* . **156**, 115–152 (1979).

22. D. E. Bergles, W. D. Richardson, Oligodendrocyte Development and Plasticity. *Cold Spring Harb. Perspect. Biol.* **8**, a020453 (2015).

23. W.-P. Ge, A. Miyawaki, F. H. Gage, Y. N. Jan, L. Y. Jan, Local generation of glia is a major

astrocyte source in postnatal cortex. *Nature*. **484**, 376–380 (2012).

24. A. Bhaduri, C. Sandoval-Espinosa, M. Otero-Garcia, I. Oh, R. Yin, U. C. Eze, T. J. Nowakowski, A. R. Kriegstein, An atlas of cortical arealization identifies dynamic molecular signatures. *Nature*. **598**, 200–204 (2021).

25. D. Velmeshev, L. Schirmer, D. Jung, M. Haeussler, Y. Perez, S. Mayer, A. Bhaduri, N. Goyal, D. H. Rowitch, A. R. Kriegstein, Single-cell genomics identifies cell type-specific molecular changes in autism. *Science*. **364**, 685–689 (2019).

26. S. Malik, G. Vinukonda, L. R. Vose, D. Diamond, B. B. R. Bhimavarapu, F. Hu, M. T. Zia, R. Hevner, N. Zecevic, P. Ballabh, Neurogenesis continues in the third trimester of pregnancy and is suppressed by premature birth. *J. Neurosci.* **33**, 411–423 (2013).

27. S. C. Noctor, V. Martínez-Cerdeño, L. Ivic, A. R. Kriegstein, Cortical neurons arise in symmetric and asymmetric division zones and migrate through specific phases. *Nat. Neurosci.* **7**, 136–144 (2004).

28. V. Matyash, H. Kettenmann, Heterogeneity in astrocyte morphology and physiology. *Brain Res. Rev.* **63**, 2–10 (2010).

29. C. Falcone, E. L. McBride, W. D. Hopkins, P. R. Hof, P. R. Manger, C. C. Sherwood, S. C. Noctor, V. Martínez-Cerdeño, Redefining varicose projection astrocytes in primates. *Glia*. **70**, 145–154 (2022).

30. C. Falcone, M. Wolf-Ochoa, S. Amina, T. Hong, G. Vakilzadeh, W. D. Hopkins, P. R. Hof, C. C. Sherwood, P. R. Manger, S. C. Noctor, V. Martínez-Cerdeño, Cortical interlaminar astrocytes across the therian mammal radiation. *J. Comp. Neurol.* **527**, 1654–1674 (2019).

31. H.-G. Lee, M. A. Wheeler, F. J. Quintana, Function and therapeutic value of astrocytes in

neurological diseases. *Nat. Rev. Drug Discov.* (2022), doi:10.1038/s41573-022-00390-x.

32. K. A. Guttenplan, M. K. Weigel, P. Prakash, P. R. Wijewardhane, P. Hasel, U. Rufen-Blanchette, A. E. Münch, J. A. Blum, J. Fine, M. C. Neal, K. D. Bruce, A. D. Gitler, G. Chopra, S. A. Liddelow, B. A. Barres, Neurotoxic reactive astrocytes induce cell death via saturated lipids. *Nature*. **599**, 102–107 (2021).

33. C. Hercher, V. Chopra, C. L. Beasley, Evidence for morphological alterations in prefrontal white matter glia in schizophrenia and bipolar disorder. *J. Psychiatry Neurosci.* **39**, 376–385 (2014).

34. Y. Hu, I. Ylivinkka, P. Chen, L. Li, S. Hautaniemi, T. A. Nyman, J. Keski-Oja, M. Hyytiäinen, Netrin-4 promotes glioblastoma cell proliferation through integrin β4 signaling. *Neoplasia*. **14**, 219–227 (2012).

35. J. Li, X. Wang, L. Chen, J. Zhang, Y. Zhang, X. Ren, J. Sun, X. Fan, J. Fan, T. Li, L. Tong, L. Yi, L. Chen, J. Liu, G. Shang, X. Ren, H. Zhang, S. Yu, H. Ming, Q. Huang, J. Dong, C. Zhang, X. Yang, TMEM158 promotes the proliferation and migration of glioma cells via STAT3 signaling in glioblastomas. *Cancer Gene Ther.* (2022), doi:10.1038/s41417-021-00414-5.

36. M. E. Hegi, A.-C. Diserens, T. Gorlia, M.-F. Hamou, N. de Tribolet, M. Weller, J. M. Kros, J. A. Hainfellner, W. Mason, L. Mariani, J. E. C. Bromberg, P. Hau, R. O. Mirimanoff, J. G. Cairncross, R. C. Janzer, R. Stupp, MGMT gene silencing and benefit from temozolomide in glioblastoma. *N. Engl. J. Med.* **352**, 997–1003 (2005).

37. G. Wang, Y. Li, D. Zhang, S. Zhao, Q. Zhang, C. Luo, X. Sun, B. Zhang, CELSR1 Acts as an Oncogene Regulated by miR-199a-5p in Glioma. *Cancer Manag. Res.* **12**, 8857–8865 (2020).

38. P. Rakic, Neuron-glia relationship during granule cell migration in developing cerebellar cortex. A Golgi and electronmicroscopic study in Macacus Rhesus. *J. Comp. Neurol.* **141**, 283–312 (1971).

39. N. A. Oberheim, T. Takano, X. Han, W. He, J. H. C. Lin, F. Wang, Q. Xu, J. D. Wyatt, W. Pilcher, J. G. Ojemann, B. R. Ransom, S. A. Goldman, M. Nedergaard, Uniquely hominid features of adult human astrocytes. *J. Neurosci.* **29**, 3276–3287 (2009).

40. C. R. Cadwell, F. Scala, S. Li, G. Livrizzi, S. Shen, R. Sandberg, X. Jiang, A. S. Tolias, Multimodal profiling of single-cell morphology, electrophysiology, and gene expression using Patch-seq. *Nat. Protoc.* **12**, 2531–2553 (2017).

41. C. Bardy, M. van den Hurk, B. Kakaradov, J. A. Erwin, B. N. Jaeger, R. V. Hernandez, T. Eames, A. A. Paucar, M. Gorris, C. Marchand, R. Jappelli, J. Barron, A. K. Bryant, M. Kellogg, R. S. Lasken, B. P. F. Rutten, H. W. M. Steinbusch, G. W. Yeo, F. H. Gage, Predicting the functional states of human iPSC-derived neurons with single-cell RNA-seq and electrophysiology. *Mol. Psychiatry*. **21**, 1573–1588 (2016).

42. S. J. Tripathy, L. Toker, C. Bomkamp, B. O. Mancarci, M. Belmadani, P. Pavlidis, Assessing Transcriptome Quality in Patch-Seq Datasets. *Front. Mol. Neurosci.* **11**, 363 (2018).

CHAPTER 3

At the time of submission of this dissertation, this article was published in Nature.

Individual human cortical progenitors can produce excitatory and inhibitory neurons

Ryan N. Delgado[1,2,3,4,8,9], Denise E. Allen[1,2,3,4,8], Matthew G. Keefe[1,2,3,4,8], Walter R. Mancia Leon[3,5], Ryan S. Ziffra[1,2,3,4], Elizabeth E. Crouch[3,6], Arturo Alvarez-Buylla[3,5] & Tomasz J. Nowakowski[1,2,3,4,5,7,9]

Affiliations:

1 Department of Anatomy, University of California at San Francisco, San Francisco, CA, USA.

2 Department of Psychiatry and Behavioral Sciences, University of California at San Francisco, San Francisco, CA, USA.

3 The Eli and Edythe Broad Center of Regeneration Medicine and Stem Cell Research, University of California at San Francisco, San Francisco, CA, USA.

4 Weill Institute for Neurosciences, University of California at San Francisco, San Francisco, CA, USA.

5 Department of Neurological Surgery, University of California at San Francisco, San Francisco, CA, USA.

6 Department of Pediatrics, University of California at San Francisco, San Francisco, CA, USA.

7 Chan Zuckerberg Biohub, San Francisco, CA, USA.

8 These authors contributed equally: Ryan N. Delgado, Denise E. Allen, Matthew G. Keefe.

9 Corresponding authors: Ryan.N.Delgado@gmail.com; Tomasz.Nowakowski@ucsf.edu

Abstract

The cerebral cortex is a cellularly complex structure comprising a rich diversity of neuronal and glial cell types. Cortical neurons can be broadly categorized into two classes—excitatory neurons that use the neurotransmitter glutamate, and inhibitory interneurons that use γ-aminobutyric acid (GABA). Previous developmental studies in rodents have led to a prevailing model in which excitatory neurons are born from progenitors located in the cortex, whereas cortical interneurons are born from a separate population of progenitors located outside the developing cortex in the ganglionic eminences[1–5] However, the developmental potential of human cortical progenitors has not been thoroughly explored. Here we show that, in addition to excitatory neurons and glia, human cortical progenitors are also capable of producing GABAergic neurons with the transcriptional characteristics and morphologies of cortical interneurons. By developing a cellular barcoding tool called 'single-cellRNA-sequencing-compatible tracer for identifying clonal relationships' (STICR), we were able to carry out clonal lineage tracing of 1,912 primary human cortical progenitors from six specimens, and to capture both the transcriptional identities and the clonal relationships of their progeny. A subpopulation of cortically born GABAergic neurons was transcriptionally similar to cortical interneurons born from the caudal ganglionic eminence, and these cells were frequently related to excitatory neurons and glia. Our results show that individual human cortical progenitors can generate both excitatory neurons and cortical interneurons, providing a new framework for understanding the origins of neuronal diversity in the human cortex.

Background

The neocortex is responsible for carrying out many higher-order cognitive functions, such as decision making, language comprehension and sensory perception. During neural development, progenitors in the cortex and ganglionic eminences produce the diverse array of neurons and glia that comprise the neocortex. Extensive developmental studies in rodents have

60

demonstrated that the two principal types of cortical neuron—glutamatergic excitatory neurons and GABAergic inhibitory interneurons—are produced by two distinct groups of progenitors[1–5]. Excitatory neurons are generated by progenitors located dorsally in the developing cortex and migrate radially to occupy their terminal positions in the cortical plate. By contrast, cortical interneurons are generated by progenitors located in the ganglionic eminences of the ventral forebrain and then migrate dorsally into the developing neocortex. A few limited studies have suggested the possibility that human cortical progenitors might also give rise to cortical interneurons, but the findings have been conflicting[6–10]. Short-term clonal labelling of progenitors in the human cortex revealed the local generation of newborn GABAergic neurons[7], but did not determine whether they were cortical interneurons or another type of interneuron such as those of the olfactory bulb, which have previously been shown to derive from cortical progenitors in mice[11–13]. Similarly, in vitro cultures derived from human cortical progenitors have been shown to generate GABAergic inhibitory neurons[10]. Conversely, a later study that labelled dividing cells in short-term cultures of human organotypic tissue slices did not find a substantial fraction of newborn inhibitory neurons in the cortex[8]. The overwhelming consensus remains that human cortical progenitors give rise to excitatory neurons but not cortical inhibitory neurons; however, this has not been thoroughly examined and the developmental potential of individual human cortical progenitors remains largely unknown.

Results

Design and validation of lineage tracer

In order to carry out high-throughput clonal lineage tracing of primary human neural progenitors, we developed STICR—an ultra-high-complexity barcoding strategy that allows us to permanently label cells and their progeny. This approach uses a lentivirus encoding a heritable, transcribed molecular barcode within the 3′-untranslated region of the enhanced green fluorescent protein (EGFP) reporter gene (Fig. 3.1a). The combinatorial, single-molecule

barcode design of STICR allows for a predefined, error-correctable barcode library with a maximum diversity of 125 million sequences (Extended Data Fig. 3.1a; see Methods). Deep sequencing of STICR plasmid and lentiviral libraries confirmed their ultra-high barcode diversity (approximately 50 million to 65 million unique barcodes per library) without any overrepresented barcodes (Extended Data Fig. 3.1b). Using the observed barcode diversity and frequency from each STICR library, we modelled the rate of barcode 'collision'—the event in which two different cells are independently labelled with the same barcode—and found that STICR could be used to label more than 250,000 cells before reaching an estimated barcode collision rate of around 0.5% (Extended Data Fig. 3.1c). To confirm the accuracy of STICR barcode recovery from single cells, we carried out a cell-mixing, or 'barnyard', experiment in which we labelled mouse and human cells with different STICR libraries that were readily distinguishable by a constant 'viral index' sequence unique to each library (Extended Data Fig. 3.1d). Following single-cell RNA sequencing (scRNA-seq), recovered STICR barcodes were 100% concordant with the STICR library used to infect each population (Extended Data Fig. 3.1e, f), indicating that our method is accurate and can be applied robustly to carry out high-throughput clonal lineage tracing.

Lineage tracing of human neural progenitors

To determine the developmental potential of individual human cortical progenitors, we derived primary human cell cultures from the cortical germinal zone of three different specimens at stages of peak neurogenesis (gestational weeks 15 and 18, GW15 and GW18) (Fig. 3.1b). Before clonal labelling, one of the specimens (GW18) was further dissected on the basis of known anatomical landmarks, allowing us to generate region-specific cultures from the germinal zones of the prefrontal cortex (PFC), primary visual cortex (V1) and medial ganglionic eminence (MGE). We then labelled cells with STICR and cultured them for six weeks in vitro before carrying out scRNA-seq. Transcriptome analysis of 121,290 cells identified three principal

cortical cell type trajectories— excitatory neurons, GABAergic inhibitory neurons and glia—on the basis of differential gene expression, including that of marker genes NEUROD2, DLX2 and GFAP, respectively (Fig. 3.1c, d and Extended Data Figs. 2, 3). We identified intermediate progenitor cells (IPCs) within both the inhibitory and the excitatory neuron trajectories, which we refer to as DLX2+ IPCs (inhibitory trajectory) and EOMES+ IPCs (excitatory trajectory) (Extended Data Fig. 3). Cluster correlation analysis of STICR datasets with an scRNA-seq atlas of the developing primary human cortex at comparable developmental time points and regions14 further supported these cell-type designations (Extended Data Fig. 3.2d). We recovered STICR barcodes in 63 ± 23% (mean ± s.d., n = 5 libraries) of cells per culture, including cells of each principal cell type (Supplementary Table 1 and Extended Data Fig. 3g). In total, we identified 1,461 unique clonal barcodes, 1,324 of which belonged to multicellular clones with a median size of 23 cells per clone (Fig. 3.1e). Although there is very little known about the output of human cortical progenitors over this time scale, we observe a maximum clone size of 1,209 cells, which is congruent with a prior study that measured the output from three individual human outer radial glia15. Notably, no STICR barcodes were shared between experimental groups (Extended Data Fig. 3h), further indicating that cells identified as members of multicellular clones here are not the result of barcode collision. Although all three principal cortical cell types (excitatory glutamatergic neurons, GABAergic interneurons and glia) were found in multicellular clones in both GW15 samples and the GW18 V1 sample, GW18 PFC clones contained relatively more GABAergic inhibitory neurons than excitatory neurons or glia (Extended Data Fig. 3i). Consistent with previous studies, the GW18 MGE culture completely lacked excitatory neurons and instead comprised almost entirely interneurons (Extended Data Fig. 3i). Thus, STICR reveals clonal lineage relationships of all principal cell types derived from human forebrain progenitors.

Clonal relationship of human cortical neurons

In order to determine the lineage relationships of cells born from cortical progenitors, we first analysed the relative proportions of principal cell types in multicellular clones. The majority (66%; 829 of 1,252) of multicellular clones from cortical cultures contained at least 1 excitatory neuron, with these clones containing a median of 3 excitatory neurons per clone (Extended Data Fig. 3.4a, b and Supplementary Table 1). Although our study of regional cortical progenitors is limited to a single GW18 sample, we found that PFC clones contained proportionally Clonal relationship of human cortical neurons In order to determine the lineage relationships of cells born from cortical progenitors, we first analysed the relative proportions of principal cell types in multicellular clones. The majority (66%; 829 of 1,252) of multicellular clones from cortical cultures contained at least 1 excitatory neuron, with these clones containing a median of 3 excitatory neurons per clone (Extended Data Fig. 3.4a, b and Supplementary Table 1). Although our study of regional cortical progenitors is limited to a single GW18 sample, we found that PFC clones contained proportionally fewer excitatory neurons than clones in the V1 sample (Extended Data Fig. 3.4b, c). This is consistent with the neurodevelopmental gradient in the cortex, where developmental milestones are reached in the PFC several weeks before the V1 region16. Across all samples, only 1.5% (19 of 1,252) of cortical clones were composed entirely of excitatory neurons, and 51% (635 of 1,252) contained a combination of excitatory neurons, inhibitory neurons and glia (Fig. 3.2a–c, Extended Data Fig. 3.4c and Supplementary Table 1). Owing to extensive aggregation of excitatory neurons within in vitro cultures and the gentle dissociation used (Extended Data Fig. 3.4d, e), production of excitatory neurons may be underestimated in the dataset derived from cell culture experiments. Nevertheless, across GW15 samples, the coincidence of excitatory and inhibitory neurons within the same clone occurred across the range of clone sizes (Fig. 3.2b). Together, these results suggest that individual human cortical progenitors reproducibly generate both excitatory and inhibitory neurons.

Analysis of cortically born inhibitory neurons

We next sought to determine the transcriptional identities of cortically born GABAergic neurons. In addition to DLX2 (Fig. 3.1d), inhibitory-trajectory cells were broadly enriched for general markers of interneurons, including GAD1, ARX, SLC32A1 and DLX6AS-1 (Extended Data Fig. 3e). Iterative subclustering and transcriptional trajectory analysis of the inhibitory cells along with the DLX2+ IPCs revealed three distinct subgroups of GABAergic inhibitory neurons that we termed IN.1, IN.2 and IN.3 (Fig. 3.2d and Extended Data Fig. 3.5a–c). IN.1 cells were enriched for markers of SST+ cortical interneurons, including SST, NPY, TAC3 and NXPH1 (Fig. 3.2e). Consistent with developmental studies in mice which showed that SST+ cortical interneurons derive primarily from the MGE[17], 73% (492 of 671 cells) of IN.1 cells are produced by MGE progenitors (Extended Data Fig. 3.5d). By contrast, IN.1 cells comprised only 0.3% (179 of 56,244 cells) of progeny born from cortical progenitors (Extended Data Fig. 3.5e). Furthermore, MGE-derived IN.1 cells expressed canonical MGE-born interneuron genes, including LHX6, NKX2-1, ACKR3 (CXCR7), PDE1A and MAF, while cortically born IN.1 cells did not (Extended Data Fig. 3.5f). Together, these data suggest that IN.1 cells are transcriptionally similar to SST+ cortical interneurons and that the majority of IN.1 cells derive from the MGE. In contrast to IN.1 trajectory cells, IN.2 and IN.3 trajectory cells were transcriptionally similar to cells born from the caudal ganglionic eminence (CGE) based on their expression of marker genes such as SCGN, SP8, PCDH9 and BTG1 (Fig. 3.2e)[18]. Furthermore, IN.2 and IN.3 cells differed from IN.1 cells in that they were derived entirely from cortical progenitors, with no contribution from MGE progenitors (Extended Data Fig. 3.5d). Top IN.2 markers included TSHZ1, PBX3, MEIS2, CALB2, CDCA7L, SYNPR and ETV1, which are enriched in mouse olfactory bulb interneurons (Fig. 3.2f)[19,20]. By contrast, top IN.3 marker genes included NR2F1, NFIX, PROX1 and NR2F2, which are enriched within the CGE, as well as SOX6 and CXCR4, which are enriched in cortical interneurons (Fig. 3.2i), suggesting that these cells are transcriptionally similar to CGE-derived cortical interneurons[14,21–25]. Comparison of IN.2 and IN.3 marker genes with

orthogonal datasets, including those from the Allen Brain Atlas[26], similarly support this distinction (Extended Data Fig. 3.5g, h). Thus, although there are currently no marker genes that can unequivocally distinguish cortical interneurons from olfactory bulb interneurons, our transcriptome-wide data suggest that IN.2 cells resemble olfactory bulb interneurons while IN.3 cells are similar to CGE-born cortical interneurons. Previous fate-tracing studies in mice have shown that cortical progenitors can produce a subset of embryonically born olfactory bulb interneurons[11]. In agreement with these findings, we found that multicellular cortical clones frequently contain olfactory-bulb-like (IN.2 branch) GABAergic inhibitory neurons (Extended Data Fig. 3.5d, e). Notably, many clones containing olfactory-bulb-like GABAergic inhibitory neurons also contained excitatory neurons (61%; 321 of 530 clones) (Extended Data Fig. 3.5j). Such clones were found in all cortical samples but were especially common in both GW15 samples (Supplementary Table 1). GW18 PFC clones contained the highest proportion of olfactory-bulb-like GABAergic neurons (41%; 1,637 of 3,949 cells) out of any sample analysed, and although relatively few of the GW18 PFC clones contained excitatory neurons (9 of 141 clones), 8 of these also contained at least one olfactory-bulb-like interneuron (Extended Data Fig. 3.5d and Supplementary Table 1). These results suggest that, similar to mice13, human cortical progenitors can generate olfactory bulb interneurons during embryogenesis, and that some excitatory neurons and embryonically born olfactory bulb interneurons are clonally related. Surprisingly, we found that most (79%; 655 of 829) multicellular clones that contained excitatory neurons also included putative cortical interneurons (IN.3 cells) (Fig. 3.2g), indicating that some human cortical progenitors can generate both excitatory neurons and cortical interneurons. Many of these clones contained multiple cells of both types, and mixed clones were especially abundant in GW15 and GW18 V1 samples. Further subclustering analysis of excitatory neurons revealed both deep-like and upper-like excitatory neuron subgroups, both of which were frequently clonally related to IN.3 cells (Extended Data Fig. 3.6). Notably, every cortical sample we analysed contained clones with both excitatory and inhibitory cortical neurons (Fig. 3.2g).

These mixed clones also frequently contained olfactory bulb interneurons as well as glial-trajectory cells (Extended Data Fig. 3.5i, j). Taken together, our results suggest that human cortical progenitors cultured in vitro are unexpectedly multipotent in their ability to generate a wide variety of principal neural cell types, including both excitatory neurons and putative cortical interneurons—two cell types previously thought to be produced by different pools of spatially restricted progenitors in the developing forebrain.

Mixed excitatory–inhibitory clones in xenografts

To confirm that the observed lineage relationship between excitatory neurons and cortical interneurons was not an artefact of in vitro culture, we STICR-labelled cortical germinal zone cells from three additional GW15 specimens as above, and transplanted them into the cortex of early postnatal immunodeficient mice, where we allowed them to develop for six weeks before analysis (Fig. 3a). Immunohistochemistry revealed xenografted human cells at the injection site itself as well as distributed throughout the adjacent tissue (Figs. 3b and Extended Data Fig. 3.7a–c). To quantify the proportion of the principal cell types in the xenografts, we carried out immunohistochemistry for the excitatory-neuron marker NEUROD2, the interneuron marker GABA, and the glia markers SOX9 and OLIG2. Excitatory neurons were the most common cell type, accounting for 75.7 ± 8.4% of xenograft-derived cells, followed by glia at 16.8 ± 6.3%, and interneurons at 7.9 ± 3.2% (mean ± s.d.; n = 7 recipient mice; Extended Data Figs. 7a–c). In parallel, we isolated EGFP+ xenograft cells by fluorescence-activated cell sorting (FACS), and carried out scRNA-seq to determine their transcriptional identities and clonal relationships. Similar to our in vitro cultures, xenograft cells formed distinct transcriptional clusters of GABAergic neurons, excitatory neurons and glia (Fig. 3c and Extended Data Fig. 3.7d–f). The proportions of excitatory and inhibitory neurons within multicellular clones from xenograft experiments differed from those observed in our in vitro cultures (Fig. 3.2g). This was likely to be due to the specific loss of excitatory neurons from in vitro culture during dissociation, as

these cells frequently associated in tightly formed masses that were difficult to dissociate (Extended Data Figs. 4d, e). Quantification of principal cell types by scRNA-seq closely matched the cell proportions observed by immunohistochemistry, and was reproducible across biological replicates (Extended Data Fig. 3.7g). These data suggest that xenograft-derived STICR-labelled cells analysed by scRNA-seq accurately reflected the cell proportions produced within the xenograft. We recovered STICR barcodes from 76 ± 10% (n = 3 libraries) of xenografted cells and identified 660 multicellular clones that ranged in size from 2 to 101 cells, with a median clone size of 5 cells (Fig. 3d). Within multicellular clones, the proportions of principal cortical cell types were highly similar across biological samples (Fig. 3e). We then further analysed cells categorized as either inhibitory neurons or DLX2+ IPCs in order to determine their transcriptional identities and clonal relationships. Subclustering analysis revealed several distinct groups of GABAergic cells (Extended Data Fig. 3.8a), including one that was enriched for genes found in cortical interneuron-like IN.3 cells— such as NR2F1, KLHL35, NFIX and SCGN—and was similar to reference GABAergic cortical inhibitory neurons (Extended Data Fig. 3.8b, c). In order to directly compare the transcriptomic identities of GABAergic neurons from xenografts to those observed following in vitro culture, we integrated scRNA-seq data of GABAergic neurons from both sets of experiments, using the in vitro cultured cells as a reference. GABAergic neurons from the xenografts clustered well with their in vitro counterparts (Fig. 3f, Extended Data Fig. 3.8d) and integrated entirely within the previously defined transcriptional trajectories (Fig. 3.2f). Consistent with their marker-gene expression, 85% (211 of 249) of GABAergic xenograft cluster-1 cells clustered within the IN.3 trajectory (Extended Data Fig. 3.8d), suggesting that they had the transcriptional identity of cortical interneurons. Few (4%; 9 of 249 cells) GABAergic xenograft cluster-1 cells clustered within the IN.2 trajectory (Extended Data Fig. 3.8d), consistent with previous transplantation studies which found that the cortical environment is not conducive to the generation of olfactory bulb interneurons27. Xenograft-derived GABAergic IN.3 neurons were found in 56 multicellular

clones, 43 (77%) of which also contained excitatory neurons (Fig. 3g, Extended Data Fig. 3.8e–

g and Supplementary Table 2). Mixed IN.3/excitatory-neuron clones were found in all three

xenografted samples, including GW15 replicate 5, which was transplanted immediately after

transduction with STICR and was never cultured (see Methods). Thus, as we had previously

observed in vitro, cortical progenitors can generate mixed excitatory/ inhibitory-neuron clones

when xenografted into the perinatal mouse brain, and these inhibitory neurons are

transcriptionally similar to cortical interneurons. Finally, to determine whether human cortical-

progenitor-derived GABAergic inhibitory neurons had additional features consistent with cortical

interneurons, we characterized their morphology and distribution in two additional xenograft

experiments from GW17 samples. For these experiments, we used a FACS-based isolation

strategy to enrich for human progenitor cells before transducing them with STICR (Fig. 3.4a,

Extended Data Fig. 3.9). Given the protracted time scale of human brain development, we

waited for 12 weeks after transplantation before analysing the xenografts by

immunohistochemistry. We observed STICR-labelled cells with well-elaborated

processes distributed throughout the cortical layers and also within the olfactory bulbs (Fig. 3.4b

and Extended Data Fig. 3.9). Of STICR-labelled cells located in the cortex, 8.3% (89 of 1,071)

expressed GABA, and these were also broadly dispersed across the cortical layers (Fig. 3.4c, d

and Extended Data Fig. 3.10). Although most STICR-labelled GABA+ cells were located at a

similar rostral–caudal level as the transplantation site, we did observe cells more distant from

the transplantation site in both directions. Together, these morphological characteristics of

cortically derived GABAergic inhibitory neurons are consistent with their classification as cortical

interneurons. Thus, individual human cortical progenitors can give rise to both excitatory

neurons and cortical GABAergic inhibitory neurons (Fig. 3.4e).

Discussion

By carrying out high-throughput lineage tracing of approximately 1,900 human cortical progenitors, we have been able to demonstrate a lineage relationship between human excitatory neurons and putative cortical interneurons—two populations that had been widely believed to arise from distinct progenitor populations. Although existing scRNA-seq informatic tools have been used to describe the transcriptional trajectories between progenitors and mature cell types (that is, pseudotime analysis[28]), extensive analysis of the developing human brain at this time has not uncovered this lineage relationship[29], nor would further analysis be likely do so given the degree to which the transcriptional trajectories of these cell types differ. Furthermore, pseudotime analysis itself does not provide direct evidence of clonal relationships, and the lineages of transcriptionally similar progenitors can differ. At present, approaches that retrospectively infer developmental lineage relationships from somatic mutations predicted to have arisen during development[30–32] lack the resolution to confidently assign daughter cells to individual progenitors, and thus would not have revealed this relationship.

Previous studies suggested that cortical progenitors generate cortical interneurons in primates, but were limited by their use of a small set of putative 'marker' genes that do not directly link progenitors to their progeny[6,7]. In these studies, differences in the expression of GABAergic marker genes between different germinal zones were presumed to persist across differentiation from progenitors to inhibitory neurons. These markers were thus used to retrospectively infer the origin of cortical interneurons. However, this type of approach cannot account for potential changes in gene expression that can occur over development, or for potentially unknown sources of cells expressing the same putative 'marker' genes. Additionally, although one study identified cortically born interneurons using a gammaretorviral labelling approach[7], their approach did not determine whether these cells were cortical interneurons or olfactory bulb interneurons, which our work and others have found to be produced from progenitors in the cortex[11–13]. Our prospective approach of clonal labelling by STICR combined with scRNA-seq-

based transcriptome-wide analysis has allowed us to distinguish between these two types of interneurons, and we have found, surprisingly, that both types can be clonally related to excitatory neurons (Figs. 2–4). As with any prospective labelling study carried out in humans, we cannot definitively rule out the possibility that our methodologies influenced progenitor behaviour. However, the observation of clonally related excitatory and inhibitory neurons in all of our experimental contexts provides strong evidence that human cortical progenitors have the capacity to generate both types of neuron.

In this study, we investigated the clonal relationships of cells born from GW15–GW18 human cortical progenitors over a six-week developmental window. A previous study in which newborn neurons were labelled with 5-bromo-2′-deoxyuridine (BrdU)[8] quantified BrdU+ ;DLX2+ cells in organotypic slice culture derived from GW17.5–GW20.5 human cortex. After roughly eight to ten days of BrdU labelling, that work[8] did not find a substantial amount of BrdU+ ;DLX2+ cells from cortical progenitors. Given the longer window of our experiments, these data suggest that the production of interneurons by cortical progenitors does not occur throughout the entire duration of corticogenesis, but instead begins at some point after midgestation. Consistent with this interpretation, six weeks after labelling we observed a higher proportion of inhibitory neurons within clones derived from GW18 progenitors than within those from GW15 progenitors (Supplementary Table 1). Furthermore, extensive chains of migrating interneurons have been observed in the perinatal human cortex[33]. Thus, production of cortical interneurons from cortical progenitors may extend beyond the period in which excitatory neurons are born. It is also possible that the different experimental systems used in these studies (that is, in vitro cell culture and xenograft here, and ex vivo organotypic slice culture in ref. 8) might contribute to some of the observed incongruencies between these three systems. Future studies aimed at detailing the output of cortical progenitors over a broader developmental period and throughout different cortical regions will help to further elucidate the contribution of this phenomenon to human brain development.

In a companion study[34], STICR was used to perform in vivo clonal labelling of embryonic mouse forebrain progenitors, and STICR-labelled cells were analysed postnatally by scRNA-seq. Although both glutamatergic excitatory neurons and cortical GABAergic inhibitory neurons were recovered, they did not occur within the same clone. The lineage relationship that we observe between cortical excitatory and inhibitory neurons herein thus raises new questions regarding the development of the human cerebral cortex. First, what are the implications of a single progenitor producing both excitatory neurons and cortical interneurons? Evolutionary expansion of the primate neocortex has been attributed to the increased proliferative capacity of cortical neural progenitors. Adaptations in cortical progenitor competence to produce both principal types of cortical neuron could help to ensure the appropriate inhibitory/excitatory balance, despite the dramatic increase in the pool of cortical excitatory neurons[35]. Recent studies have revealed that although the inhibitory/excitatory balance increases from mice to humans, the relative composition of cortical interneuron types remains relatively constant across evolution. Although existing technical limitations prevent us from confidently estimating the precise cellular contributions of cortical progenitors to the mature human brain, future studies that quantify the relative contributions of progenitors from the cortex and ganglionic eminences will be helpful in understanding the cellular basis of how normal human cortical function is achieved.

Our study opens many avenues for future investigation. The molecular mechanisms that regulate the production of locally born cortical inhibitory neurons are unknown at present. In mice, signalling through the Sonic hedgehog pathway[36] is required for individual cortical progenitors to undergo a GABAergic 'switch' and to generate inhibitory neurons[37] that migrate to the olfactory bulb[13]. Does a similar molecular mechanism govern the production of cortically-derived cortical interneurons in humans? Furthermore, what molecular markers, if any, can distinguish them from cortical interneurons born in the CGE or MGE? Previous studies have found that NR2F1 and NR2F2 are expressed not only in the CGE but also in cortical progenitors[9,10]. Given the transcriptional similarity of cortically-derived IN.3 cells to CGE-derived

interneurons observed here, it is possible that a similar developmental program is used. These are just some of the questions raised by our new understanding of the human cortical lineage, and addressing these will help to further decipher the origins and mechanisms underlying human brain development.
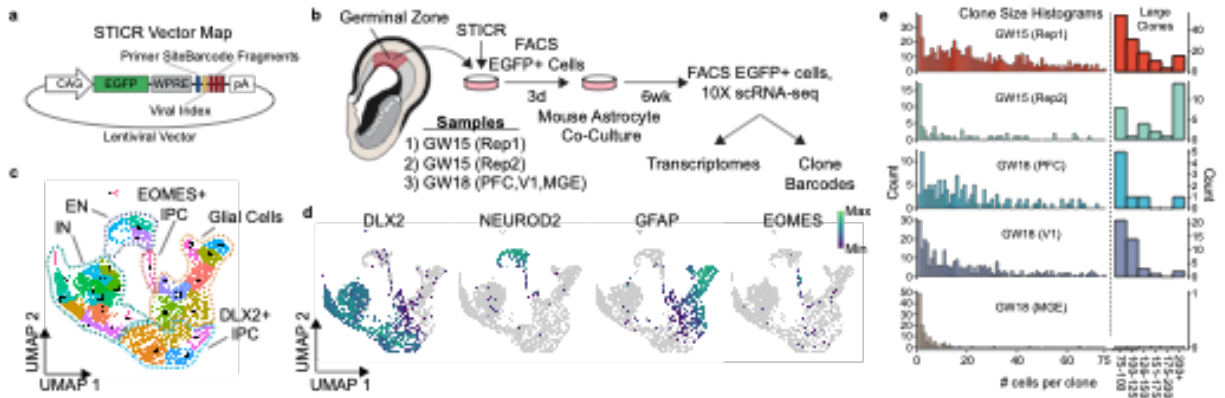
FIGURES

Figure 3.1 | STICR-labelled progenitors generate all three principal cortical cell types. a, Design of the STICR lentiviral vector. CAG, cytomegalovirus/chicken β-actin promoter; WPRE, woodchuck hepatitis virus post-transcriptional regulatory element; pA, polyadenylation signal. **b,** The experimental design used to label and capture samples. GW, gestational week; Rep, replicate. **c,** Uniform manifold approximation and projection (UMAP) embedding and Leiden clustering of STICR-labelled cells following scRNA-seq. IPC, intermediate progenitor cell. **d,** Feature plots showing the principal-cell-trajectory marker genes DLX2 (GABAergic inhibitory neurons and DLX2+ IPCs), NEUROD2 (excitatory neurons), GFAP (glia), and EOMES (EOMES+ IPCs). **e,** Histogram showing clone sizes within each sample. Left, clone sizes of 1–75 cells. Right, clone sizes of more than 75 cells, in 25-cell bins.
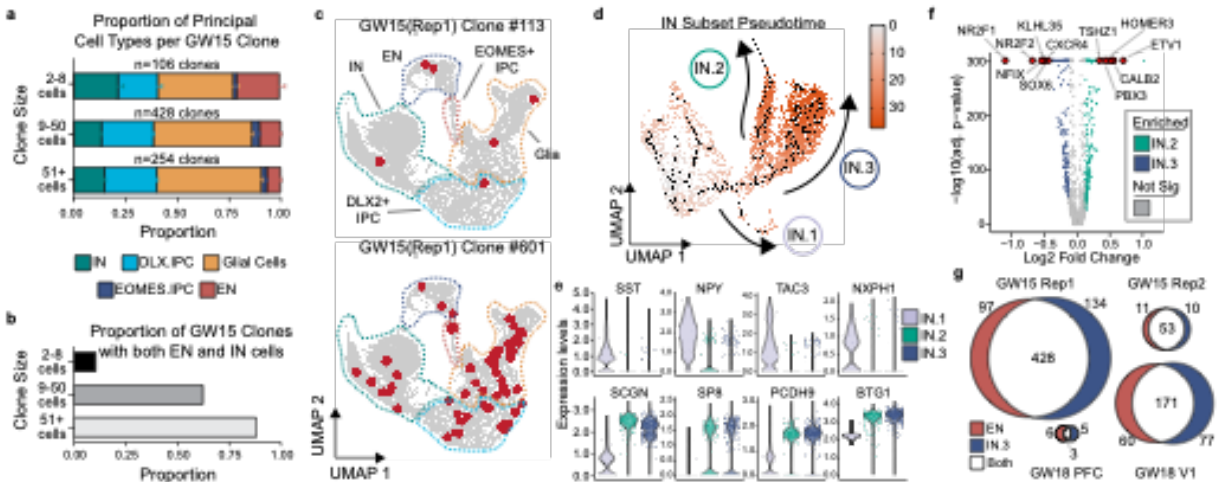
Figure 3.2 | Individual human cortical progenitors can generate both excitatory and inhibitory cortical neurons in vitro. a, Stacked boxplot depicting the average (mean ± s.d.) proportion of GABAergic inhibitory neurons, DLX2+ IPCs, EOMES+ IPCs, excitatory neurons and glia in different sized clones. **b,** Bar graph depicting the proportions of clones of different sizes that contain both excitatory and inhibitory neurons. **c,** Representative 6-cell (GW15, replicate 1, clone 113) and 61-cell (GW15, replicate 1, clone 732) clones depicted in UMAP space. Cells within each clone are coloured in red. Dashed lines depict borders of principal cell types from Fig. 3.1c. **d,** Pseudotime transcriptional trajectories of subclustered inhibitory neurons. Three different interneuron transcriptional trajectories (IN.1, IN.2 and IN.3) are indicated with arrows. **e,** Violin plots depicting expression of IN.1 marker genes (SST, NPY, TAC3 and NXPH1) and general CGE marker genes (SCGN, SP8, PCDH9 and BTG1) in IN.1, IN.2 and IN.3 trajectories. **f,** Volcano plot comparing differences in gene expression between IN.2 and IN.3 trajectory cells. Not sig, genes with log2 fold-change less than 0.125 and/or adjusted P value greater than 0.1. **g,** Venn diagram depicting the number of multicellular cortical clones that contain IN.3 cells or excitatory neurons.
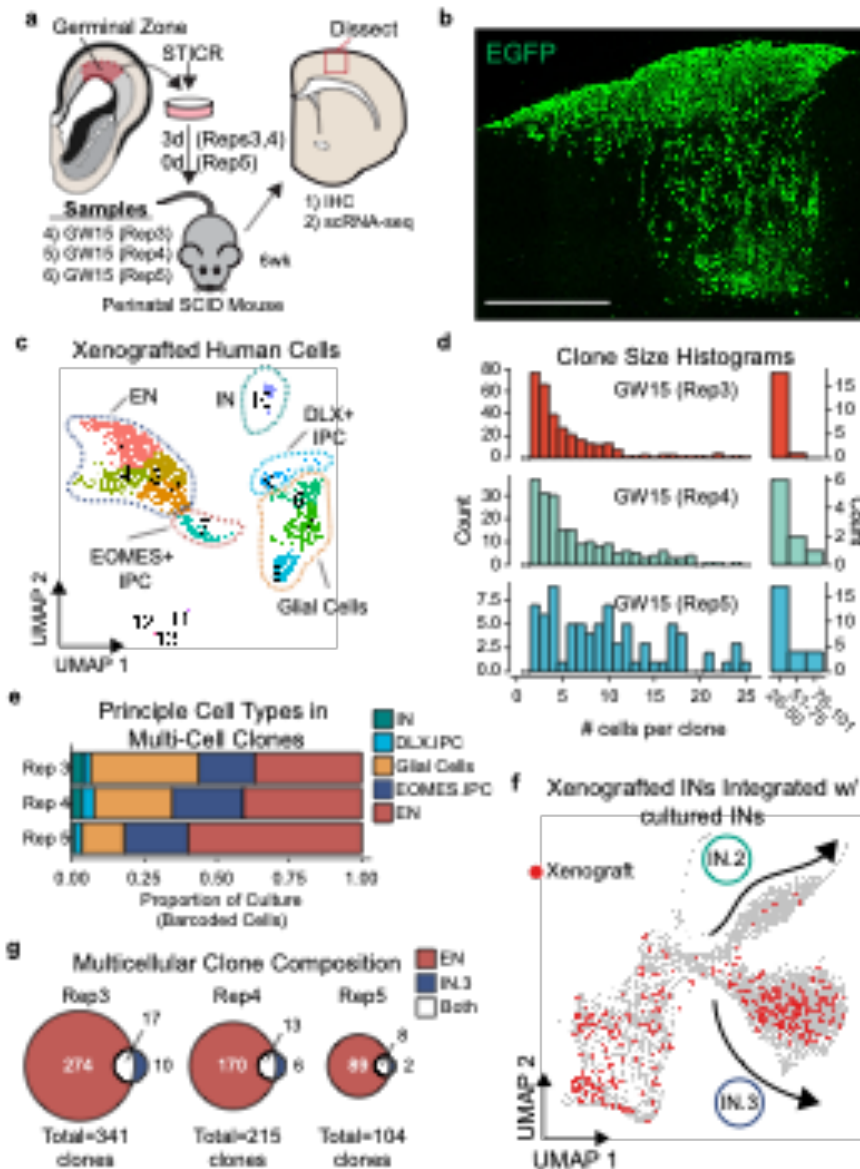
Figure 3.3 | Xenografted human cortical progenitors generate both excitatory and inhibitory cortical neurons in the same clone. a, Experimental design and analysis of STICR-labelled progenitors by immunohistochemistry (IHC) and scRNA-seq following transplantation into the postnatal murine cortex. **b,** Representative image of transplanted human cortical cells. EGFP expression from STICR is in green. Scale bar, 500 μm. **c,** UMAP embedding and Leiden clustering of xenografted cells following scRNA-seq. **d,** Histogram of clone sizes within each xenograft sample. Left, clone sizes from 1–25 cells. Right, clone sizes of greater than 25 cells, in 25-cell bins. **e,** Stacked barplot depicting the relative proportions of principal cell types within multicellular clones of each sample. **f,** UMAP embedding of both cultured and xenograft-derived inhibitory neurons. Xenograft-derived cells that are members of multicellular clones are highlighted in red. IN.2 and IN.3 trajectories are depicted with arrows. g, Venn diagrams showing the number of multicellular clones containing excitatory neurons and/or IN.3 neurons.
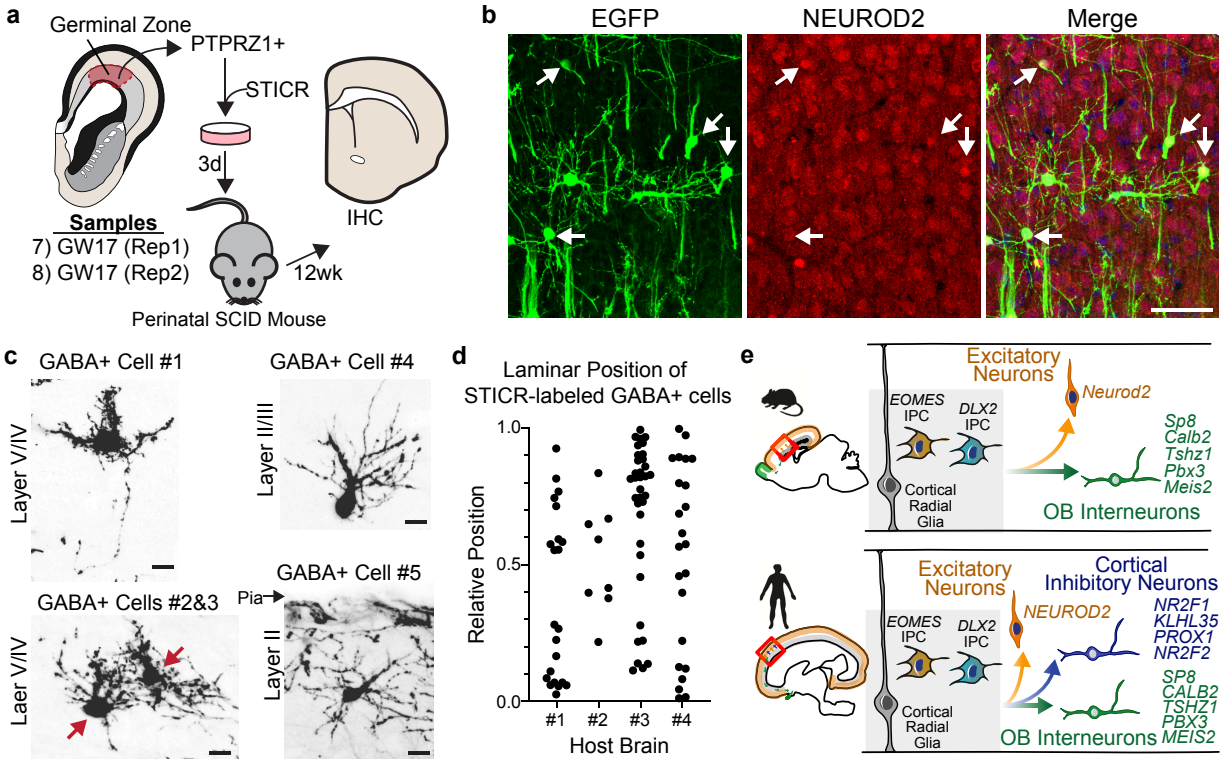
Figure 3.4 | Xenografted human cortical progenitors generate GABAergic inhibitory neurons that distribute across the cortical laminae. a, Experimental design and analysis of PTPRZ1-enriched, STICR-labelled progenitors by IHC following transplantation into the postnatal murine cortex. **b,** Representative images of transplanted human cortical cells analysed by IHC. EGFP expression from STICR is depicted in green, NEUROD2 in red, and DAPI counterstain in blue. Arrows indicate NEUROD2+ /STICR+ double-positive cells. Scale bar, 50 µm. **c,** Representative images of STICR-labelled GABA+ cells throughout the cortical plate. Red arrows depict soma of two cells in the same image. Scale bars, 10 µm. **d,** Relative laminar positions of GABA+ /STICR+ cells in the cortical plate from the corpus callosum (value of 0.0) to the pial surface (value of 1.0) in each host mouse brain. Host brains 1 and 2 were transplanted with GW17 (replicate 1); host brains 3 and 4 were transplanted with GW17 (replicate 2). **e,** Schematic showing the differences in the developmental potential of human and mouse cortical progenitors. OB, olfactory bulb.

77
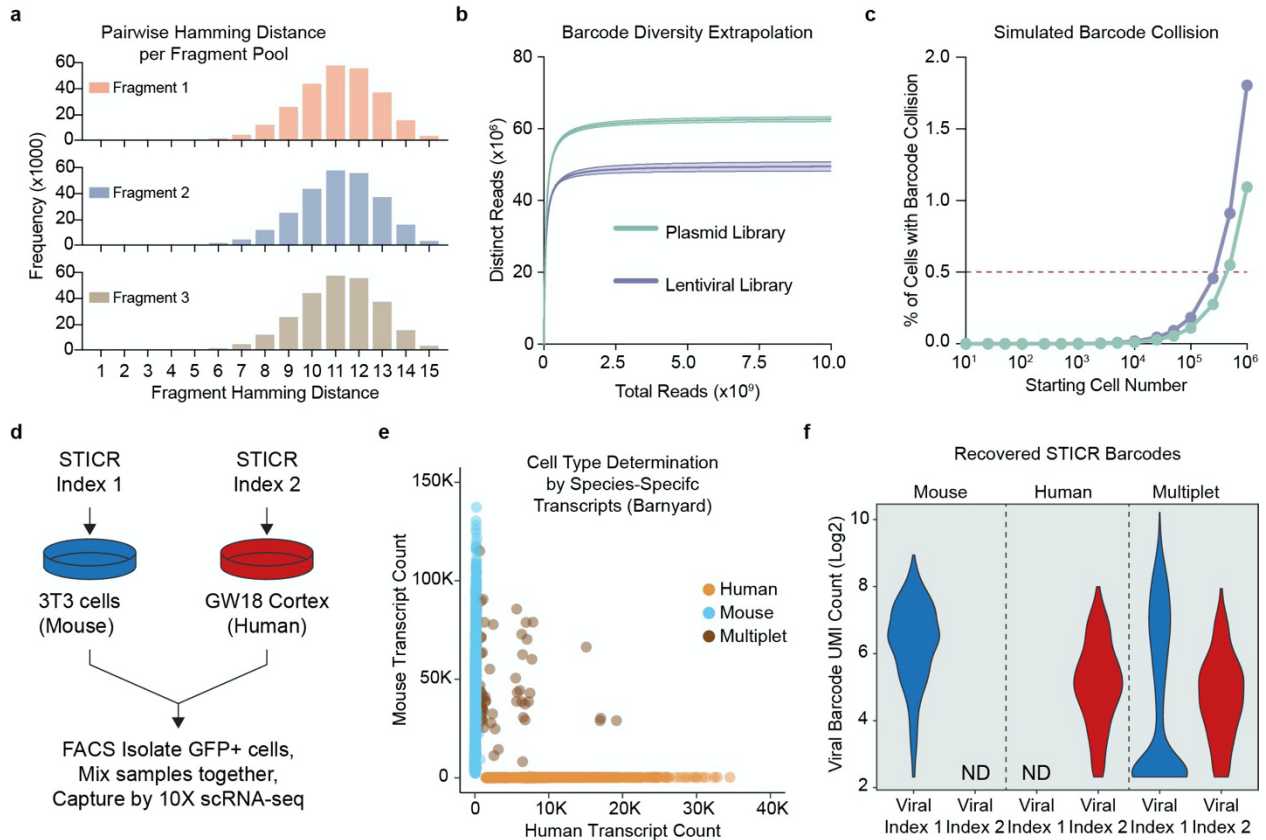
Extended Data Figure 3.1 | Validation of the STICR barcode design. a, Histogram showing pairwise hamming distances between every sequence in each STICR fragment pool. **b,** Barcode diversity extrapolations derived from sequencing a representative STICR plasmid or lentiviral library. Mean ± 95% confidence range for each library is shown. **c,** Simulated barcode collision frequencies (mean ± s.d.) for a range of starting cell numbers, based on the barcode diversity estimated in b. Barcode sampling was performed with replacement using measured proportions of barcodes within the representative plasmid and lentiviral libraries depicted in b. Each simulation was performed 20,000 times. Most error bars (depicting standard deviations) are not visible as they are smaller than the dots (depicting mean values). **d,** The 'barnyard' species-mixing experiment. **e,** Plot depicting species-specific transcript counts from barnyard experiment. Each dot depicts a single cell and the dot colour indicates whether the cell was determined to be a 3T3 cell (mouse), cortex cell (human), or mixed droplet (multiplet). **f,** Violin plots depicting the number of unique STICR barcode molecules recovered from droplets identified as either mouse, human, or multiplet. ND, not detected.
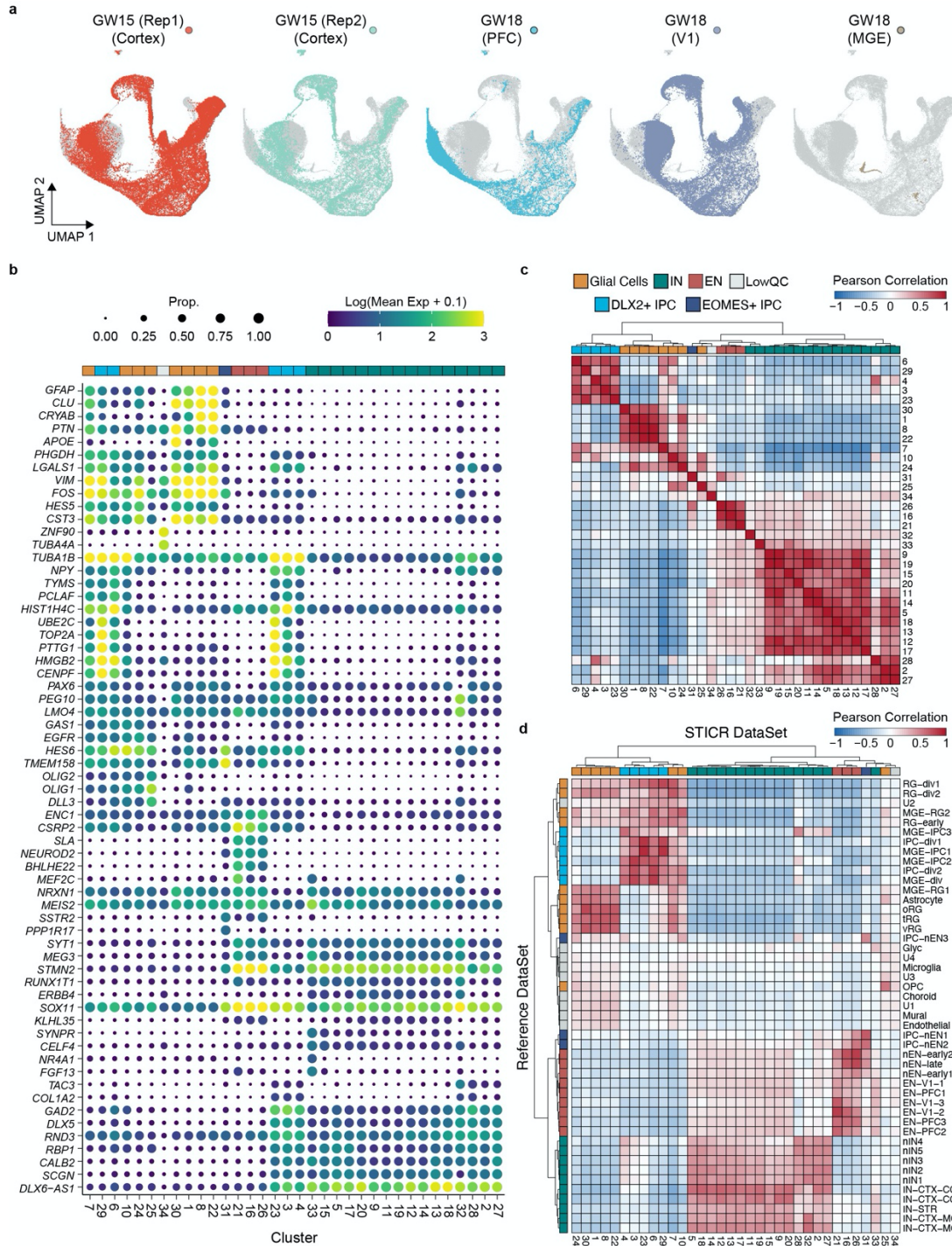
Extended Data Figure 3.2 | Cluster analysis of in vitro STICR data sets. a, UMAP plots of each individual biological sample highlighted. **b,** Top marker-gene expression for each cluster. The size of each dot corresponds to the proportion of cells in the cluster that express the gene, while the colour of the dot corresponds to the average expression level per cluster. **c, d,** Heatmap depicting pairwise transcriptional cluster correlation of in vitro cultured cells with itself (c) and the 2017 Nowakowski scRNA-seq atlas (d) 14. The principal cell-type designation is depicted next to each column and row. The dendrogram depicts hierarchical clustering distance.
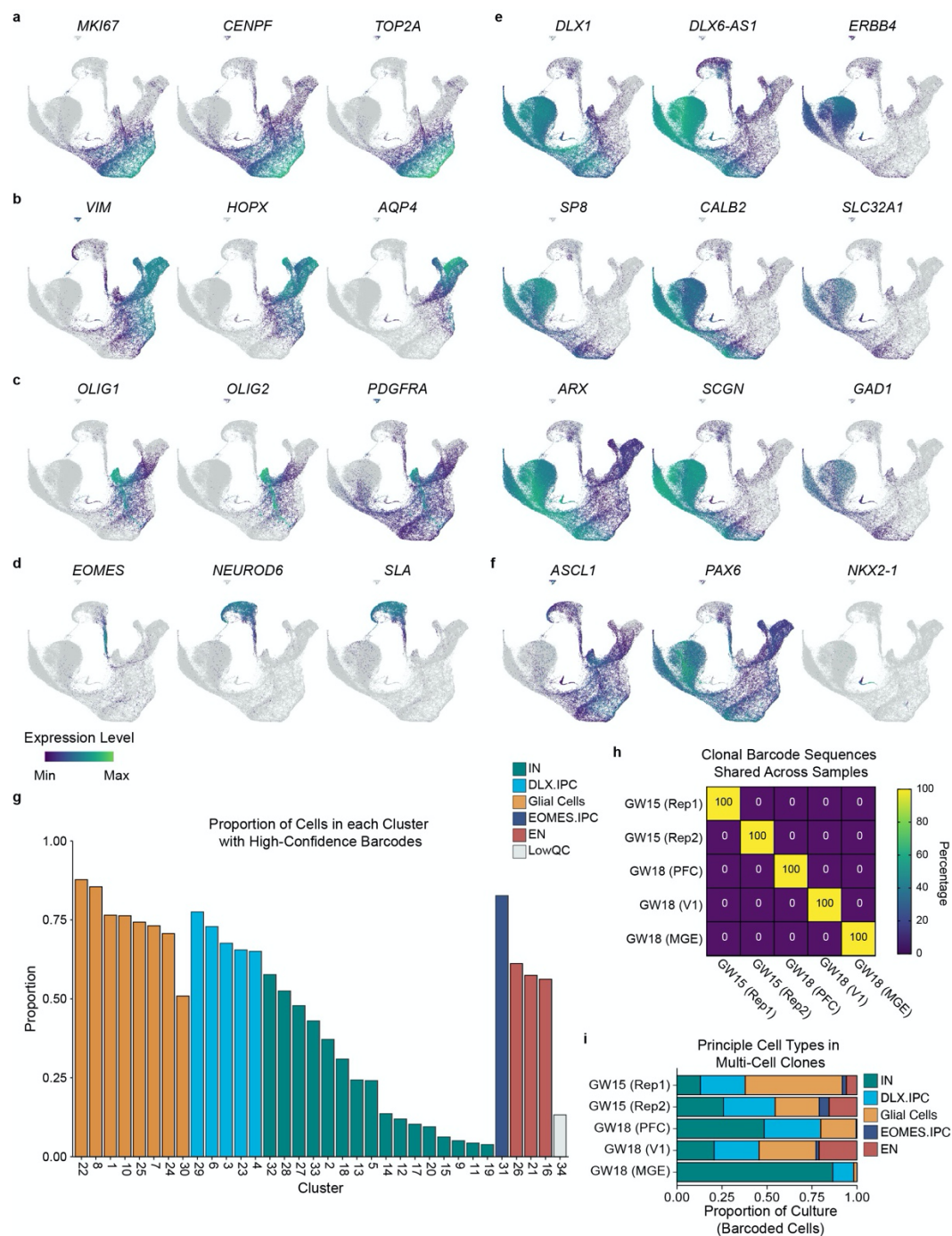
Extended Data Figure 3.3 | Transcriptional analysis of in vitro STICR data sets. a–f, Feature plots depicting expression of genes corresponding to cell cycle (a), glia (b), oligodendroglia (c), excitatory neurons (d), interneurons (e), and regional markers (f). **g,** Bar plot depicting the proportion of cells within each cluster with a recovered STICR barcode. **h,** Heatmap depicting the percentage of STICR barcodes shared between biological samples. GW15 (Rep1) and all GW18 samples were labelled with the same viral stock, while GW15 (Rep2) was labelled with a different stock (see Methods). **i,** Stacked barplot depicting relative proportions of principal cell types within each sample, restricted to cells that are members of multicellular clones.
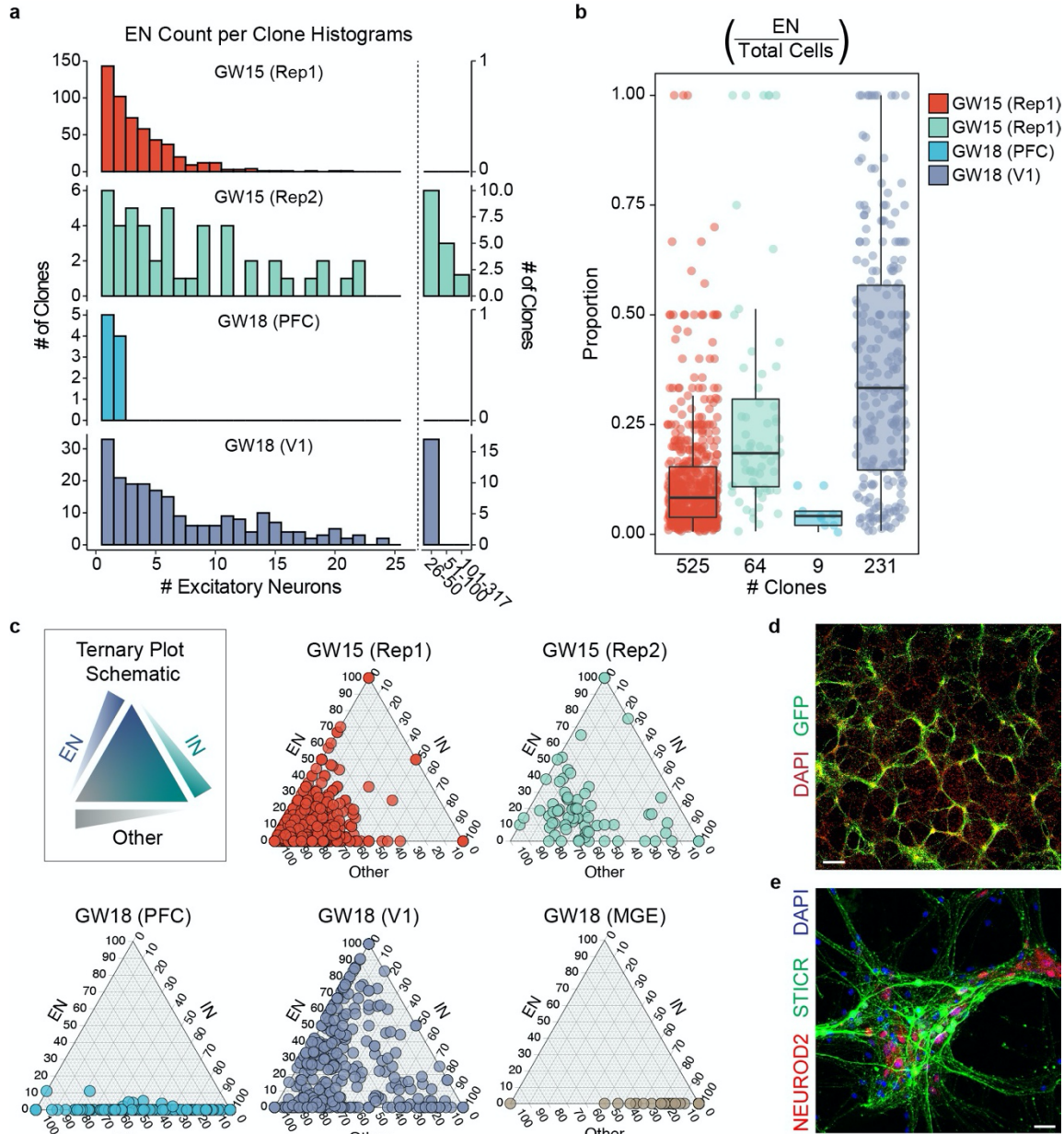
Extended Data Figure 3.4 | Clonal analysis of cortical clones containing excitatory neurons. a, Histogram of excitatory neuron (EN) counts within each multicellular cortical clone. Left, clone sizes from 1–25 cells in single-cell bins. Right, clone sizes of more than 25 cells in the indicated bin sizes. **b,** Box-and-whisker plot depicting the proportion of EN cells within individual multicellular clones for each biological sample. Maxima and minima of boxes depict third and first quartiles, while box centres depict medians. Whiskers depict 1.5× the interquartile distance. Individual clone values are shown as dots. The number of clones is listed below each sample group. **c,** Ternary plots depicting the relative proportions of inhibitory neurons, excitatory neurons and all other cell types ('Other') within individual clones. **d, e,** Immunohistochemistry of in vitro cultures derived from GW15 germinal zone cells labelled with STICR. d, Low-magnification image to show distribution; scale bar, 25 μm. e, High-magnification image showing a cluster of ENs; scale bar, 250 μm.

81

a Inhibitory Neuron Subclusters

b *MKI67* *STMN2* *CENPF* *ERBB4*

c STICR IN/DLX2+ IPC Subset — Cluster Marker Correlation

d IN Branch Composition (Multi-Cell Clones) — GW15_Rep1, GW15_Rep2, GW18_PFC, GW18_V1, GW18_MGE

e Sample Composition (Multi-Cell Clones) — IN.1, IN.2, IN.3, IN.early, DLX.IPC

f MGE-Derived Cells — *NKX2-1*, *LHX6*, *ACKR3*, *MAF*, *PDE1A*; Cortex, MGE

g BrainSpan Laser Capture Microdissection — Z-score; Cortex, OB/RMS, Globus Pallidus, Striatum

h Allen Mouse Brain ISH Atlas — IN.2, IN.3; IN.2 markers: *Pbx3* 0.62, *Tshz1* 0.50, *Cdca7l* 0.55, *Synpr* 0.52; IN.3 markers: *Chl1* 0.39, *Nr2f1* 1.58, *Enc1* 0.50, *Prox1* 0.41 — Expression

i Normalized Clone Composition — GW15 Rep1 (712); GW15 Rep2 (76), GW18 PFC (141), GW18 V1 (323), GW18 MGE (72); EN, IN.1, IN.2, IN.3, Glia, IN.early, DLX.IPC, EOMES.IPC

j Clonal composition of EN-Containing Multicellular Clones — IN.2 (321 Clones), IN.3 (655 Clones), 303; IN.2, IN.3
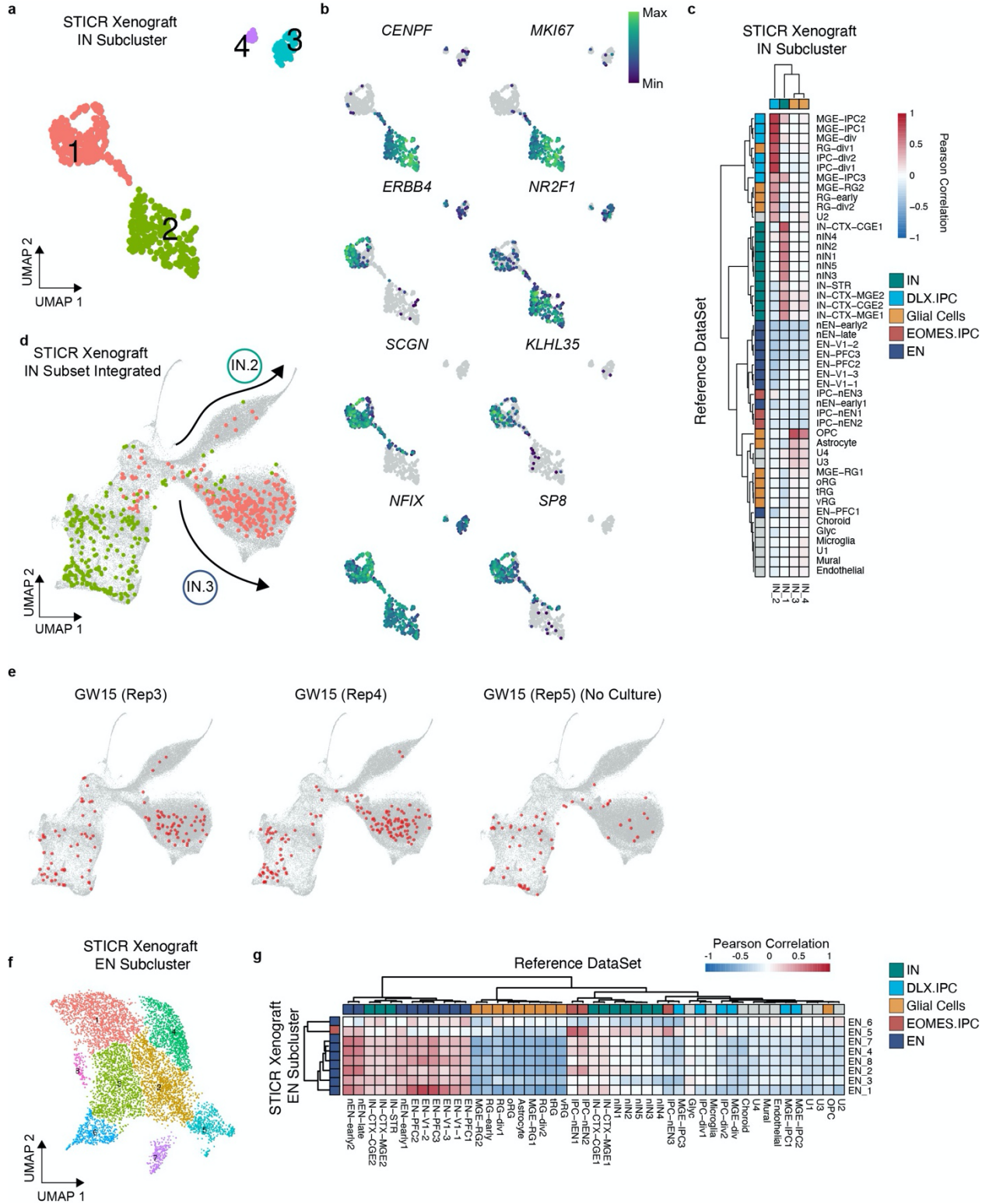
82

Extended Data Figure 3.5 | Clonal and transcriptional analysis of inhibitory neurons and DLX2+ IPCs in vitro. a, UMAP embedding and Leiden subclustering of GABAergic inhibitory neuron (IN) trajectory cells. **b,** Feature plots depicting expression of MKI67, STMN2, CENPF and ERBB4. **c,** Heatmap depicting pairwise transcriptional cluster correlation of this data set with itself. **d,** Stacked barplot depicting relative proportion of multicellular clones from each sample that comprise each IN trajectory. **e,** Stacked barplot depicting the relative proportions of different IN trajectory cells within multicellular clones of each sample. **f,** Feature plots depicting MGE-derived cells (red) and expression of NKX2-1, LHX6, ACKR3, MAF and PDE1A. The enlarged insets below show IN.1 trajectory cells. **g,** Heat plot depicting differential expression of IN.2 and IN.3 marker genes in the developing human cortex, olfactory bulb/rostral migratory stream and basal ganglia. Data are derived from the Allen BrainSpan Laser Capture Microdissection database. Dendrograms reflect hierarchical clustering of genes and samples while colours represent quantile-normalized z-scores. **h,** Paired violin plots and in situ hybridization (ISH) images of P60 mouse brains from the Allen Brain Atlas for select genes. The log2 fold difference between IN.2 (olfactory-bulblike) and IN.3 (cortical-interneuron-like) cells is depicted above each violin plot. **i,** Stacked barplots depicting relative proportions of IN.1, IN.2, IN.3, EN and glia trajectory cells within multicellular clones. The number of clones is listed below each sample. **j,** Venn diagram depicting the number of EN-containing multicellular cortical clones that also contain IN.2 and/or IN.3 cells.
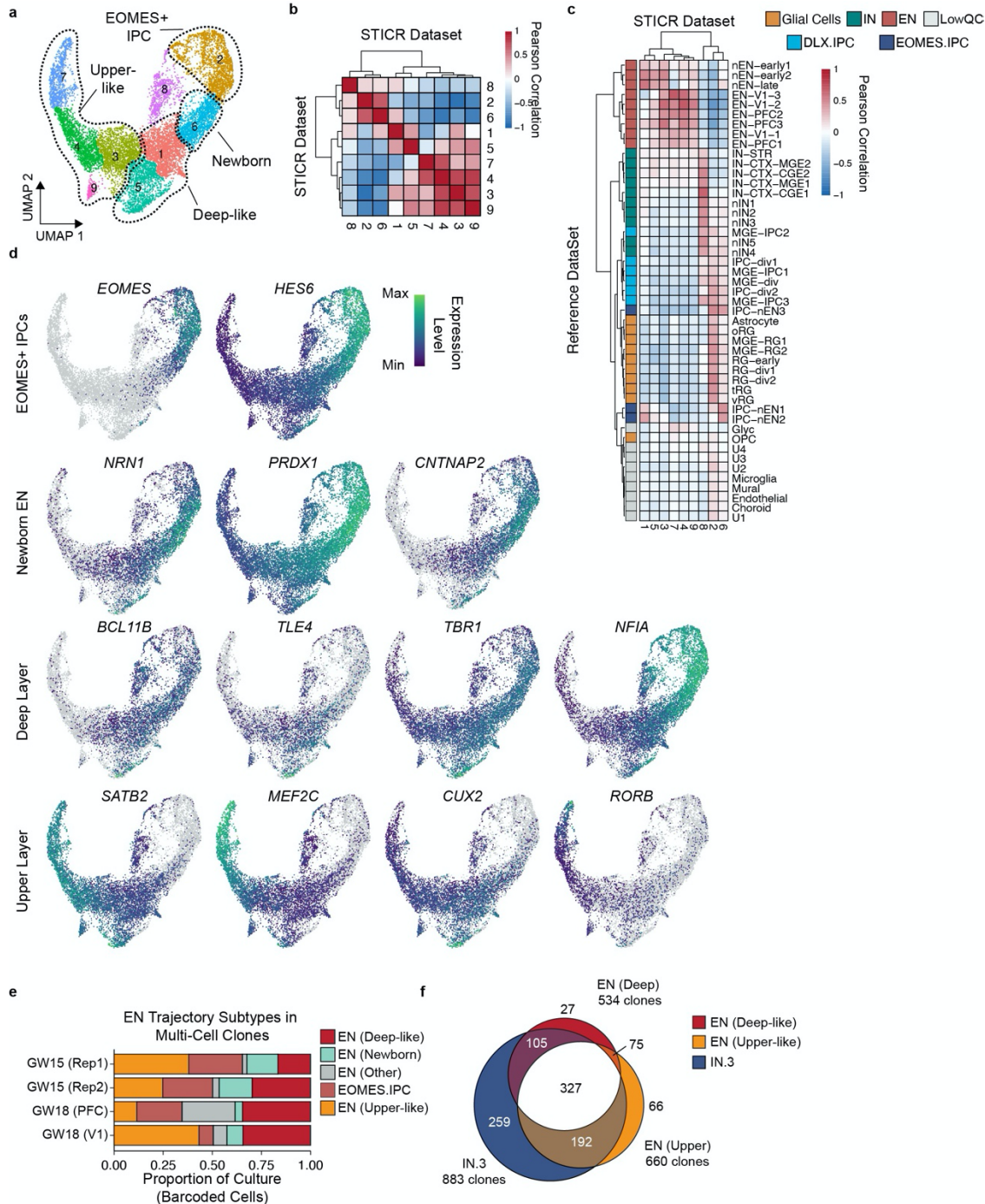
Extended Data Figure 3.6 | Clonal and transcriptional analysis of excitatory neurons and EOMES+ IPCs in vitro. a, UMAP embedding and Leiden subclustering of excitatory neuron (EN) trajectory cells. **b, c,** Heatmap depicting pairwise transcriptional cluster correlation of subclustered EN trajectory cells with self (b) and with the 2017 Nowakowski developing human brain scRNA-seq atlas (c) 14. **d,** Feature plots depicting the expression of genes corresponding to labelled subclustered EN trajectory subtypes. **e,** Stacked barplot depicting relative proportions of EN subtypes within EN trajectory cells of multicellular clones. **f,** Venn diagram depicting the number of multicellular cortical clones containing deep-like ENs, upper-like ENs, and IN. 3 cells.

84

Extended Data Figure 3.7 | Characterization of human cortical progenitor xenografts at six weeks. a, b, Representative images of transplanted human cortical cells analysed by IHC for principal cell-type markers six weeks after transplantation. EGFP expression from STICR is in green, with NEUROD2 or GABA expression in red. Scale bars: a, 50 μm; b,10 μm. **c,** Barplot depicting the proportion (mean ± s.d.) of transplanted cells expressing principal cell-type markers as assessed by IHC. n = 7 sections derived from 6 xenografted mice, 3 of which were transplanted with donor cells from GW15 Rep1 and 3 with cells from GW15 Rep2. **d,** Top marker-gene expression for each cluster from xenografted cells. Sizes of dots correspond to the proportion of cells in the cluster expressing the gene, while dot colours correspond to the average expression level per cluster. **e,** UMAP embedding of xenografted cells and feature plots depicting expression of NEUROD2, EOMES, DLX2, MKI67 and GFAP. **f,** Heatmap depicting pairwise transcriptional cluster correlation of subclustered excitatory neuron (EN) trajectory cells with the 2017 Nowakowski developing human cortex scRNA-seq atlas14. **g,** Comparison of principal celltype quantification (mean) in transplanted cells by analysis method (IHC versus scRNA-seq) and biological replicate.

Extended Data Figure 3.8 | Transcriptional analysis of excitatory and inhibitory neurons from xenografts. a, UMAP embedding and Leiden subclustering of inhibitory neuron (IN) trajectory cells from xenografts. **b,** Feature plots depicting expression of CENPF, MKI67, ERBB4, NR2F1, NFIX, SP8, SCGN and KLHL35. **c,** Heatmap depicting pairwise transcriptional cluster correlation of subclustered xenograft IN and DXL2+ IPC trajectory cells with the 2017 Nowakowski developing human cortex scRNA-seq atlas14. **d,** UMAP embedding depicting cells in multicellular clones from xenograft IN subclusters 1 (salmon) and 2 (lime), integrated with an in vitro cultured STICR IN subset. **e,** UMAP embedding depicting individual interneuron trajectory cells from multicellular clones from xenograft experiments integrated with interneuron trajectory cells from in vitro cultures, split by biological replicate. Members of such clones are highlighted in red. **f,** UMAP embedding and Leiden subclustering of excitatory neuron (EN) and EOMES+ IPC trajectory cells from xenografts. **g,** Heatmap depicting pairwise transcriptional cluster correlation of subclustered xenograft EN and EOMES+ IPC trajectory cells with the 2017 Nowakowski developing human cortex scRNA-seq atlas.
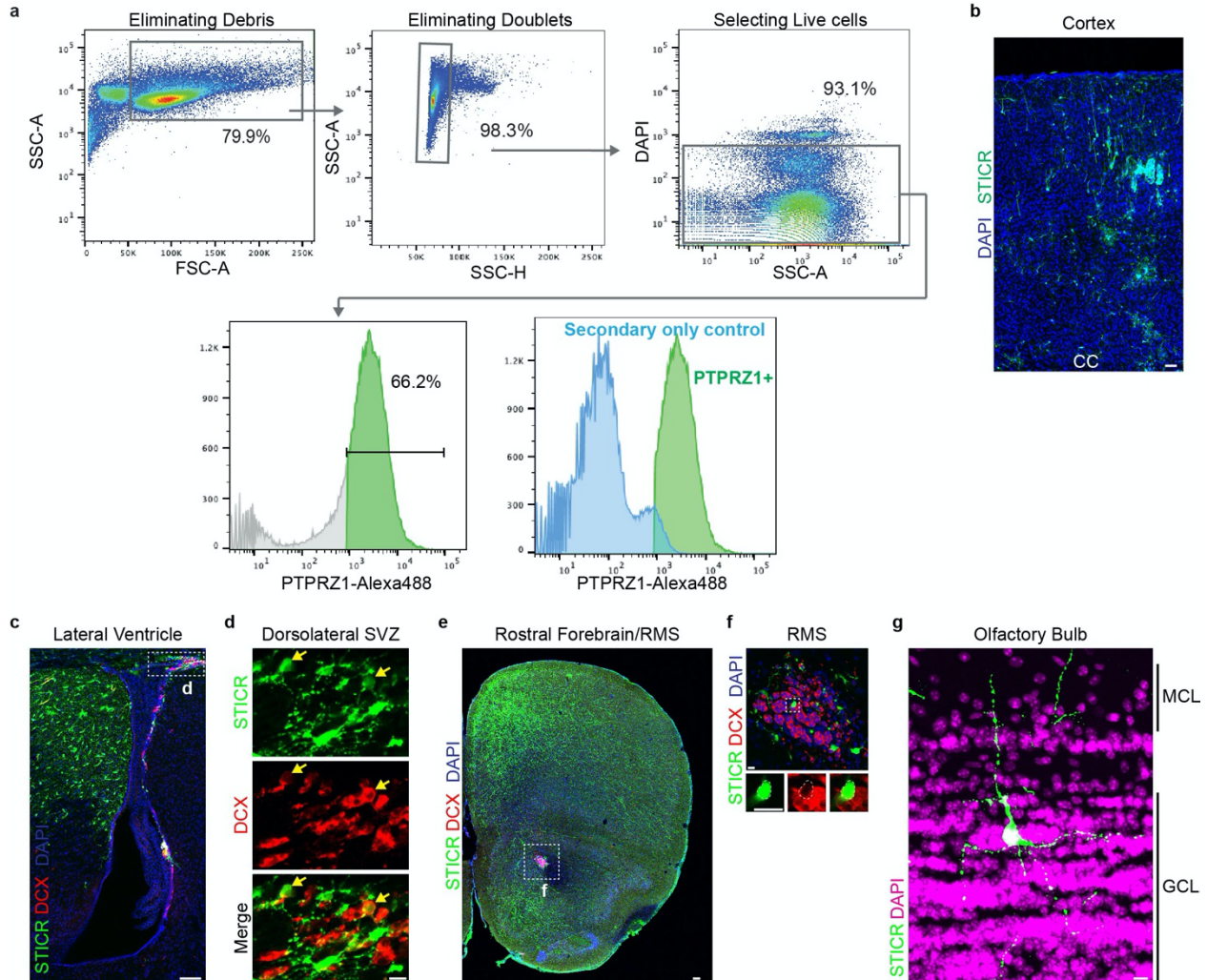
Extended Data Figure 3.9 | Analysis of PTPRZ1-sorted STICR+ cells in the cortex, subventricular zone, rostral migratory stream and olfactory bulb at 12 weeks. a, Representative FACS plots depicting isolation of PTPRZ1+ cells from the cortical germinal zone. **b,** Representative image of transplanted human cortical cells in the cortex of a 12-week-old host mouse. EGFP expression from STICR is in green, with DAPI in blue. Scale bar, 50 μm. CC, corpus callosum. **c,** Representative images of PTPRZ1-sorted, STICR-labelled cells in the dorsolateral corner of the lateral ventricle of a 12-week-old host mouse analysed by IHC. EGFP from STICR is in green, DCX in red, and DAPI in blue. Scale bar, 100 μm. **d,** High-magnification inset of the region boxed in c. Scale bar, 10 μm. SVZ, subventricular zone. **e,** Representative images of PTPRZ1-sorted, STICR-labelled cells in the rostral forebrain, analysed by IHC. The rostral migratory stream (RMS) is outlined by the white box labelled f. GFP expression is in green, DCX expression in red, and DAPI in blue. Scale bar, 100 μm. **f,** High-magnification insets of the RMS depicted in the white box in e. Scale bar, 10 μm. The cell outlined by the white box is magnified below. **g,** Representative images of PTPRZ1-sorted, STICR-labelled cells that migrated from the transplantation site in the cortex to the olfactory bulb, analysed by IHC. EGFP expression is in green and DAPI in magenta. GCL, granule cell layer; MCL, mitral cell layer. Scale bar, 10 μm.
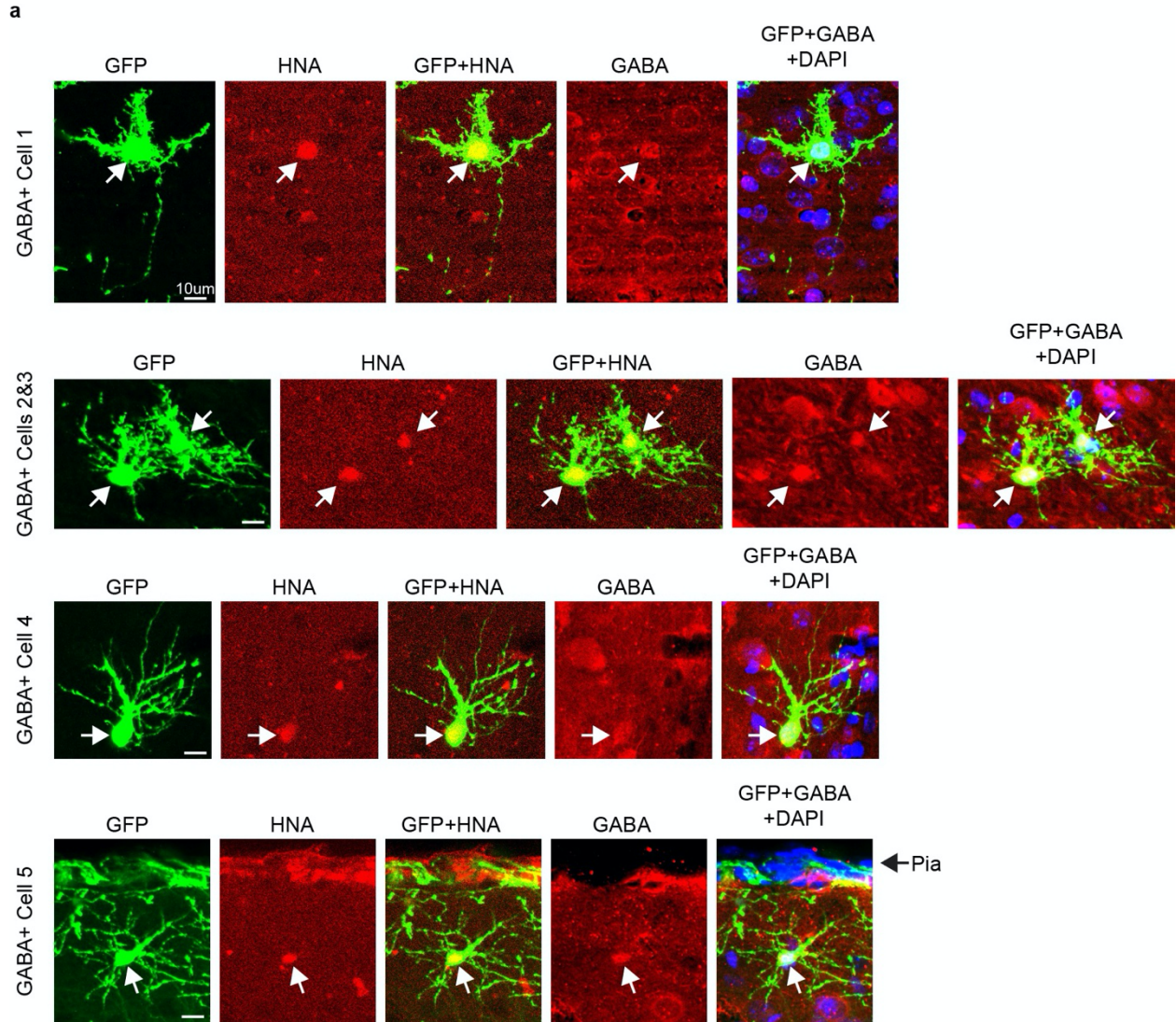
Extended Data Figure 3.10 | Immunohistochemistry of STICR-labelled cortical INs from xenografts at 12 weeks. Representative images of STICR-labelled GABA+ cells throughout the cortical plate, analysed by IHC. EGFP from STICR is in green, GABA in red, and DAPI in blue. Same cells from Fig. 3.4d. Arrows point to STICR-labelled GABA+ cells. Scale bar, 10 µm.

MATERIALS AND METHODS

STICR barcode design

STICR barcode fragment sequences were generated using the Barcode Generator script written by L. Comai and T. Howell (http://comailab. genomecenter.ucdavis.edu/index.php/Barcode_generator), with a sequence length of 15 base pairs (bp) and a minimum hamming distance of 5. Sequences containing the restriction-enzyme sites matching STICR's multicloning site (MCS) or homopolymer repeats longer than 4 bp were excluded. In total, 3 non-overlapping sets of 500 sequences meeting these design criteria were selected (Supplementary Table 3).

STICR library creation

STICR barcode libraries were created using a modified pSico lentiviral plasmid (Addgene, catalogue number 11578) in which the sequence between the central polypurine tract (cPPT)/central termination sequence (CTS) and 3′ long terminal repeat (LTR) was replaced with a DNA fragment containing the hybrid cytomegalovirus/chicken β-actin promoter (CAG) promoter, EGFP transgene, WPRE, multiple cloning site (MCS), and bGH polyadenylation signal. The MCS consisted of three adjacent pairs of restriction-enzyme sites (EcoRI–BamHI– NheI– XhoI) between which the STICR barcode fragments would be added. Each double-stranded DNA (dsDNA) STICR barcode fragment was synthesized as a pair of single-stranded DNA (ssDNA) oligonucleotides (Genewiz), annealed, and then pooled together with all the other barcode fragments belonging to its set. Each set of barcode fragments were added into the STICR MCS individually over three rounds of restriction-enzyme cloning. After each round of barcode fragment addition, a 'negative selection' digestion was performed using a restriction enzyme that targeted the stuffer sequence in the MCS that should have been replaced with a barcode fragment, so that undigested/barcoded molecules were removed from the library. Following each 'negative selection', the resulting STICR barcode plasmid library was amplified

by transformation into MegaX DH10B electrocompetent Escherichia coli (Thermo, C640003) and grown overnight on LB agar plates at 37 °C. In order to maintain even sequence distribution, transformed bacteria were plated at high density (roughly 75 million colonies per large-format plate), which helped to restrict colony size and make them grow more uniformly. In order to maximize the diversity of barcode sequences, we transformed enough barcoded plasmid to get at least ten times as many colonies as there were potential barcode sequences. Additionally, the STICR plasmid contained a sequencing primer site upstream of the STICR barcode. Each STICR library contained a 3-bp viral 'index' immediately downstream of the sequencing primer binding site. The viral index is unique to each library and allowed us to differentiate STICR barcodes from different libraries.

Generation of STICR NGS libraries for modelling

 To generate a STICR plasmid library for next-generation sequencing (NGS), we first digested 1 µg of STICR plasmid library with XhoI and then ligated a polymerase chain reaction (PCR) adaptor containing a unique molecular identifier (UMI) to this site (Supplementary Table 4). Ligation products were amplified by PCR using a Q5 Hot Start high fidelity 2× master mix (NEB, catalogue number M0494) using primers targeted to the STICR sequencing primer site and the adaptor sequence (Supplementary Table 4), using the following program: 1, 98 °C, 30 s; 2, 98 °C, 10 s; 3, 62 °C, 10 s; 4, 72 °C, 10 s; 5, repeat steps 2–4 15 times; 6, 72 °C, 2 min; 7, 4 °C, hold. Following PCR amplification, a 0.8–0.6× dual-sided size selection was performed using Ampure XP beads (Beckman Coulter, catalogue number A63881). The resulting libraries were sequenced to the depth of roughly 30 million reads. To generate a STICR lentiviral library for NGS, we carried out an RNA extraction from 1/20 of the total lentiviral prep using 300 µl Trizol (Thermo Fisher, catalogue number 15596026). After incubating for 5 min, 60 µl chloroform was added, incubated for 3 min, then centrifuged at 12,000g at 4 °C for 15 min. The aqueous phase was extracted and mixed with an equal volume of 100% ethanol, then loaded onto a Zymo

Direct-zol RNA Microprep (Zymo, R2061) column. The Zymo protocol was followed from there to bind and rinse the RNA, which was eluted in a final volume of 7 µl. Complementary DNA (cDNA) was generated from 5 µg of template RNA with the SuperScript IV kit (Thermo Fisher, catalogue number 18090010) using 1 µl of 2 µM STICR viral library reverse transcription (RT) primer (Supplementary Table 4). To add a UMI and a primer-binding handle to the individual cDNA molecules, one-cycle PCR was performed with cDNA and STICR viral library cDNA UMI primer (Supplementary Table 4) with 25 µl Q5 high fidelity 2× MasterMix (NEB, catalogue number M0492S), 2.5 µl primer (10 µM), 2.5 µL H2O, and 20 µl cDNA with the following program: 1, 98 °C, 40 s; 2, 62 °C, 20 s; 3, 72 °C, 2 min; 4, 4 °C, hold. Primers were removed with a left-sided 0.8× SPRISelect cleanup (Beckman Coulter, catalogue number B23318). Finally, cDNA was amplified using the same methods as for the lentiviral library above, and the library was sequenced to a depth of roughly 100 million reads.

Diversity and collision modelling of STICR libraries

STICR barcode sequences were extracted from fastq files with custom scripts that removed PCR duplicate reads using the UMI (for a general description, see the section 'scRNA-seq analysis and STICR barcode analysis' below). As it is prohibitively expensive to sequence high-diversity libraries to saturation, we extrapolated the total number of unique STICR barcodes using the Preseq38 command lc_extrap using default settings. Together with the measured relative barcode abundances, we used the extrapolated STICR barcode library size to model barcode collisions with the R (v4.0.1) programming language. Using base R functions, we simulated the labelling of a starting population of cells with a range of sizes from 101 to 106 and repeated each simulation 20,000 times. We then quantified the mean number of unique barcodes chosen for each starting cell population size. The difference between the starting cell population size and the number of unique barcodes present represented the number of collisions that happened at that population size.

Cell lines

We used NIH/3T3 cells (ATCC) and Lenti-X HEK293T cells (Takara Bio) here; we did not test them for mycoplasma or authenticate them.

Lentivirus production

STICR lentivirus was produced using a third-generation lentivirus packaging system: pMDLg/pRRE (Addgene, catalogue number 12251), pRSV-Rev (Addgene, 12253) and VSVG envelope (Addgene, 12259). Plasmids were transfected into Lenti-X HEK293T cells (Takara Bio, 632180) using jetPRIME (Polyplus, 114-15). In order to improve the viral titre, which is reduced in part because of the reverse orientation of the STICR EGFP-barcode transcript relative to the external lentiviral promoter, we also co-transfected pcDNA3.1 puro Nodamura B2 plasmid (Addgene, 17228) along with the other plasmids. Lenti-X 293T cells were grown and transfected in Dulbecco's modified Eagle's medium (DMEM) (Fisher, MT10017CV) supplemented with 10% fetal bovine serum (FBS) (Hyclone, SH30071.03) and 1% penicillin streptomycin (Fisher, 15070063). Twenty-four hours after transfection, media was replaced with Ultraculture media (Lonza, BE12-725F) supplemented with sodium pyruvate (final concentration 0.11 mg ml−1, Sigma P2256-25G) and sodium butyrate (0.005 M final molarity, Sigma B5887-1G) and penicillin streptomycin. Seventy-two hours after transfection, media were collected, passed through a 0.45 µM filter (Corning, 431220), and then ultracentrifuged at 22,000g for 2 h. Pellets were resuspended in 100 µl of sterile phosphate-buffered saline (PBS) (Thermo, 14190250) overnight at 4 °C and then aliquoted and stored at −80 °C.

Tissue procurement and STICR transduction

Deidentified tissue samples were collected with previous patient consent in strict observance of the legal and institutional ethical regulations. Protocols were approved by the Human Gamete,

94

Embryo, and Stem Cell Research Committee (Institutional Review Board) at the University of California, San Francisco. In order to visualize tissue for microdissection, samples were embedded in 3% low-melting-point agarose (Fisher, catalogue number BP165-25) and then cut into 300-µm sections perpendicular to the ventricle on a Leica VT1200S vibrating blade microtome in oxygenated artificial cerebrospinal fluid containing 125 mM NaCl, 2.5 mM KCl, 1 mM MgCl2, 1 mM CaCl2 and 1.25 mM NaH2PO4. The germinal zone was then isolated by microdissection using a scalpel and fine forceps under a Leica MZ10F dissecting microscope. In order to dissociate cells into a single-cell suspension, microdissected tissue was incubated in 200 µl of 0.25% trypsin (Fisher, catalogue number 25200056) and 2,000 units ml−1 of DNase I (NEB, M0303) for 20 min at 37 °C and then gently mechanically triturated with a 1,000 µl pipetteman 10 times. We then added 800 µl of DMEM supplemented with 10% FBS to the sample to neutralize the papain, and removed the trypsin/ DNase/FBS solution by centrifuging the sample for 5 min at 300g. The sample was rinsed in Dulbecco's PBS (Thermo, 14190250) and then centrifuged again for 5 min at 300g to remove the rinse. Cells were then immediately resuspended in NES media (defined in ref. 39), supplemented with ROCK inhibitor (Tocris, 1254) at a final concentration of 10 µM to reduce cell death and 1% penicillin streptomycin (Fisher, 15070063), and then plated on a 24-well tissue culture dish coated with 0.01% poly-l-ornithine (Sigma, P4957), 5 µg ml−1 laminin (Invitrogen, 23017-015) and µg ml−1 fibronectin (Corning, 354008) at a density of 500,000 cells per well. STICR lenitivirus was added to culture media at a roughly 1:250 to 1:500 dilution so that approximately 30% of cells were infected. After 24 h, virus-containing medium was removed and replaced with fresh medium. Seventy-two hours after infection, cultures were dissociated using papain (Worthington, LK003163), and EGFP+ cells were isolated by FACS and then used in a barnyard assay or in vitro culture assay. In order to maximize cell viability in the in vivo mouse transplantation assay, EGFP+ cells were not FACS-isolated before transplantation, but instead the entire culture (containing both EGFP− and EGFP+ cells) was used. See below for descriptions of each assay. In an effort to label

95

subclones, the GW15 replicate-2 in vitro sample was initially infected with a STICR viral library derived from a fully barcoded STICR plasmid encoding a GFP-T2a-TVA transgene and a viral index 'E' and then plated on mouse astrocytes as above. This culture was infected two and four weeks later with EnvA-pseudotyped (Addgene, 74420) STICR libraries with viral index '1' (two weeks) and '3' (four weeks) with a 1:100 dilution of virus:medium. During analysis, we found that labelling with EnvA-pseudotyped libraries was poor, so STICR barcodes with a '1' or '3' index were not considered during analysis. Therefore, clonal analysis of GW15 replicate 2 was carried out using only the initial 'E' index STICR barcodes transduced six weeks before analysis, as for all the other in vitro libraries.

Barnyard experiment

To confirm that transcribed STICR barcodes can be accurately recovered using scRNA-seq, we carried out a 'barnyard experiment' in which we infected separate cultures of human cortical cells (GW18 sample) and mouse 3T3 cells (ATCC) with different STICR libraries. These libraries could be distinguished from each other by a constant sequence unique to each library ('viral index') (Fig. 3.1a). After three days, we dissociated cultures with papain and FACS-isolated EGFP+ cells. EGFP+ cells from both species were then mixed together and processed using a Chromium Single Cell 3' GEM, Library & Gel Bead Kit v3 (10X Genomics, PN-1000075). Following sequencing (see below), transcript libraries were aligned with Cell Ranger (version 3.0.2) to a hybrid mouse/ human genome and droplets were determined to be either a mouse cell, a human cell, or a multiplet. Quantification of recovered STICR viral index (see below) for mouse, human and multiplate droplets is included in Extended Data Fig. 3.1f.

In vitro culture assay

Long-term in vitro experiments were performed using an astrocyte co-culture system. Primary mouse cortical astrocytes were isolated from CD-1 mice at postnatal day 1 and cultured as

described40, but with additional subdissection to remove the subventricular zone surrounding the lateral ventricles. Astrocytes were plated at a density of 400,000 cells per 3.5 cm2 in 12-well cell-culture plates containing DMEM (Fisher, MT10017CV) supplemented with 10% FBS (Hyclone, SH30071.03) and 1% penicillin streptomycin (Fisher, 15070063) 3–5 days before the addition of human cells. For each experiment, roughly 1,000–2,000 STICR-labelled (EGFP+) cells were added to a 12-well plate already containing mouse cortical astrocytes and cultured in DMEM, 1% B-27 supplement (Invitrogen, 12587-010), 1% N-2 supplement (Invitrogen, 17501-048) and 1% penicillin streptomycin. Cell culture medium was half-changed every three to four days for six weeks. Cultures were then dissociated into single-cell solution using papain, and EGFP+ cells were isolated by FACS. Following FACS isolation, EGFP+ cells were concentrated by centrifugation (300g for 10 min), and prepared for scRNA-seq.

PTPRZ1-based FACS

For enrichment of progenitors before xenograft transplantation, we used an adaptation of the protocol in ref. 41 in order to isolate cells that express the cell-surface protein PTPRZ1, which is enriched on the surface of cortical progenitors15,42. Cortical tissue was dissociated to a single-cell suspension as above, then resuspended in a solution of 1% bovine serum albumin (BSA) (Sigma Aldrich, A7979-50ML) and 0.1% glucose in Hanks' balanced salt solution (HBSS) (Life Technologies, 14175-095) for staining. Cells were incubated with mouse anti-PTPRZ1 primary antibody (Santa Cruz Biotechnology, sc-33664) at 1:50 dilution for 20 min on ice, washed with HBSS/BSA/glucose, incubated in goat anti-mouse IgM 488 secondary antibody (Thermo Fisher Scientific, A-21042) at 1:500 dilution for 20 min on ice, washed, and resuspended in HBSS/BSA/glucose. PTPRZ1+ cells were then isolated via FACS, plated, labelled with STICR, and cultured as above. Representative FACS plots are shown in Extended Data Fig. 3.9a.

Xenograft transplantation assay

Mouse transplantation assays were performed in CB17. Cg-PrkdcscidLystbg-J/Crl mice (Envigo) at postnatal day 3–5. STICR-labelled cultures were dissociated with papain, centrifuged at 300g for 5 min, rinsed once with DPBS, and then resuspended in ice-cold L15 media (Fisher, 11-415-064) with 180 Kunitz per millilitre of DNase (Fisher, 50-100-3290). Following anaesthetization, 100 nl of cell mixture (roughly 40,000 to 80,000 cells) was injected through a bevelled glass needle using a stereotactic rig into L −1, A 2.5, D −0.8 mm from lamda. At least five mice were injected with cells from each human specimen. Both male and female mice were used. In order to minimize clumping of xenograft cells, 20 mM EGTA (Sigma, E4378) was added to the cell mixture for GW17 replicate-1 and replicate-2 samples as well as GW15 replicate 3. After 6 weeks, mice were euthanized and one brain from each set of transplantations was cut into 1-mm coronal sections using a brain mould (Stoelting, 51386). EGFP+ regions of the cortex were dissected from slices using a fluorescent dissecting scope and then dissociated into a single-cell solution using papain. EGFP+ cells were then isolated by FACS, concentrated by centrifugation (300g for 10 min), and prepared for scRNA-seq. Mice were housed in a barrier facility with a 12-h light/12-h dark cycle and temperature and humidity control (70 °F, 50% rack humidity). All protocols and procedures followed the guidelines of the Laboratory Animal Resource Center at the University of California, San Francisco, and were conducted with approval of the Institutional Animal Care and Use Committee (IACUC).

IHC of in vitro cultures and xenografts

At the experimental endpoint, transcardiac perfusion of sterile PBS followed by 4% PFA (Fisher, 50-980-487) was used to rinse and then fix the specimens. Brains were dissected out and drop-fixed overnight in 4% PFA at 4 °C. Brains used for cryosections (Fig. 3.3 and Extended Data Fig. 3.7) were then cryopreserved in a 1:1 solution of OCT (VWR, 25608- 930) and 30% sucrose, embedded in cryomolds containing the same solution, frozen on dry ice, and stored at

98

−80 °C. Brains were then cryosectioned at 12 µm onto glass slides and stored at −80 °C. Blocking and permeabilization were performed using a blocking solution consisting of 10% normal donkey serum, 1% Triton X-100 and 0.2% gelatin in PBS for 1 h. Primary and secondary antibodies were diluted and incubated in this same blocking solution. Cryosections were incubated with primary antibodies at 4 °C overnight, washed 3 × 10 min with washing buffer (0.1% Triton X-100 in PBS), incubated with secondary antibodies for 2 h at room temperature, and washed 3 × 10 min with washing buffer; coverslips (Azer Scientific, 1152460) were then mounted using Prolong Gold Antifade reagent (Invitrogen, P36930). Brains used for morphological analysis of GABAergic cells (Fig. 3.4 and Extended Data Figs. 9, 10) were fixed as above but stored in PBS at 4 °C. Brains were then sectioned on a Leica VT1000 S vibrating blade microtome to 40 µm and slices were stored in PBS. Slices were incubated in blocking solution composed of 10% normal donkey serum and 0.1% Triton-X in PBS at room temperature for 2 h and then incubated in primary antibodies diluted in blocking solution overnight at 4 °C, washed 5 × 30 min in 0.1% PBST, incubated in secondary antibodies in blocking solution overnight at 4 °C, washed 5 × 30 min in 0.1% PBST, mounted on glass slides, and coverslipped as above using ProLong Gold Antifade reagent. In vitro cultures comprising human cortical cells co-cultured with mouse astrocytes were prepared as described above, but without FACS isolation of EGFP+ STICR-labelled cells, and plated on eight-well chamber slides (Thermo Scientific, 154534PK). After 6 weeks, cultures were fixed with 4% PFA for 1 h at 4 °C. Cultures were then washed three times and stored in PBS. Immunohistochemistry was performed as described for the 40-µm mouse brain sections but with 10-min washes. Slides were coverslipped with ProLong Gold Antifade reagent. The antibodies used include: chicken anti-GFP (Aves, GFP-1020; 1:1,000), mouse anti-human nuclear antigen (Novus, NBP2-34342; 1:100), rabbit anti-GABA (Millipore Sigma, A2052-100ul; 1:250), rabbit anti-NEUROD2 (Abcam, ab104430; 1:500), guinea pig anti-DCS (Millipore Sigma, AB2253; 1:200), rabbit anti-GFAP (abcam, ab7260; 1:1,500), rabbit anti-SOX9 (Abcam, ab104430; 1:250) and mouse anti-OLIG2

(Millipore Sigma, MABN50; 1:200). Secondary antibodies used include AlexaFluor anti-chicken 488 (Jackson Immunoresearch 703-545-155; 1:500), anti-mouse 488 (ThermoFisher A-21042; 1:500), anti-rabbit 594 (ThermoFisher A-21207; 1:500), anti-guinea pig 647 (Jackson Immunoresearch 706-605-148; 1:500), anti-mouse IgG1 488 (ThermoFisher A-21121; 1:500), and anti-mouse IgG2a 647 (ThermoFisher A-21241; 1:500). Confocal imaging was performed using a Leica SP8 confocal microscope with either a 10× or a 20× air objective; a 2-µm optical z-step was used for all images. Images were processed using ImageJ/Fiji. For quantification of the major cell types in xenografted mice (Fig. 3.3 and Extended Data Fig. 3.7), tilescans of the transplanted region were z-projected with average intensity; channel intensity was normalized across images; and cells expressing EGFP and/or human nuclear antigen in addition to cell-type markers were counted manually using the CellCounter plugin for ImageJ/Fiji. GABA+ STICR-labelled cells in the cortex of host mice (Fig. 3.4) were quantified by imaging four 40-µm-thick brain slices derived from within 400 µm of the transplantation site of each animal. Two mice for each of the two GW17 specimens were analysed in total. GFP+ cells in the cortex were counted using the CellCounter plugin for ImageJ/Fiji. The relative laminar positions of GABA+ /GFP+ double-positive cells were measured using the Measure tool from ImageJ/Fiji to draw a line from the top of the corpus callosum straight up to the pial surface through the soma of each cell. The relative location of the cell's soma to the top of the corpus callosum was then divided by the total length of the line drawn from the corpus callosum to pial surface.

scRNA-seq library preparation

Preparation of scRNA-seq libraries was carried out using a Chromium Single Cell 3' GEM, Library & Gel Bead Kit v3 (10x Genomics, #PN1000075). Libraries were prepared according to the manufacturer's protocol.

STICR barcode library recovery

STICR barcodes were subamplified from each 10X cDNA library using the Q5 Hot Start High Fidelity 2× master mix (NEB, M0494). In brief, 10 µl of cDNA was used as template in a 50 µl PCR reaction containing STICR barcode read 1 and 2 primers (0.5 µM, each) targeting the region immediately upstream of the STICR viral index/barcode as well as the partial Illumina Read1 sequence added during cDNA library preparation (Supplementary Table 4), using the following program: 1, 98 °C, 30 s; 2, 98 °C, 10 s; 3, 62 °C, 20 s; 4, 72 °C, 10 s; 5, repeat steps 2–4 11 times; 6, 72 °C, 2 min; 7, 4 °C, hold. Following PCR amplification, a 0.8–0.6 dual-sided size selection was performed using Ampure XP beads (Beckman Coulter, #A63881).

Sequencing

10X transcriptomic libraries and STICR barcode libraries were sequenced using Illumina Novaseq 6000 or Illumina HIseq 4000 machines. 10x transcriptomic libraries were sequenced to the average depth of approximately 50,000 reads per cell (2,200 genes per cell). STICR barcode libraries from lineage-tracing experiments were sequenced to a depth of roughly 30 million reads per library. STICR plasmid libraries for diversity and collision modelling were sequenced to a depth of approximately 30 million reads per library, while a single STICR lentiviral library was sequenced to a depth of approximately 100 million reads.

scRNA-seq analysis and STICR barcode analysis

10X transcriptomic libraries were aligned to the hg38 genome using CellRanger (version 3.0.2). Aligned cell/transcript counts were processed by Seurat43 (version 3.2.0.9014 for initial in vitro cultures and version 4.0 for subsequent xenografted cultures and integration of xenograft data into in vitro data) to remove cells containing fewer than 1,000 genes, fewer than 1,250 transcripts, or a high abundance of mitochondrial reads (greater than 7% of total transcripts). Cells passing these thresholds were then processed with Cellbender44(version 0.1) in order to

identify and remove background reads and instances of barcode swapping. We then identified

and remove multiplets that arose during cell capture using Solo45(version 0.1). Additionally,

xenograft libraries were aligned to a chimaeric hg38/mm10 genome using CellRanger (version

3.0.2) in order to identify potential cross-species multiplets; cells identified as such were

removed from analysis. Libraries were integrated using Seurat's SCTransform and

FindIntegrationAnchors functions, to identify integration features. First, in vitro cultured STICR

experiments were integrated together. Subsequently, transcriptomic libraries form xenograft

experiments were integrated with in vitro cells using integration anchors identified from in vitro

cultures. Leiden cell clustering, pseudotime analysis and data visualizations (that is, creation of

UMAP images) were performed using Monocle346 (version 0.2.1.9). Pseudotime analysis of

cells of the inhibitory-neuron trajectory was performed by setting the root node within mitotic

inhibitory neuron clusters. IN.1, IN.2 and IN.3 trajectories were defined as major branches of the

principal graph that led to distinct sets of clusters. Within subclustered inhibitory neurons, cluster

17 appeared to be at the beginning of both IN.2 and IN.3 trajectory cells in pseudotime analysis

and was thus termed 'IN.early'. Differential gene expression analysis was conducted using

Seurat FindMarkers/FindAllMarkers functions. With the exception of cluster 34 from the in vitro

culture data (Fig. 3.1c), we identified marker-gene expression consistent with previously

described cell types. Cluster 34 expressed markers of multiple cell types and did not show

strong cluster correlation with reference data (Extended Data Fig. 3.2d). Thus, we refer to cells

in this cluster as 'Unknown' in Supplementary Table 1. Iterative subclustering of

excitatory neuron and EOMES+ IPC trajectory cells from in vitro STICR cultures revealed

'upper-like' (subclusters 3, 4, 7, 9), 'deep-like' (subclusters 1, 5), 'newborn excitatory neurons'

(subcluster 6), 'EOMES+ IPCs' (subcluster 2), and 'EN.other' (subcluster 8) subgroups based

on gene expression. The EN.other group did not show a strong correlation with excitatory

neurons, and thus cells in this category were removed from further clonal analysis. Within

xenografted excitatory-neuron and EOMES+ IPC trajectory cells, iterative subclustering

revealed 'excitatory neuron' and 'EOMES.IPC' subclusters. STICR barcode analysis was performed using custom scripts. First, BBMap (BBMap, Bushnell B; sourceforge.net/projects/bbmap/) was used to remove low-quality reads and then extract reads containing STICR barcode sequences. Then, BBMap was used to extract individual STICR barcode fragments that were then aligned to our predefined fragment reference sets using Bowtie47 (version 5.2.1), allowing for up to two mismatches per fragment. Aligned STICR barcodes were compiled into a file containing their corresponding 10x cell barcode and 10x UMI sequences using Awk. Finally, UMI-tools48 (version 0.5.1) was used to remove duplicate STICR barcode/cell barcode (CBC) reads by UMI, allowing for 1-bp mismatches in the UMI. STICR barcodes/CBC) pairings with at least five distinct UMIs were retained. Cells with a single STICR barcode meeting this criterion were retained for clonal analysis. Possible instances of STICR barcode superinfection (multiple STICR barcodes per starting progenitor) were identified by calculating jaccard similarity indexes of all STICR barcodes pairings found to co-occur within a single cell. Those pairings with a jaccard similarity index of 0.55 or greater that occurred in 10 cells or more were considered to be a valid superinfection clone and retained for clonal analysis. Cells that contained multiple STICR barcodes that were not determined to be valid superinfections were further analysed for the relative abundance of individual STICR barcodes. Cells that contained a 'dominant' STICR barcode with five or more times the number of barcode counts compared with the next most abundant STICR barcode, as determined by UMI, were retained and assigned that dominant barcode. Those cells that did not contain barcodes meeting these criteria were not considered for clonal analysis.

Validation of marker genes with Allen Brainspan data

Expression data (Z-score) for IN.2 and IN.3 marker genes were downloaded from the Allen Brainspan database (https://www.brainspan.org). Biological samples from the cortex, basal ganglia and rostral migratory stream/olfactory bulb were retained for further analysis.

Hierarchical clustering was performed using the R package pheatmap(v.1.0.12), with the ward.D clustering method. For visualization purposes, the data ranges were divided into 20 quantiles.

Transcriptional cluster correlation

Marker genes from each cluster in the Nowakowski primary human brain reference atlas14 were calculated using Seurat FindAllMarkers, restricting genes to those present in at least 25% of cells of that cluster. The top 100 marker genes by fold-expression for each cluster were then retained for further analysis. The average expression for each cluster's top marker genes was calculated for each cluster in the Nowakowski reference atlas as well as in each cluster within the STICR transcriptional datasets. We then calculated pairwise Pearson correlations between every reference and STICR cluster, and depicted the result using a heatmap.

Statistics and reproducibility

Images shown in figures are representative of results found in multiple replicates: Fig. 3b (two specimens each xenografted into three mice), Fig. 3.4b, c (two specimens each xenografted into two mice), Extended Data Fig. 3.4d, e (four specimens), and Extended Data Figs. 9, 10 (two specimens each xenografted into two mice).

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

scRNA-seq transcriptomic data and STICR barcode data are available at the database of Genotypes and Phenotypes (dbGAP; https://www. ncbi.nlm.nih.gov/gap/) under accession number phs002624.v1.p1; and at the Gene Expression Omnibus (GEO;

https://www.ncbi.nlm.nih.gov/ geo/) under accession number GSE187875. An interactive

browser of single-cell data can be found at the University of California, Santa Cruz (UCSC) cell

browser49(https://human-cortical-lineage.cells.ucsc.edu). Publicly available reference genomes

hg38 and mm10 were used for analysis. Source data are provided with this paper.

Code availability

Custom codes used in this study are available at the following GitHub repository:

https://github.com/NOW-Lab/STICR.

AUTHOR CONTRIBUTIONS

R.N.D conceived of the project, designed and generated the STICR barcode library, designed and conducted experiments, analysed the data, and wrote the manuscript. D.E.A. helped to design experiments, conducted experiments, analysed data, and helped to write the manuscript. M.G.K helped to design experiments, conducted experiments and helped to write the manuscript. W.R.M.L carried out xenograft transplantations. R.S.Z. helped to construct the STICR library. E.E.C. carried out PTPRZ1 FACS. A.A.-B. helped to supervise the research. T.J.N conceived of the project, helped to design experiments, assisted in the interpretation of data, and helped to write the manuscript.

COMPETING INTERESTS

A.A.-B. is co-founder and on the Scientific Advisory Board of Neurona Therapeutics.

REFERENCES

1. Cai, Y., Zhang, Y., Shen, Q., Rubenstein, J. L. & Yang, Z. A subpopulation of individual neural progenitors in the mammalian dorsal pallium generates both projection neurons and interneurons in vitro. *Stem Cells* **31**, (2013).

2. Kohwi, M. A subpopulation of olfactory bulb GABAergic interneurons is derived from Emx1- and Dlx5/6-expressing progenitors. *J. Neurosci.* **27**, (2007).

3. Herrera, D. G., Garcia-Verdugo, J. M. & Alvarez-Buylla, A. Adult-derived neural precursors transplanted into multiple regions in the adult brain. *Ann. Neurol.* **46**, (1999).

4. Bhaduri, A. An atlas of cortical arealization identifies dynamic molecular signatures. *Nature* **598**, (2021).

5. Bandler, R. C. et al. Single-cell delineation of lineage and genetic identity in the mouse brain. Naturehttps://doi.org/10.1038/s41586-021-04237-0 (2021).

6. López-Bendito, G. Chemokine signaling controls intracortical migration and final distribution of GABAergic interneurons. *J. Neurosci.* **28**, (2008).

7. Stuart, T. Comprehensive integration of single-cell data. *Cell* **177**, (2019).

8. Gorski, J. A. Cortical excitatory neurons and glia, but not GABAergic neurons, are produced in the Emx1-expressing lineage. *J. Neurosci.* **22**, (2002).

9. Bandler, R. C., Mayer, C. & Fishell, G. Cortical interneuron specification: the juncture of genes, time and geometry. *Curr. Opin. Neurobiol.* **42**, (2017).

10. Zhang, Y. Cortical neural stem cell lineage progression is regulated by extrinsic signaling molecule Sonic hedgehog. *Cell Rep.* **30**, (2020).

11. Stumm, R. K. CXCR4 regulates interneuron migration in the developing neocortex. *J. Neurosci.* **23**, (2003).

12. Bystron, I., Blakemore, C. & Rakic, P. Development of the human cerebral cortex: Boulder Committee revisited. *Nat. Rev. Neurosci.* **9**, (2008).

13. Mayer, C. Developmental diversification of cortical inhibitory interneurons. *Nature* **555**, (2018).

14. Alzu'bi, A. Distinct cortical and sub-cortical neurogenic domains for GABAergic interneuron precursor transcription factors NKX2.1, OLIG2 and COUP-TFII in early fetal human telencephalon. *Brain Struct. Funct.* **222**, (2017).

15. Anderson, S. A., Marín, O., Horn, C., Jennings, K. & Rubenstein, J. L. Distinct cortical migrations from the medial and lateral ganglionic eminences. *Development* **128**, (2001).

16. Guo, T. Dlx1/2 are central and essential components in the transcriptional code for generating olfactory bulb interneurons. *Cereb. Cortex* **29**, (2019).

17. Fuentealba, L. C. Embryonic origin of postnatal neural stem cells. *Cell* **161**, (2015).

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

19. Paredes, M. F. Extensive migration of young neurons into the infant human frontal lobe. *Science* **354**, (2016).

20. Crouch, E. E. & Doetsch, F. FACS isolation of endothelial cells and pericytes from mouse brain microregions. *Nat. Protocols* **13**, (2018).

21. Xu, Q., Tam, M. & Anderson, S. A. Fate mapping Nkx2.1-lineage cells in the mouse telencephalon. *J. Comp. Neurol.* **506**, (2008).

22. Fleming, S. J., Marioni, J. C. & Babadi, M. CellBender remove-background: a deep generative model for unsupervised removal of background noise from scRNA-seq datasets. bioRxivhttps://doi.org/10.1101/791699 (2019).

23. Lein, E. S. Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **445**, (2007).

24. Anderson, S. A., Eisenstat, D. D., Shi, L. & Rubenstein, J. L. Interneuron migration from basal forebrain to neocortex: dependence on Dlx genes. *Science* **278**, (1997).

25. Schildge, S., Bohrer, C., Beck, K. & Schachtrup, C. Isolation and culture of mouse cortical astrocytes. *J. Vis. Exp.* **71**, (2013).

26. Ludwig, L. S. Lineage tracing in humans enabled by mitochondrial mutations and single-cell genomics. *Cell* **176**, (2019).

27. Sussel, L., Marin, O., Kimura, S. & Rubenstein, J. L. Loss of Nkx2.1 homeobox gene function results in a ventral to dorsal molecular respecification within the basal telencephalon: evidence for a transformation of the pallidum into the striatum. *Development* **126**, (1999).

28. Lareau, C. A. Massively parallel single-cell mitochondrial DNA genotyping and chromatin profiling. *Nat. Biotechnol.* **39**, (2021).

29. Daley, T. & Smith, A. D. Modeling genome coverage in single-cell sequencing. *Bioinformatics* **30**, (2014).

30. Pollen, A. A. Molecular identity of human outer radial glia during cortical development. *Cell* **163**, (2015).

31. Hansen, D. V. Non-epithelial stem cells and cortical interneuron production in the human ganglionic eminences. *Nat. Neurosci.* **16**, (2013).

32. Letinic, K., Zoncu, R. & Rakic, P. Origin of GABAergic neurons in the human neocortex. *Nature* **417**, (2002).

33. Petanjek, Z., Berger, B. & Esclapez, M. Origins of cortical GABAergic neurons in the cynomolgus monkey. *Cereb. Cortex* **19**, (2009).

34. Bhaduri, A. Outer radial blia-like cancer stem cells contribute to heterogeneity of glioblastoma. *Cell Stem Cell* **26**, (2020).

35. Qiu, X. Reversed graph embedding resolves complex single-cell trajectories. *Nat. Methods* **14**, (2017).

36. Bernstein, N. J. Solo: doublet identification in single-cell RNA-seq via semi-supervised deep learning. *Cell Syst.* **11**, (2020).

37. Lodato, M. A. Somatic mutation in single human neurons tracks developmental and transcriptional history. *Science* **350**, (2015).

38. Azim, E., Jabaudon, D., Fame, R. M. & Macklis, J. D. SOX6 controls dorsal progenitor identity and interneuron diversity during neocortical development. *Nat. Neurosci.* **12**, (2009).

39. Nowakowski, T. J. Spatiotemporal gene expression trajectories reveal developmental hierarchies of the human cortex. *Science* **358**, (2017).

40. Speir, M. L. et al. UCSC cell browser: visualize your single-cell data. Bioinformaticshttps://doi.org/10.1093/bioinformatics/btab503 (2021).

41. Young, K. M., Fogarty, M., Kessaris, N. & Richardson, W. D. Subventricular zone stem cells are heterogeneous with respect to their embryonic origins and neurogenic fates in the adult olfactory bulb. *J. Neurosci.* **27**, (2007).

42. Batista-Brito, R. The cell-intrinsic requirement of Sox6 for cortical interneuron development. *Neuron* **63**, (2009).

43. Tripodi, M., Filosa, A., Armentano, M. & Studer, M. The COUP-TF nuclear receptors regulate cell migration in the mammalian basal forebrain. *Development* **131**, (2004).

44. Cao, J. The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**, (2019).

45. Alzu'bi, A. The transcription factors COUP-TFI and COUP-TFII have distinct roles in arealisation and GABAergic interneuron specification in the early human fetal telencephalon. *Cereb. Cortex* **27**, (2017).

46. Li, J. Transcription factors Sp8 and Sp9 coordinately regulate olfactory bulb interneuron development. *Cereb. Cortex* **28**, (2018).

47. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, (2009).

48. Smith, T., Heger, A. & Sudbery, I. UMI-tools: modeling sequencing errors in unique molecular identifiers to improve quantification accuracy. *Genome Res.* **27**, (2017).

49. Onorati, M. Zika virus disrupts phospho-TBK1 localization and mitosis in human neuroepithelial stem cells and radial blia. *Cell Rep.* **16**, (2016).

Publishing Agreement

It is the policy of the University to encourage open access and broad distribution of all theses, dissertations, and manuscripts. The Graduate Division will facilitate the distribution of UCSF theses, dissertations, and manuscripts to the UCSF Library for open access and distribution. UCSF will make such theses, dissertations, and manuscripts accessible to the public and will take reasonable steps to preserve these works in perpetuity.

I hereby grant the non-exclusive, perpetual right to The Regents of the University of California to reproduce, publicly display, distribute, preserve, and publish copies of my thesis, dissertation, or manuscript in any form or media, now existing or later derived, including access online for teaching, research, and public service purposes.

DocuSigned by:

Denise Allen

F6D45EEFD46F4DE...   Author Signature

5/25/2022

Date